THE IMPACT OF FEEDBACK INTERACTIONS ON

ONLINE LEARNER SATISFACTION AND ACHIEVEMENT

Erin A. Crisp

Submitted to the faculty of the School of Education
in partial fulfillment of the requirements
for the degree
Doctor of Education
in the Department of Instructional Systems Technology
Indiana University
July 2019

Accepted by the School of Education Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Education.

Doctoral Committee

(Curtis Bonk, PhD)

(Gamze Ozogul, PhD)

(Daniel Hickey, PhD)

April 18th, 2019

To Eric, whose super powers include
removing obstacles and reducing anxiety,
to my three sons for providing constant encouragement,
to Cassie and Rachel for unwavering friendship,
and to my IWU friends and colleagues.
I feel truly blessed.

Erin A. Crisp

THE IMPACT OF FEEDBACK INTERACTIONS ON ONLINE LEARNER SATISFACTION AND ACHIEVEMENT

The design of an online course--taught by different faculty members and accessed by groups of adult, online learners--has the potential to produce or prevent learner satisfaction and achievement. Feedback interactions between online learners and the instructor are central to successful learning; nevertheless, there are gaps in what is known about the frequency, distribution, timeliness, and content of feedback that can impact student achievement and satisfaction. This study uses analytics from a learning management system (LMS) along with analyses of the feedback instructors provide to students in comments on written assignments. It compares those data with achievement exam scores and learners' responses on an institutional end-of-course (EOC) survey. The intent is to determine whether there are relationships between easily accessible learning analytics data sources and student achievement or satisfaction. Findings indicate that while LMS data were not predictors of student achievement or satisfaction, there was evidence that individualized and content-specific comments from instructors to students had an impact on student achievement and satisfaction. As a result, instructional designers could target course improvements that facilitate the instructor's ability to provide meaningful, individualized feedback to students.

(Curtis Bonk, PhD)

(Gamze Ozogul, PhD)

(Daniel Hickey, PhD)

Table of Contents

The Impact of Feedback Interactions on Online Learner Satisfaction and Achievement

Chapter 1: Introduction

Statement of the Problem

For adult learners who are not engaged in the 18-22 year-old, residential college experience, it is clear that online or blended learning is the preferred modality in the United States (“Integrated Postsecondary Education Data System,” n.d.). Even as for-profit universities close their doors or pursue non-profit status, adult learners continue to seek out online and blended formats for learning as evidenced by the rapid increases in enrollments at primarily adult-serving institutions like Western Governors University, Grand Canyon University, and Southern New Hampshire University where enrollments have increased between 8 and 27% since 2012 (Ginder, Kelly-Reid, & Mann, 2017). An increase in the supply of learner-consumers choosing online postsecondary education (and using federal funds to do so) has led to increases in the demands placed on the providers of this education. Two accreditors in particular, the Higher Learning Commission (HLC) and the Western Association of Schools and Colleges (WASC), have faced public and federal scrutiny over their handling of postsecondary quality (Kreighbaum, 2018).

As more learners turn to computer-assisted methods of gaining skills, credentials, and degrees (Carey, 2016), the accessibility, affordability, and appeal of online learning continue to be major considerations of instructional designers and academic leaders (Honebein & Honebein, 2015). Providing high-quality, accessible, affordable, and enjoyable online learning opportunities for adults has never been more competitive or important. Furthermore, in the employment marketplace, the need for life-long access to learning opportunities continues to increase (Gallagher, 2017). Academic leaders in postsecondary institutions are squeezed amid the competing priorities of learners, accreditors, federal financial aid policymakers, and employers.

In response to this competitive landscape, it is more important than ever that instructional designers who design learning experiences for postsecondary, adult and online learners, are able to balance the three sides of the instructional design iron triangle: effectiveness, efficiency, and appeal (Honebein & Honebein, 2015). Effective designs provide evidence of learning outcome mastery, efficient designs lead to quick results, and appealing designs are enjoyed by the learner. The instructional designer experiences the same tensions as the institution as she considers the three competing stakeholders: accreditors, employers, and learners. Accreditors want to see that an institution’s academic programs are effective; they need evidence that students are demonstrating mastery of the stated learning outcomes. Employers prioritize investing in tangible skills, so they are most interested in efficiency, and similarly, employees need to be able to translate academic experiences into workplace skills immediately. Finally, the learning experience needs to appeal to learners; they need to be motivated to persist through completion and therefore, experience the benefits of a completed credential. Instructional designers, then, experience a microcosm of the same tensions being played out on a national stage.

One challenge of addressing the tensions between effectiveness, efficiency, and appeal (Honebein & Honebein, 2015) in online instructional design at most higher education institutions is that designers and leaders lack the kinds of metrics that would allow them to make proactive decisions related to curriculum and instruction. Often, by the time academic leaders participate in a program review where data are analyzed and decisions are made, many graduates have already left their programs and are now sharing information and perspectives with their colleagues and friends about both the strengths and the weaknesses of the instructional design of the courses in that program. Referrals continue to be the number one source of new students at the institution involved in the research, so it is imperative that learners leave having had an effective, efficient, and appealing experience. Using the near real-time data available through the analytics in the learning management system (LMS), it might be possible to gain proactive insight into a student's experiences. Are there online learner behaviors that could be observed as early warning signs of problems with the effectiveness of the instruction, the efficiency of the design, or the overall appeal of the courses? What are the key ingredients of learning, and could those be measured and monitored on a large scale?

In an effort to find efficiency without losing effectiveness or appeal, instructional design leaders are realizing the potential of technologies that deliver timely, relevant feedback. Feedback is a powerful construct in the design of effective instruction, so it seems logical that feedback delivery technology could be leveraged to: increase efficiency by delivering immediate feedback, improve quality by delivering accurate feedback, and maintain appeal by being user-friendly. Robust studies of the impact of user-implemented feedback delivery technologies, performed without conflicts of interest (without the vendor's funding) and measuring impact on student achievement, are not plentiful.

Studies produced since 2010 tend to view feedback as a process rather than merely a concept (Dawson et al., 2018). A new paradigm of education recognizes that learner-centered instruction (more active, more authentic, and more socially connected) often hinges on designing effective feedback processes (Kuh & O'Donnell, 2013; Prensky, 2016; Reigeluth, Beaty, & Myers, 2016). Yet, many questions remain surrounding the aspects of a feedback process that most impact learner satisfaction (appeal) and achievement (effectiveness). For example, is the ideal formula for feedback simply that it is timely and relevant? If that is the case, artificially intelligent feedback delivery systems should be universally more effective than humans in situations where it is possible to program accurate feedback delivery for 100% of a student's attempts. Are there instructional design situations where human feedback is more effective than programmed feedback? What are those situations? What is the ideal balance between human and non-human sources of feedback or between peer and instructor sources? How does the learner's overall feedback experience in a course (i.e., the combination of timeliness, individualization, relevance, credibility, etc.) affect either her satisfaction or her ability to attain the identified learning outcomes? Do most learners feel similarly positive or negative about different types of feedback? These questions have yet to be fully explored.

As instructional technologies that provide immediate, individualized feedback become more widely implemented, educators and instructional designers need to increase their understanding of the dimensions of feedback and the relationships between feedback delivery processes, learner achievement, and learner satisfaction. Because formative feedback is a critical, multi-dimensional element of learner-centered instruction, we need to know more about how the dimensions of feedback work together as a whole to impact a learner's achievement and satisfaction. Providing individualized feedback to learners, both informally in discussions and formally on papers or projects, is the most time consuming element of an instructor's role. Knowing more about the relationships between student achievement and the dimensions of a learner's feedback experience could lead to instructional designs that make the best use of every moment an instructor spends providing feedback.

Background and justification. In 2015, I was an instructional designer working to support online course development at a non-profit, private university serving primarily adult learners. Our college was tasked with implementing a new adaptive learning tool in English composition and college algebra courses. The adaptive learning technology made big promises to provide learners with immediate and accurate pre-programmed feedback for each of their attempts. Our team worked closely with the vendor's team in the design of the courses, following what we knew to be sound instructional design principles. We scaffolded concepts, provided next steps for learning in the feedback, used diagnostic assessment to create branching, individualized learning pathways for students, and balanced workload through the course modules to provide a consistent experience for the learner.

After launching the newly designed courses, my team and I quickly learned that our courses had a design flaw. Even though students greatly appreciated the timeliness of the automated feedback, the courses as a whole were lacking the human element. We had invested so much time and effort in programming the feedback delivered via the software application that we had neglected to consider the student's overall feedback experience in the course. In the end-of-course surveys, learners expressed a number of complaints about the new technology. They felt trapped in an endless cycle of doing more problems to pass a module. Many of them disagreed with the computer-assisted responses, so distrust for the source of the feedback grew. Several learners made comments that they were not even sure if a human existed on the other side of the course content. Furthermore, comments showed that some learners felt isolated, ignored, and abandoned by their human instructors.

My experience is not an isolated one. Recently, University of Central Florida (UCF) administrators experienced the same kind of negativity but on a much larger scale (Lieberman, 2018). UCF had redesigned many of their business courses to incorporate the same adaptive learning engine piloted by our team. Backlash from students was swift. More than 1,800 students signed a petition “criticizing the college’s recent shift” (Lieberman, 2018). Their primary complaints were very similar to what we heard in our initial experiments with adaptive learning technology implementation.

Taking a step back from the overall design of the course, we immediately recognized our problem. The programmed feedback seemed to be working well to accomplish cognitive growth goals (exam scores were equal or higher to the former design), but we had neglected a crucial design element--the human instructor. We had neglected the affective goals that are, perhaps, just as essential for effective learning to occur. This oversight prompted our team to develop a rubric to capture the design of a learner’s feedback experience in a course (Table 1).

Table 1
Assessing the Six Dimensions of Feedback

	1	2	3
1) Timeliness- How quickly is feedback provided to the learner?	On average, > 72 hours after the interaction	On average, 24-72 hours after each interaction	On average, immediately to 24 hours after each interaction
2) Frequency- How often is feedback received?	Feedback is provided after fewer than 50% of learner interactions.	Feedback is provided after 51-89% of learner interactions.	Feedback is provided after 90% of learner interactions.
3) Distribution- To what extent are interactions dispersed throughout the experience?	The only designed interaction occurs at one point (usually the end) of the learning experience. Other interactions are student initiated.	The designed interactions are massed around 2 or 3 points in time (midterm and final for example).	Interactions are equally dispersed throughout the learning experience so that learners are receiving feedback at regular and predictable intervals.
4) Source or Credibility- To what extent does the learner	> 80% of the feedback instances are from similarly knowledgeable peers	50-79% of feedback instances are from similarly	> 80% of feedback instances are from a highly qualified subject matter expert or a

trust the source of the feedback?	or sources lacking trust from learners.	knowledgeable peers.	source trusted by most learners.
5) Individualiza-tion- To what extent is the feedback connected to the learner's unique strengths, needs, or interests?	< 50% of feedback instances are uniquely connected to individual learner's strengths, needs, or interests.	50-79% of feedback instances are uniquely connected to individual learner's strengths, needs, or interests.	> 80% of feedback instances are uniquely connected to an individual learner's strengths, needs, or interests.
6) Content- To what extent is the content of the feedback useful for the learner?	> 80% of the feedback is either motivational or provides a simple knowledge of response.	50-79% of feedback instances are of the type described in level 1 and the remaining instances provide next steps for learners.	> 50% of feedback instances provide next steps for learners to either extend their learning or correct misconceptions.

Each row of the rubric (Table 1) measures one of six dimensions of feedback that we defined after a review of literature pertaining to learner-centered instructional designs (Crisp & Bonk, 2018). Our team decided that after developing a course, we would apply the admittedly untested rubric and predict what level could be attained in each of the six dimensions. This was the course design team's crude method of evaluating our instructional design work. Does our design provide ample opportunity for both *automated feedback* (scoring high in the timeliness, frequency, and distribution categories) and also for *human feedback* (scoring high in the credibility, content, and personalization categories)? Our initial design, the flawed design, would not have scored well on the human feedback dimensions of the rubric (Table 1) if we had used the rubric at the time, so using the rubric on future course designs could prevent future errors; however, the tool and all of its benchmarks remained untested. Even now, we do not know if courses that score high on our rubric actually produce any better learning gains for learners than courses that are evaluated as lower scoring. This study stems from a desire to better understand the relationships between feedback and achievement.

Additionally, we learned that teaching with adaptive learning technology embedded in the course design was different than teaching without it. Instructors would need a better understanding of the value that they as instructors bring to the student. For instructors who were accustomed to providing right/wrong feedback as their main source of interaction with online students, new strategies for engaging learners would need to be provided. Again, the six dimensions of feedback rubric (Table 1) was useful because faculty development personnel could use it to teach instructors about the high value of providing more personalized feedback in exchange for less frequent feedback because this function was replaced by the technology. Clearly though, more research is needed to understand the impact of different types of feedback interactions.

Purpose of the Research

In this research, the purpose is to determine the extent to which existing learning management system (LMS) and end-of-course survey (EOC) data could be useful metrics in the proactive measurement and monitoring of feedback interaction data in online courses from business degree programs. Findings will show correlations between LMS, EOC data, and standardized exam data.

First, the feedback interactions that occur in a learner's final three courses of his/her degree program are counted manually or captured using the LMS system analytics and categorized by four dimensions of feedback: (1) timeliness, (2) distribution, (3) frequency, and (4) individualized content. Next, the counts of the timeliness, distribution, frequency, and individualized content of learners' feedback interactions are compared with their achievement as measured by a standardized exam. These same dimensions are also compared with learners' end-of-course survey results to discover relationships between feedback and learner satisfaction. The four dimensions of feedback quantified in this study are: (1) timeliness of the feedback, (2) distribution of the feedback--to what extent is it evenly distributed across the weeks in a course, (3) the frequency--the number of feedback interactions that occur with both peers and the instructor, and (4) the number of feedback interactions that provide individualized and content-specific feedback.

Are there certain designs for feedback that are more highly predictive of achievement? If it could be shown that certain levels of LMS feedback interaction metrics are predictive of success on summative assessments, these data could be monitored as a measure of the effectiveness of the course design. If it is true that certain learner experiences (i.e., high achievement and high satisfaction) correlate with certain results, then leveraging learning analytics to monitor those dimensions could provide course and program evaluators with data to inform design improvement. The results could also inform instructional designers who seek to implement feedback delivery technologies (e.g., intelligent tutors, artificially intelligent adaptive learning tools, etc.).

In short, the purpose of this study is to discover relationships between adult learners' online experiences with feedback and their achievement and satisfaction. If it is found that certain feedback experiences are predictive of achievement and satisfaction, it may be possible, with additional study, to measure feedback experiences (holistic evaluation of several dimensions of feedback) as one proxy for a high quality learning experience.

The idea of using proxies for measurement is common in medical research. For example, a patient who cannot respond to a survey about quality of life would have a proxy respondent--a person who would represent their best interests and interpret the situation to approximate responses for the patient. In the social sciences, examples of researchers who use observable or tangible evidence as a rough approximation of the cognition or emotion that is occurring within the subject are often seen. For example, Szupnar, Moulton, and Schacter (2013) discuss the use of measures of retention as proxies for student attentiveness. Because it is not feasible to look inside the brains of a classroom of learners to measure the degree to which each one is paying attention during a 30-minute lecture, cognitive psychologists may use a measure of the degree to which the students retained information from the first ten minutes as a proxy for the construct--attentiveness (Szupnar, Moulton, & Schacter, 2013). Similarly, one of the purposes of this study is to understand whether observable data around feedback interactions could be used as a proxy--an incomplete but useful measurement--for quality in online education.

Many researchers have confirmed, through original research and meta-analyses, that feedback is an essential component of the learning process (Brown, Collins, & Duguid, 1991; Gibbs & Simpson, 2004; Hattie & Timperley, 2007; Jaehnig & Miller, 2007; Merrill, 2013). Yet, key questions remain about how and why feedback impacts learning since some research shows inconsistent or negative results for immediate feedback (Kluger & DeNisi, 1996). As educational technologies become easy to implement, widely available, inexpensive, and prolific, they could be incorporated into online courses in ways that fail to fulfill their intended purposes. Furthermore, to take full advantage of the learning analytics data that are now readily available through the learning management systems (LMS), it is essential to know which LMS data are worth collecting and providing to course and program leaders for the purposes of quality assurance and continuous improvement.

The Roots of Adaptive Learning

Dating back to the 1960s, psychologists asserted that time rather than fixed intelligence was the primary variable impacting student achievement (Bloom, 1968; Carroll & Spearett, 1963; Keller, 1974). The time that students are willing to spend focused and invested in learning something is directly impacted by that student's relationship with the instructor (Carroll & Spearett, 1963). Keller (1985) spoke at a graduation ceremony near the end of his life and announced that he had been "living in a dream" and that personalized systems of instruction (PSI) "are unrealistic and involve too many different people" (p. 4). Keller's discouragement is understandable. While Keller's PSI produced remarkable learning gains in controlled settings, scaling the design failed repeatedly during his lifetime.

I propose that perhaps Keller's mastery learning notebooks (which provided individualized, carefully scaffolded instruction, assessment, feedback, and practice for each learner) were not scalable at the time because, without a human connection in the learning process, students quickly lost motivation and perseverance. Perhaps small scale trials were successful because students could look up from their individualized learning paths into the eyes of a human researcher or teacher and feel encouraged to persevere. The online implementations of adaptive learning technologies today face the same struggles as the PSI or mastery learning approaches of the 1960s and 1970s. It is reasonable to hypothesize that if a notebook or software application is providing a learner with the majority of his/her feedback, the learner feels dissatisfied, abandons the work more quickly, and, therefore, learns less than if he/she experienced a human connection.

Despite the plethora of research on feedback, there is a notable gap related to the impact of feedback as a process (Dawson et al., 2018) on student satisfaction and achievement. In their meta-analyses of feedback research studies, Kluger and DeNisi (1996) conclude that only one third of the studies conducted (as of 1996) measured the impact of feedback on performance. More recently, Dawson et al. (2018) echo a very similar concern in their review of the feedback literature since 1999.

Audience. Instructional designers, instructors, and adaptive learning software providers all stand to gain from the results of this research. To avoid making the same mistakes as the mastery learning proponents of the 1970s, designers today could benefit from following a design framework for the feedback process that humanizes instruction in online courses. Before we adopt this feedback design framework, it is first necessary to understand more about the relationship between the dimensions of feedback and student satisfaction and achievement.

Definition of Terms

Feedback, as described in this research, is the procedure “used to tell a learner if an instructional response is right or wrong” (Kulhavy, 1977, p. 211). Furthermore, feedback provides “information about the correctness of the response” and extends or expands a learner’s knowledge state (Jaehnig & Miller, 2007, p. 220). To further operationalize the definition of feedback, researchers often refer in some way to three conditions for effective feedback, defined by Sadler as: (1) the learner’s understanding of the learning goal or objective, (2) comparison of the learner’s attempt with the defined standard, and (3) an opportunity for the learner to take action that closes the gap between the defined standard and the attempt (Sadler, 1989).

More recently, to build upon Sadler’s (1989) research on dialogic feedback loops, Carless (2018) provides evidence for the value of studying learners’ “feedback literacy” and “feedback spirals”--long-term investigations into individual students’ experiences receiving and applying feedback from a variety of sources over a period of years. Related to Carless’ study (2018), there are four dimensions of feedback measured in the present research. They are (1) timeliness, (2) frequency, (3) distribution, and (4) individualized content, defined in Table 2. All of the dimensions of feedback chosen for this study also appear in Carless' qualitative data as key elements worthy of consideration (2018). For example, in Carless' interview with learners, he identifies timeliness as a significant barrier. His learners reported that feedback often "comes back too late" for it to be of significant value in improving future efforts (p. 2). The participants in Carless' study articulated the value of frequent feedback that can be acted upon for future attempts (p. 3); this connects to the frequency dimension in this study.

Carless' students also wanted the content of instructor feedback to provide clear suggestions for improvement--elsewhere referred to as Feedback Mark 1 (Boud & Molloy, 2013). While Boud and Molly differentiate between Feedback Mark 1 as teacher-directed and relevant mostly for short term gains, and Feedback Mark 2 as increasing student self-regulation that leads to long term gains, Carless seems to indicate from his longitudinal study that college students can extract both Mark 1 and 2 feedback from the same interaction. In an interview with one student, she describes how one instructor's feedback to her concerning coverage of breadth versus depth in a paper not only helped her improve the attempt in question, but it also impacted her demonstrations of learning in subsequent years (Carless, 2018, p. 4). The quality dimension of the feedback that is consistent across both Feedback Mark 1 and Mark 2 is that, in both frameworks, the content of the feedback provides next steps for success and is specific to the learner--individualized (Boud & Molly, 2013). Data for both of these dimensions of feedback--content and individualization--are collected in this study (Table 2).

The independent variables in this study are listed in Table 2 along with definitions for each variable. These variables are referred to as "dimensions" of feedback because of Dr. David Dirlam's work around defining the dimensions of a construct and then measuring each dimension (Dirlam, 2017). *Achievement* and *satisfaction* are both referred to as dependent variables in this study and will be described in later chapters. The measure of *achievement* used in this study is students' scores on a standardized exam delivered through an outside vendor--Peregrine. Validity and reliability for this instrument are discussed in the *instruments* section. The variable *satisfaction* refers to students' satisfaction with their courses. Satisfaction is measured using a survey with questions pertaining solely to the curriculum and instruction in the course. This instrument is also later described in greater detail.

Table 2

Four Dimensions of Feedback

Dimension	Description
Timeliness	Responses on submitted assignments are provided within 7 days of submission, and learners' questions are answered in 24-48 hours. The assumption is that timely feedback is best.
Frequency	The number of feedback interactions between the learner and either peers or the instructor. The assumption is that higher frequencies of feedback are best.
Distribution	The extent to which feedback interactions are dispersed across the weeks in a course. The assumption is that distributed interactions are best.
Individualized Content-specific	The feedback is specific to the individual learner's goals, strengths, needs, or questions; it either provides the learner with next steps to correct misunderstandings or prompts the learner to extend their learning in some new and novel way. The assumption is that individualized and content-specific feedback is best.

The dimensions described in Table 2 are described from the learner's perspective as opposed to the instructor's perspective. For example, I do not discuss the time it takes an instructor to provide feedback even though this is an important quantifiable variable within the feedback process. If findings suggest that learners who receive the most content-specific, individualized feedback learn the most, recommendations from the results of the research would pertain to instructional design adjustments that would afford more time to instructors at key time periods in the course design. Recommendations might include leveraging instructional technology for some assignments in the course in order to provide instructors with more time for individualized feedback on other assignments.

With a clear understanding of the dimensions of feedback analyzed in this study, and a summary of contemporary questions surrounding feedback in instructional design, it is beneficial to also understand the knowledge claims and evidence in historical and current research.

Chapter 2: Literature Review

Current State of Knowledge on Feedback

Several domains of research are in agreement that feedback is a construct that is central to learner-centered instructional design. From a neuroscientific perspective, Harvard neuroscientist and educational researcher Todd Rose states clearly that learning cannot happen without feedback (CAST, 2011). From a psychological perspective, Kruger and Dunning (1999), conclude that without feedback, learners, and humans in general, fail to recognize their own misconceptions and therefore fail to progress. Similarly, learning theorists assert that feedback is a critical component of all learning theories (Smith & Dillon, 1999). Numerous studies over a century of research have provided evidence that when learners receive feedback, it improves their performance (Kulik, Kulik, & Bangert-Drowns, 1990; Pennebaker, Gosling, & Ferrell, 2013).

Beyond findings around the importance of feedback, advances in educational technology have prompted the investigation of new delivery mechanisms for feedback. These delivery mechanisms include artificially intelligent tutoring systems (Koedinger & Corbett, 2006), adaptive learning technology (Murray & Perez, 2015), and peer feedback delivery mechanisms (Xie, 2013). Furthermore, a new paradigm of education recognizes that learner-centered instruction (more active, more authentic, and more constructivist) is a worthy 21st century goal (Reigeluth, Beatty, & Myers, 2016).

Despite an idea that is so widely accepted, scant empirical research attempts to test the impact of certain feedback processes on learner achievement (Dawson et al., 2018). The proliferation of educational technologies embedded throughout online instruction has prompted some to call for a renewed emphasis on humanizing online learning (Reupert, Maybery, Patrick,

& Chittleborough, 2009). Finding the appropriate balance between the human and technological aspects of online education is challenging. While it makes sense to leverage technology to do the things that it does best so that humans are able to do the things we do best (Prensky, 2016), there is much to be learned about how to achieve this balance in instructional design for online learning.

As instructional technologies that provide immediate, individualized feedback become more widely implemented (Adkins, 2016), educators and instructional designers need to more fully understand the varied dimensions of feedback (Crisp & Bonk, 2018). There is also some benefit in gaining a better understanding of the impact of different feedback designs on learner achievement and satisfaction. For example, if an online course provides high frequency feedback but from an untrusted source, how does that compare with an online course that provides lower frequency feedback but from a highly trusted source? What situational factors should be considered by the instructional designer to choose one design over another? Amid these emerging questions, a century of feedback research does reveal some consistent findings.

Feedback improves performance. First, feedback does appear to improve performance in many, but not all, instructional situations. Several meta-analyses of feedback research have confirmed that formative feedback (assessment for learning or formative assessment) improves performance (Black & Wiliam, 1998; Wiliam, 2011). Learners who receive, understand, and apply formative feedback show greater achievement gains than those who do not receive formative feedback (Black & Wiliam, 1998; Gibbs & Simpson, 2004; Spiller, 2013; Wiliam, 2011). This thread of feedback research has become known as research on feedback as a process (Dawson et al., 2018).

Alternatively, other researchers, in their meta-analyses, have asserted that feedback interventions produce highly inconsistent results (Kluger & DeNisi, 1996). These researchers tend to define feedback as a construct, and they answer questions around the timeliness and relevance of feedback. For example, in one quasi-experimental study, educators hypothesized that learning would be most effective if students had knowledge of results (KR) immediately after responding to a test question (Pressey, 1950). That study involved the use of punchboard self-scoring devices with test papers covering the holes. If a learner tried to punch in an incorrect response with his pencil, no hole was punched, so the learner must try again until the correct hole was punched in the paper (Pressey, 1950). Through a series of experiments, Pressey concluded that this immediate feedback does increase learners' ability to memorize information (Pressey, 1950), but very little attention was given to transfer or generalizability of the knowledge or skill gained. Also, when experimental results were inconsistent (as they were with a series of Russian words that were taught to English speakers), no explanation was offered by Pressey (1950). Because of Pressey's study, Kluger and DeNisi (1996) rightly question why the knowledge of result (KR) feedback did not improve subsequent performance in some of the instances and question why this inconsistency was never addressed by Pressey.

The Pressey study is significant because it offers design similarities to the educational technology provided to students today. In Pressey's study, there were few negative consequence described for "guessing" behaviors since students could use their pencils to punch as many holes as they wanted until they found the correct answers. This is similar to the eLearning tutorials that remain popular in online learning. During formative assessment practice sessions, learners are free to click any response until they find the right one and are then able to continue with the programmed tutorial. Just as Pressey found inconsistent results when there was evidence of “guessing,” similar findings exist more recently using digital versions of the same experiment (Rha, 1988). It is not surprising that feedback has little impact on learning when there is evidence of “guessing” behavior among learners.

Similarly, “hunt and peck” behaviors also confound results in feedback research. In studies where learners are able to browse a reading passage, for example, to pick out the answers to the multiple-choice questions, feedback interventions produced inconsistent results (Rha, 1988). It appears that when learners are not thoughtful about their choices during formative assessment, the intended results are not achieved.

Several other concerns related to feedback research have also been raised. Feedback is often not used by learners in subsequent attempts (Hounsell, 1987), or if it is read by the learners, it is not understood by them (Lea & Street, 1998). Different evaluators apply rubrics differently, and in some studies, feedback has had negative effects on student achievement (Bangert-Drowns, Kulik, Kulik, & Morgan, 1991). In their frequently cited meta-analysis of feedback intervention (FI) research, Kluger and DeNisi (1996) question whether it is the feedback that produces the effect or whether it might be the climate that the agent created for the receiver of the feedback that produces positive effects--a climate of high expectations and accountability. Their meta-analysis reveals that two thirds of all empirical FI literature does not shed light on the question of FI effects on performance (Kluger & DeNisi, 1996).

Despite these reservations, overall consensus is clear that there is a connection between feedback and learning. Even Kluger (1993), who describes several confounding variables in a century’s worth of feedback research, eventually concludes that the existing empirical data does in fact point to the effectiveness of feedback interventions in certain situations. These findings are echoed in several other significant meta-analyses.

What do we know about feedback? First, when learners receive feedback, it improves their performance (Kulik et al., 1990; Pennebaker, Gosling, & Ferrell, 2013). Next, we know that providing learners with feedback that reveals misconceptions and provides next steps for learning results in learning gains (Kulik et al., 1990). Finally, instructional strategies that incorporate formative feedback are among those with the highest effect sizes (Fraser, Walberg, Welch, & Hattie, 1987; Gibbs & Simpson, 2004; Hattie, 2015; Jaehnig & Miller, 2007).

A search of the “What Works Clearinghouse” reveals 22 studies in which students are presented with variations of the three conditions: (1) no feedback or correct response only (this is also reported as KR- knowledge of results); (2) verification feedback: the correct response along with the respondent’s choice and the original question (KCR- knowledge of correct response); or (3) elaboration feedback: verification feedback plus rationale or recommended next steps for the learner (What Works Clearinghouse, 2016). With only a few exceptions, the elaboration feedback condition produces the greatest learning gains in multiple subject areas and with a variety of age categories (Jaehnig & Miller, 2007; Marsh, Lozito, Umanath, Bjork, & Bjork, 2012). When undesirable behaviors like “guessing” or “hunt and peck” are minimized, research that reveals a positive correlation between feedback and learning outweighs research that reveals inconsistent, negative, or no effect. When meaningful feedback is provided in a timely manner on work that has value to the learner, it has been shown to improve performance over students who experienced minimal or no feedback.

Formative feedback is central to postsecondary learner satisfaction. Not only does feedback often lead to greater learning gains, feedback also leads to learner satisfaction. The perceptions of college learners are significantly impacted by the degree to which they feel that the feedback they receive is valuable. In student satisfaction surveys, the most criticized aspect of postsecondary education is a lack of timely, relevant feedback (Boud & Molloy, 2013; Nicol, Thomson, & Breslin, 2014). Learners, and online learners especially, consistently rank the following three items as most valuable for a high-quality learning experience: (1) faculty responsiveness, (2) quality instruction, and (3) timely feedback (Herbert, 2006).

Interestingly, there is evidence that instructors and students may perceive the ingredients for satisfaction differently. Instructors believe that students are most satisfied by effectively communicated course content that illustrates clear expectations coupled with specific and timely feedback on performance (Dennen, Darabi, & Smith, 2007). In short, instructors emphasize learner-content interaction as the key predictor of satisfaction (Moore, 1989). Conversely, students tend to place more emphasis on the importance of interpersonal communication and being treated as individuals (Dennen, Darabi, & Smith, 2007). Students tend to emphasize learner-instructor interactions followed by learner-content and finally learner-learner interaction as central to their satisfaction (Moore, 1989). Even if one agrees that feedback does positively impact satisfaction, is the impact a result of the fact that it is interaction with the instructor or is there something inherently satisfying about feedback itself regardless of the source? Distance educators have described their perspectives on the value of different types of interaction (synchronous, asynchronous, self, learner, and instructor) in a Delphi study and articulate the importance of the learner-self interaction that seems to happen naturally in the learning process, but they prioritize learner-instructor and learner-learner interactions as most valuable (Soo & Bonk, 1998).

If it is true that feedback alone leads to satisfied learners, then fully automated eLearning courses where every response or submission from the learner results in elaborated feedback from an artificially intelligent tutor should provide the highest of all satisfaction ratings. One review of existing literature on this topic finds that there are few studies that examine the relationship between learner satisfaction and eLearning (Mohammadi, 2015). However, from the limited research that does exist, Mohammadi concludes that human support and feedback are of “immense importance” in an eLearning environment (2015, p. 371). Mohammadi’s review reveals the importance of training instructors to establish a welcoming, friendly online environment where users are encouraged to actively support one another in the midst of using eLearning to interact with content (Mohammadi, 2015). Timely and frequent feedback from an automated source, then, is likely to provide learner-content interactions, a high contributor to student satisfaction, but feedback from and interaction with a trusted human source is equally likely to produce higher learner satisfaction (Mohammadi, 2015).

Similar to Mohammadi’s discussion of establishing a welcoming online environment, other research illuminates the value of humanizing online instruction or improving social presence (Reupert, Maybery, Patrick, & Chittleborough, 2009). Qualitative research involving interviews of faculty members who teach both online and face-to-face reveals that in the absence of face-to-face instruction, with ubiquitous use of online discussion forums, feedback becomes even more central to a learner’s satisfaction. Faculty indicate that prompt feedback is important to students because it helps students see instructors as “more caring and more human if they receive prompt responses to emails and assignments” (McGuire, 2016, p.68). Faculty interviewed for this study indicate that the impact of the feedback intervention is not that it functions to correct cognitive misunderstandings or reinforce correct behaviors quickly. The impact of feedback on student satisfaction for online learners is, potentially, that it connects them to the human, social presence in the online course space (McGuire, 2016). This assertion begs the question, then, that if that feedback intervention is replaced by an artificial intelligent coach, to what extent would it impact student satisfaction or achievement?

To conclude, there is ample evidence that students both online and face-to-face welcome formative feedback. Students and instructors both acknowledge that satisfaction in online learning is largely dependent on interactions with one another, and feedback delivery is a logical vehicle for interacting. It is unclear from the research how much of a student's interaction-dependent satisfaction is attached to the humanity of the source and how much of the satisfaction is derived from gaining insight into one's own learning progress--a function that could be replicated by artificial intelligence.

Theoretical Framework

Feedback is not reinforcement. We know that feedback is important, but additionally, psychologists have asked how feedback works. Psychologically, what makes feedback effective? Dating back to the early 1900s, Thorndike's Law of Effect dictates that any behavior followed by favorable consequences is repeated while behaviors followed by negative consequences are abandoned (Thorndike, 1913). There was a time when educational researchers subscribed to this theory, or other similarly behaviorist theories, as a method of explaining how learning occurs (Johnson, 2014). However, considerable opposition arose when it became apparent in practice that reinforcement in the form of positive feedback (correct responses), grades, and rewards do not consistently increase the likelihood of correct responses in the future (Anderson, 1979; Kulhavy, 1977; Moore & Smith, 1964; Pysh, Blank, & Lambert, 1969). Furthermore, when feedback interventions are implemented from a purely behaviorist theoretical framework, learned helplessness can result (Mikulincer, 1994). Learned helplessness occurs when learners perceive that the results of their efforts are outside of their control. They feel that something other than their own effort will ultimately decide their success or failure. Learned helplessness is closely related to test anxiety in children (Fincham, Hokoda, & Sanders, 1989). In my own experience with adult learners, they too express feeling trapped, isolated, and anxious in an unending cycle of feedback/try again when automated feedback interaction designs lack a human element.

Conversely, when instructional designers structure feedback interactions from a constructivist and cognitivist theoretical framework, they acknowledge the power of anxiety, helplessness, and motivation in the design of the experience. Instructional designers may seek to mitigate negative emotions by asking human providers of feedback to do the things that cannot be replicated through artificial intelligence. This includes providing encouragement and connecting personally with learners.

As instructional technologies that provide immediate feedback become more widely implemented, educators and designers should consider a constructivist and cognitivist theoretical framework for their learner-centered instructional designs (Karagiorgi & Symeou, 2005). Many implementations of instructional technology focus on the stimulus-response (behavioristic or associationist) features of feedback delivery without attending to the learner as an individual with thoughts, feelings, perceptions, and motivations (constructivism).

In defense of a behaviorist approach, one could argue that perhaps grades and rewards are inadequate positive reinforcement for students and that perhaps this is why it appears that some feedback is ineffective. However, money is one of the most universally, well-documented reinforcements known, and yet the link between payment and satisfaction is weak (Judge, Piccolo, Podsakoff, Shaw, & Rich, 2010). Satisfaction and motivation are closely related emotions, and in a large meta-analysis of studies related to the correlation between pay level and satisfaction, researchers found that “pay level is only marginally related to satisfaction” (Judge et al, 2010). People find positive emotion, satisfaction, and motivation in meaningful contributions, effective leadership, enjoyment of colleagues, and many other social and emotional factors (Anderson, 2001). This research from the business arena provides further evidence that motivation and satisfaction--key ingredients in achievement and learning--are more complex than a simple stimulus-response behavior can explain.

Further evidence of the inadequacy of behaviorism to provide a rationale for feedback effectiveness is revealed in Mazzolini and Maddison's 2003 study. Their investigation into the ideal frequency for feedback interactions revealed that "increased instructor posting did not result in increased student participation" in all cases (as cited in Dennen, Darabi, & Smith, 2007, p. 68). Researchers concluded that, "there seems to be a threshold at which an instructor's heavy-handed or overwhelming amount of communication inhibits or discourages learner communication and participation" (Dennen et al., 2007, p. 75). If this finding is generalizable to other contexts, it could follow that over-use of automatically delivered feedback from instructional technology could have a similar negative effect. If behavioral reinforcement alone explains improvement, reaching a threshold should produce a leveling-off of satisfaction but not a decline in communication, participation, learning, and satisfaction. Constructivist and cognitivist theories provide much better explanations for why feedback improves learning.

Constructivism and cognitivism to explain effective feedback implementation. This research study assumes components of both a cognitivist and constructivist epistemology to be foundational for an understanding of feedback. The research question being posed assumes that a stimulus/response, behaviorist theoretical foundation is insufficient to explain the relationship between feedback and learning. It also assumes that an *information processing* model explains some but not all dimensions of the learning experience (Johnson, 2014). The shift away from thinking about feedback as a transmission has been appearing in learning research since the mid-1990s (Barr & Tagg, 1995), yet it is only now impacting research on feedback as a dialogic, socially constructed construct (Nicol, Thomson, & Breslin, 2014).

Constructivists hold that meaning is derived from socially-constructed, connected experiences (Jonassen, 1991). These created meanings are as unique and unpredictable as the individuals themselves; nevertheless, consideration for a wide variety of factors can increase the probability of certain results as is evident in instructional designs that use socially-constructed activity theory frameworks (Amory, 2010). Constructivism stands in contrast to the belief that behaviors are the result of stimulus and response and also to the belief that learning can be explained entirely through an information processing model (Vianna & Stetsenko, 2006). While some find constructivism to be too multi-faceted and ill-defined to be of use in instructional design (Mayer, 2004), others contend that learning is a complex process and is, therefore, most accurately understood through a multi-faceted theoretical framework like constructivism (Vianna & Stetsenko, 2006).

Constructivism describes that the mind of the learner filters input from the world to produce its own unique reality (Jonassen, 1991). Unlike information-processing theory (Simon, 1995) or behaviorist theories, constructivism would not subscribe to the idea that units of knowledge can be mapped into a schema of branching skills and subskills to be experienced similarly by a majority of learners (Johnson, 2014). Herbert Simon, award-winning, information-processing theory researcher, asserted that psychologists of the future must work to reconnect the science of cognition with the study of affect and motivation (Simon, 1995). Humans are much more likely to create meaning as opposed to acquiring it (Jonassen, 1991). According to both Piagetian and Vygotskian streams of constructivism, the knowledge contained in one’s mind emerges over time and is constantly open to change (Vianna & Stetsenko, 2006). Piaget and Vygotsky diverge, however, when it comes to a view of designing instruction (Vianna & Stetsenko, 2006) with Vygotsky's (and those who have followed him) theories most closely aligning to the assumptions that not all types of instruction are equally conducive to learner development (1978). Cognitivism and constructivism are connected as suggested by Piaget's work on cognitive constructivism (1970). The connection between cognitivism and constructivism diverges from pure or radical constructivism that opposes direct teaching or the correction of errors (Johnson, 2004).

In the 1950s and 1960s, behaviorally designed studies of feedback were plentiful. Cognitive constructivism, however, provides a better framework for why and how feedback works. For example, in a recent Australian study where 4,500 college learners and faculty responded to an open-ended response survey, 90% of learners and 89% of instructors who responded indicated that the primary purpose of feedback is to improve performance (Dawson et al., 2018). Other purposes included affective encouragement, identification of strengths and weaknesses, and justification for a grade. Furthermore, Dawson et al. (2018), describe that feedback is a dialogic process between students and instructors. If feedback is not acted upon by the learner, it is not truly feedback. This thinking aligns with Hattie and Timperly's (2007) recommendation that feedback should focus on improving self-regulation--a foundational element of both Piagetian and Vygotskyian constructivism. Using stimulus/response theories to explain why feedback works does little to address context, learner self-regulation, or socio-cultural factors.

A Cognitivist/Constructivist Framework for Feedback

Based on both the “feedback as a concept” and “feedback as a process” research that has been presented, I propose the following as a cognitive constructivist framework for defining five dimensions of feedback and describing proficiency in each dimension. The underlying purpose for this framework is to present a method that could measure feedback as a proxy for quality in online learning. For this reason, other frameworks that have been developed, such as the sustainable feedback model (Carless, Salter, Yang, & Lam, 2011) and Feedback Mark 2 (Boud & Molloy, 2013) do not serve this purpose. While other feedback models inform instructors in effective pedagogy, the purpose of this framework is to inform those tasked with designing and/or evaluating a learning experience.

I present each of the four dimensions of feedback that will be analyzed in this study along with research to support the assertion that there are more and less effective methods for implementing each dimension. If any singular dimension is implemented without regard for the others, the learner’s experience is diminished in some way.

The following four dimensions of feedback work in concert to produce the highest probability of success for the student.

1. Timeliness: Feedback should be delivered immediately or soon after the learner’s attempt, but care should be taken to ensure that the learner is making a thoughtful attempt. Guessing behaviors and “hunt and peck” behaviors where students scan a piece of text to find a word that they can use to fill in a blank negate the value of timely feedback (Brown, Collins, & Duguid, 1989; Gibbs & Simpson, 2004; Hattie, 2015; Jaehnig & Miller, 2007; Merrill, 2013). Timely feedback solidifies the learner’s cognition or addresses a misconception before the thoughts of the learner move too far astray to other topics (Brown, Collins, & Duguid, 1989; Merrill, 2013).

2. Frequency: The more frequently a learner receives feedback, the faster their learning progresses (Francom, 2016; Kulhavy, 1977; Reigeluth, Beaty, & Myers, 2016; Voorhees & Voorhees, 2016). However, frequent feedback should not depend on incorrect responses from the learner. When the instructor provides a question that serves to deepen the learner's engagement and produce dialogue, this is an effective method of providing feedback. When feedback is dialogic, the frequency of feedback provided is dependent on the quantity of attempts made by the learner, and a learner will make more of the right kinds of learning attempts when he/she is engaged in meaningful dialogue (Boud & Molloy, 2013; Carless et al., 2011). Frequent feedback gives the learner knowledge of her individual progress (Reigeulth et al., 2016).

3. Distribution: Practice attempts and feedback should be distributed rather than massed directly before a learner's summative assessment (Prensky, 2016; Reigeluth et al., 2016). While designing the distribution of feedback, the design can be universal (all students receive the same distribution), triggered (students who do x receive y), or requested by the student (Reigeluth et al., 2016). Distributed practice has been analyzed in textbook materials and, when present, contributes to greater student success (Surma, Vanhoyweghen, Camp, & Kirschner, 2018). Prensky (2016) recommends leveraging technology to do what it does best (automated feedback) and developing human capacity to interact in more complex dimensions.

4. Individualized Content: In an effective feedback experience, success is unique to every learner (Reigeluth et al., 2016). Feedback is customized to each learner in some way: skill development, learner interest, specific goals, prior outcomes, or perhaps it offers learners a choice for their preferred method of assessment and feedback (Reigeluth et al., 2016).

However, feedback can also come from the learner herself when instructional resources are introduced as a fixed point of comparison for the learner. In their research on competency-based education, for example, Voorhees and Voorhees (2016) find that learners are more successful when they use a rubric to self-assess than when the rubric is used solely by the instructor. After drafting a product, if a learner moves systematically through an analytic rubric, articulating the comparison of her work to the rubric criteria described, her own self-assessment of her work becomes individualized feedback for herself. As this occurs, she is making her justifications visible so that a trusted peer or instructor can identify misconceptions and offer additional individualized feedback that fits the learner's pre-existing rationale for her choices.

Similarly, personalized reflection can be an effective instructional strategy and can provide unique opportunities for the individual formative feedback of others (Watson & Watson, 2016). Advances in personalized learning meet the need for students to receive individualized feedback generated from an artificially intelligent program (Jarrett, 2013). Finally, maker-based research indicates that learning is more effective when learners are shown the value of the learning to the outside world and also when they are made aware of the necessity for change (McKay & Glazewski, 2016). Connecting each learner to the world and connecting learning to each individual will be different for every learner, thereby pointing to the need for individualization in a learner's overall feedback experience.

Finally, content can move the learner forward and solidify accurate understandings, or it can be motivational. In feedback research during the 1970s, the categories developed included three kinds of content. These were:

- Knowledge of results which is mostly motivational feedback (e.g., "Good job!" or "You did excellent work once again.").
- Verification feedback (e.g., "You selected B but C is the best answer").

- Elaborated feedback (e.g., “You should review Chapter 10 and consider the laws of motion.”) (Kulhavy, 1977).

Research from Raymond Kulhavy (1977) found that elaborated feedback is the most effective of the three, whereas knowledge of results feedback produces almost no effect (Kulhavy, 1977). Beyond these previously existing broad categories, more recent research provides more granular detail into the content dimension of feedback.

There have been several useful findings related to the content that is provided in high-quality feedback. First, the content of high-quality feedback is connected to the learner’s existing knowledge or skill (their knowledge state). Related to cognitivism, rather than the behaviorist tendency to see feedback as behavioral reinforcement, the content of feedback should help learners process and store their own thinking; in effect, it helps the learner think about their thinking (Reigeluth, Myers, & Lee, 2016). Second, the content of feedback should include emotional, social, and character development as well as input on the cognitive and physical knowledge or skill to be mastered (Reigeluth et al., 2016). Third, there are models for the content of feedback that can be followed to improve learner success. In one example from “social serious game” design, the content of feedback delivered to players is categorized as “question, information, hint or solution” and then delivered to the player/learner when certain conditions are met (Konert, Gobel, & Steinmetz, 2012, p. 2).

Other considerations related to the content of feedback include the finding that granular feedback is better suited for formative assessments, whereas broad feedback is more effective for summative competencies (Voorhees & Voorhees, 2016). According to Francom (2016), feedback related to the actual task accomplished by the learners rather than the topic of instruction tends to have greater relevance and effectiveness. He also notes that the content of feedback instances should range from simple to complex and then fade with independence (Francom, 2016). Among some of the other relevant findings, Watson and Watson (2016) suggest that mentoring is a method of identifying the strengths and interests of the learner so that the content of the feedback can be authentically connected. Moreover, in the maker-based instructional model, the content of feedback should help learners articulate a question that will guide their learning and prompt them to reflect and consider their own design thinking (McKay & Glazewski, 2016). Finally, the content of feedback in a learner's overall feedback experience should clearly reflect the purpose of the learning.

Defining the dimensions of feedback that can be measured provides a method for instructional designers to evaluate the effectiveness of the implementation of instructional designs such as by using a rubric. Without the use of a multi-dimensional feedback rubric, the evaluation of an instructional design may ignore critical elements. For example, evaluating the timeliness and frequency of the feedback students receive after the implementation of a new adaptive learning tool may reveal a positive impact. However, if learners do not feel that it is being individually aligned to their needs, it may have a negative impact. In effect, the design of the instruction might trade one problem (a lack of feedback) for another (feelings of isolation and dissatisfaction). Knowing more about the relationship between the dimensions of feedback and achievement allows the instructional designer and instructor to make simple course corrections to improve the efficacy of the course.

Purpose Statement

This study follows a causal design for the purpose of understanding how a learner's achievement and satisfaction are impacted by differing amounts/levels of feedback on the four dimensions described earlier (Table 2). Observations of interactions in the learning management system reveal, quantitatively, the relationships between the feedback a learner receives and that learner’s achievement. Data from learner survey responses will be compared with LMS analytics data to reveal relationships between students’ perceptions of feedback and observation of the feedback that they received.

Knowing more about the impact of feedback helps course designers and educational technology providers create effective learning experiences. Instructional designers who collaborate with faculty can coach them toward designs that leverage every feedback opportunity so that high-quality feedback delivery efforts are not wasted. If certain measured dimensions of feedback are highly predictive of student achievement, research results could also lead to better ways of measuring online course quality.

The research questions to be investigated are as follows: (1) What is the relationship between the adult, online learner’s feedback interactions in an online course and the learner’s achievement in his/her disciplinary program of study? (2) How do learners’ perceptions of instructor feedback (their satisfaction) compare to the number of individualized, content-specific instances of instructor feedback that they received in the online course? If learners receive more individualized, content-specific feedback, are they significantly more satisfied?

Chapter 3: Methods

The purpose of this research is to inform the design of accelerated, online courses delivered to adult learners. The institution involved in the study uses master courses. The master courses are highly developed and facilitated by teams of adjunct faculty members who do not write their own content but who instruct using copied sections of the master course. Faculty course developers and instructional designers make hundreds of granular decisions when designing and improving these master courses. Iteration and refinement is important because the design of the course must work for hundreds of different faculty members as well as thousands of different learners.

Course design teams could benefit from a better understanding of the degree to which feedback impacts student achievement and satisfaction so that they can implement instructional technology in ways that preserve the most influential dimensions. In this chapter, I will discuss the student design, the institutional context for the research, participants selected for this study, the sources of data collected, the instrument used to collect survey data, procedures for analyses, statistical analyses performed, as well as limitations and delimitations imposed.

Study Design

A correlational study design was chosen because the research questions seek to explain the relationships between the frequency, distribution, timeliness, content, and individualization of the feedback a learner receives and that same learner’s achievement and satisfaction. The sequence of events (i.e., feedback received then posttest or survey taken) makes it clear that feedback could have influenced the posttests and satisfaction, but even if a correlation is found, other variables could also be the underlying cause for the effect. Regardless, the results of a causal study are still useful for understanding more about potential relationships to be investigated in the future.

Course Design Context

There are some common features of the courses involved in this study. The majority of the learners enrolled in business programs are working adults; the average age is 36 years old. Learners experience a scaffolded sequence of 5-6 week courses, and they take one course at a time, progressing through a sequence of collaboratively developed learning outcomes related to a particular business domain (e.g., marketing, accounting, management, etc.). Each week of coursework is designed to take 15-20 hours of time for the learner. Most weeks contain a reading assignment and/or video, a discussion forum, a brief paper or project, and some other type of learning activity (e.g., quiz, publisher-created courseware activity, self-paced eLearning, worksheet, practice problems, scenario response, exploration activity, etc.). Coursework is entirely asynchronous, so learners are studying and completing assignments at any time of the day or night. If instructors incorporate a synchronous activity (e.g., a webinar, Q & A session, or live lecture), the activity is recorded and learners can optionally attend the live session or watch the recording. There are no synchronous learning requirements.

Each course is collaboratively designed and developed by a team of four to six people including faculty subject matter experts, instructional designers, academic deans, assessment personnel, and a librarian. The course development process takes six to nine months and results in a series of modules and workshops, each aligned to a course learning outcome. Rubrics are created by the course development team for all open-response assignments, and assessments are embedded strategically in key courses across the curriculum to collect data for the purpose of review and continuous improvement. The business education programs utilize Peregrine assessments as measures of learners’ knowledge and skill at the beginning of their discipline specific coursework and again at the end of their coursework. The Peregrine assessments appear to the learner as assignments in their coursework, and completion of the assessment counts toward the learner’s course grade. Learners receive 50 out of 1000 course points for completing the assessment, and their assignment grade is not dependent on test performance. The Peregrine assessments are taken by learners as the first assignment in their first course of the program.

Peregrine assessments have been integrated into most courses in the DeVoe School of Business in preparation for accreditation through the Accreditation Council for Business Schools and Programs (ACBSP). ACBSP requires that schools use both formative and summative assessments of learners’ skills, preferably by incorporating externally validated assessment instruments. Peregrine assessments have a strong reputation among ACBSP-accredited institutions. Peregrine’s test items undergo a rigorous process for reliably assessing the intended learning outcomes, and new test items are continuously being validated and added to the test bank. Tests are conducted online, authenticated using the learner’s university login, and tests are timed but not proctored in any way (no video camera and no live proctoring). Many questions are scenario and case-based to reduce the ease with which learners can search online for correct multiple choice responses. Peregrine data are the measure of learner achievement used in this study.

Indiana Wesleyan University’s (IWU) DeVoe School of Business began offering courses for adult learners in the 1980s in the evenings and on weekends. Throughout the 1990s and 2000s, as IWU’s distance education degrees went online, enrollment grew at double-digit pace, and by 2010, IWU had a total enrollment over 15,000 students with 6,000 in the School of

Business and Leadership. This made IWU the largest private school in Indiana and the School of Business and Leadership one of the largest in the Midwest.

Since 2010, enrollment in the School of Business and Leadership, now the DeVoe School of Business, has declined to approximately 4,500 students, primarily due to increased competition in the adult, online market from both public schools and other private, online universities like Anderson University locally and Liberty University and Grand Canyon University nationally. Nevertheless, the DeVoe School of Business is the largest Wesleyan business school, the largest independent business school in Indiana, and it continues to innovate new programs, degree offerings, industry partnerships, and business community connections. Twice annually the school produces the *DeVoe Report*, a publication that reaches a wide and expanding national audience of business leaders who share an interest in the Virtuous Leader Model® for business education.

The DeVoe School of Business (DSB) philosophy has always put learners at the center, and this is increasingly evident in their methods for following sound instructional design practice when developing curriculum. Administrative leaders use the ADDIE model--Analyze, Design, Develop, Implement, and Evaluate--for their core curriculum development processes (Chevalier, 2011). Analysis occurs through partnerships between faculty leaders and a board of industry advisors. Design and development occur collaboratively as teams of faculty and instructional designers engage in robust course development for every DeVoe offering. Next, implementation is supported through a team of learning management system troubleshooting staff members and faculty development personnel. In addition, a multi-faceted, team-based program evaluation occurs annually for every program in DSB. For example, currently DeVoe is designing a Doctorate in Business Administration degree that incorporates problem-based learning as a threaded competency throughout the coursework, scaffolding learners' problem-solving skill throughout the program and culminating with learners' authentic final dissertation projects that will also follow a problem-based learning framework.

This background information surrounding the DeVoe School of Business is relevant because it demonstrates their historically unique perspective on postsecondary education. Courses are carefully engineered by teams of highly-qualified individuals from diverse fields and then facilitated by industry practitioners who may have less scholarly expertise but more practical experience in a particular industry. A single master course is developed and copied for use by both online and face-to-face instructors, and courses start year-round on almost every day of the week. Course sections are kept intentionally small with no more than 20 learners per instructor, and the school maintains an average course section size of 15 learners. Since 80% of DSB's courses are now taught entirely online, this study focuses solely on the online delivery modality.

Curriculum leaders have articulated regularly that they need better methods of measuring the impact of the curriculum development team's instructional design choices. The teams use popular existing models to assist in building quality courses (e.g., the Quality Matters rubric, the Degree Qualifications Profile, the Online Learning Consortium rubric, and VALUE rubrics from AAC&U), but none of these models attempt to answer questions about the granular design choices that faculty and instructional designers make regularly.

I have been a team member on many course and program development projects, usually serving as the individual with instructional design and/or assessment expertise. While the existing instructional design frameworks and principles have often been helpful in building quality courses (i.e., courses that produce satisfied and highly capable learners), there are limited frameworks to inform the integration of adaptive learning software applications. These technologies have become increasingly affordable and are heavily marketed at the conferences that faculty attend. A faculty member attending a conference returned with a handout detailing 20 choices for adaptive learning technology (Appendix A). Most technologies make a similar promise--to provide immediate, personalized feedback for every learner.

When course design teams are integrating instructional technology that provides timely, personalized feedback to learners, designers have many questions and few empirical answers related to the appropriate frequency and distribution of the technology to maintain an ideal balance between instructor-student, student-student, and technology-student interactions. Furthermore, we can imagine that the content of these interactions matters, but we lack empirical evidence around the degree to which it matters. Faculty instructors have a limited amount of time to provide students with individualized feedback, so where in the online course design is that time most effectively invested: discussions, papers, or adaptive technology? Is there a threshold for feedback effectiveness--a point at which more feedback from the human instructor is not producing an effect in student learning? Studying analytics data could start to inform answers to these questions and more. This study is one step in the direction of greater knowledge about the impact of feedback in online learning.

Participants

Institutional population. The population of online, adult learners in the United States numbers in the millions (U.S. Department of Education, 2018). This research involves a convenience sample from one institution in the Midwest. The university has been offering online courses and programs since 1999 and distance education since the 1980s, so there are well-established methods for creation and delivery of course content. This institution serves adult learners, average age 32 (four years younger than the average age of the DSB students selected for this study). Learners enroll in both online and face-to-face modalities (80% online, 20% face-to-face), but only learners from the online modality were included in this study. The sample (N=206) for this study was selected from a population of approximately 6,000 online, adult learners enrolled in associate, bachelor's, and master's business education programs at a faith-based, private, non-profit, liberal arts university.

The racial, ethnic, and gender demographics of the university student population mirrors the population of much of the Midwestern United States with approximately 18% African American, 9% Latino/a, 48% male/52% female, 6% veterans, and <5% other races/ethnicities. The institution does not track English language learners, first generation student status nor does it gather data on high school or prior college GPA. Most learners are transferring credits into the institution when they enroll, and almost half of all learners are enrolled in the business school. The university operates on a non-term, rolling enrollment basis. Learners take one course at a time, and courses are 5, 6, or 8 weeks in length depending on the degree program and level. The courses analyzed for this study were the last three courses in each learner's degree program.

Sample Group. The first parameter to determine study participants was a date parameter. Only learners who were enrolled in business courses after December 2017 were eligible for the study because the university transitioned learning management systems (LMS) at this time, and prior LMS data would be very difficult to collect and analyze. Study participants could have taken their business program pretest before December 2017, but at least their final course and posttest had to have occurred after January 2018. This parameter narrowed the group of eligible study participants significantly from 4,500 to approximately 1,200.

The second parameter that narrowed the sample was that learners needed to have studied online (not face-to-face) for at least their final three courses (nine credit hours) of their programs. In addition, their end-of-course survey data needed to be accessible. Combined, these parameters narrowed the eligible population down to 560 students. Finally, both the learner's pretest and posttest assessment data needed to be available. This parameter narrowed the eligible sample group significantly since some learners entered their degree program before the required pre-test existed, and some learners had not yet completed their program of study and the required post-test.

These requirements left 270 potential study participants. To mirror the university's population of degree earners, random stratification was used to select 83 associate (AS) level students, 87 bachelor's (BS) level students, and 39 master's (MS) level students. Categorizing a learner as either an AS, BS, or MS degree seeker was complicated by the fact that learners often enter the university seeking one degree (an associate degree) but continue on through two or three degrees, finishing as a master's level student. For the sample group, the designation as AS, BS, or MS student was made at the time the learner completed the pre-assessment.

With all of these parameters taken into consideration, the final sample of participants included in the study was 206 adult, online college students. The average age of the group was 39 with a standard deviation of 9.4, a maximum age of 65, and a minimum age of 21. The demographics of the study participants were a representative reflection of the larger population of students at the university. When learners reported race/ethnicity data, they had the option to choose more than one category for themselves, so the following percentages will not add up to 100%. For race, learners in the study categorized themselves as follows: 70.7% White, 33.8% African American/Black, 1.2% no race reported, 1.2% American/Alaskan Native, and 0.5%

Asian; for ethnicity, students were 95.3% Non-Hispanic/Latino, 4.3% Hispanic/Latino, and 0.4% no ethnicity reported.

Sixty-two percent (62%) of the study participants were female while 38% were male. This ratio roughly matches the overall university student population. The university does not gather data for first generation college students. Learners self-declare their status as veterans, and 4.3% of the participants in this study were receiving some category of veteran benefit. Study participants in the undergraduate degree programs had declared management, accounting, or business as their majors. All of the graduate degree program study participants were enrolled in the MS in Accounting degree. This occurred because the accounting program was the first to require the Peregrine pretest. Other programs have since done the same, but their learners' posttest data were not available at the time of the study. Finally, the average GPA for all learners in the study was 3.35 with a standard deviation of 0.567, a maximum of 4.0, and a minimum of 1.28.

Sources of Data

LMS analytics and observation data. Using the sample group of students who met all of the requirements for the study, a list of the three final courses taken by each learner was generated. These were the three courses each student took directly before taking the business program posttest. In general, business school students take courses in sequence, so it was likely that learners graduating at a similar time would have belonged to the same course sections prior to completing their academic programs. A trained research assistant collected data from the selected course sections, following the data collection described F1-F4 of Table 3.

Table 3

Dimensions of Feedback with Sources of Data

<table>
<tr>
<td>F1- Frequency
(a) Total # of posts read,
(b) Total # of threads created,
(c) Total # of Replies posted</td>
<td>F2- Distribution
To what extent were the feedback interactions (discussion forum activity) distributed throughout the course?</td>
<td>F3- Content
How many comments from instructors provided next steps or reinforced concepts?</td>
<td>F4- Individual-ization How many comments from instructors related directly to the individual student?</td>
<td>F5- Timeliness
Students’ self-report of the timeliness of the instructor’s feedback (End-of- course survey responses to Questions 8 and 9)</td>
</tr>
<tr>
<td rowspan="2">For each a, b and c, the total # of interactions as recorded by the LMS system data</td>
<td rowspan="2">Range= Max # of interactions – Min # of interactions. The lower the range, the more equally distributed were the student’s feedback interactions.</td>
<td>Count of comments related to the content of the assignment</td>
<td>Count of comments individualized to the learner</td>
<td rowspan="2">Students responded on a 5 point Likert scale (1 Not at all – 5 Agree entirely) to the following statements: The instructor responded within 48 hours, and assignments were graded within 7 days.</td>
</tr>
<tr>
<td colspan="2">A single comment from an instructor to a learner was counted twice if it met both criteria--related to the content and individualized.</td>
</tr>
</table>

The “Class Progress” tool in the LMS courses was used to count the F1-frequency and F2-distribution of students’ feedback interactions (Figure 1). The rectangular boxes in Figure 1 show the locations in the LMS where these data could be collected.

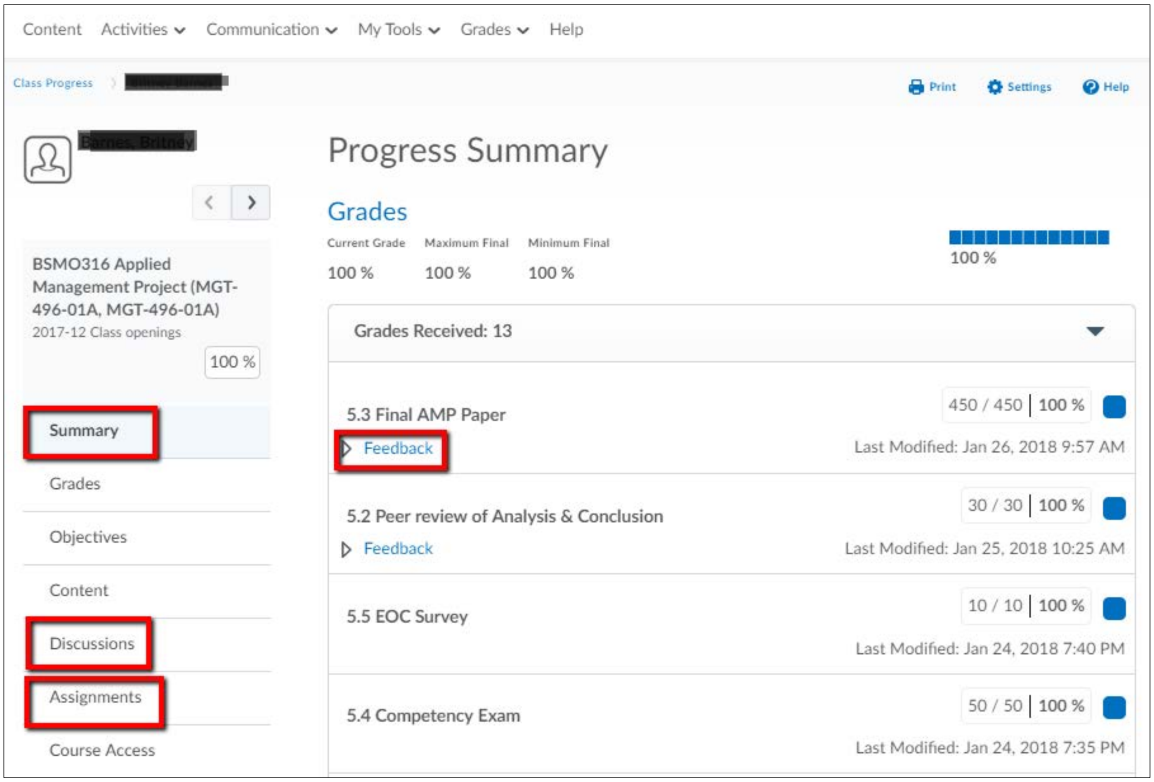


Figure 1. Class Progress Tool. This screenshot shows the summary screen in the Class Progress tool of the LMS. Rectangles indicate locations where data were accessed.

After accessing the Class Progress screen for each study participant, the following data were recorded on a spreadsheet: the number of graded items the learner had received since each grade is an instance of instructor feedback and the number of *discussion posts read*, *threads created*, and *replies posted* for each week in the course. In the Figure 2 example, for this assignment, the learner posted one original post, responded to two other posts and read 41 posts. The LMS automatically counts a discussion post as read when a learner keeps a post on their screen for longer than five seconds. Learners can check the selection box next to posts, click “mark all as read,” and the LMS will falsely indicate that the learner has read all of the posts, so it is recognized that these data are not entirely valid indicators of “posts read.”

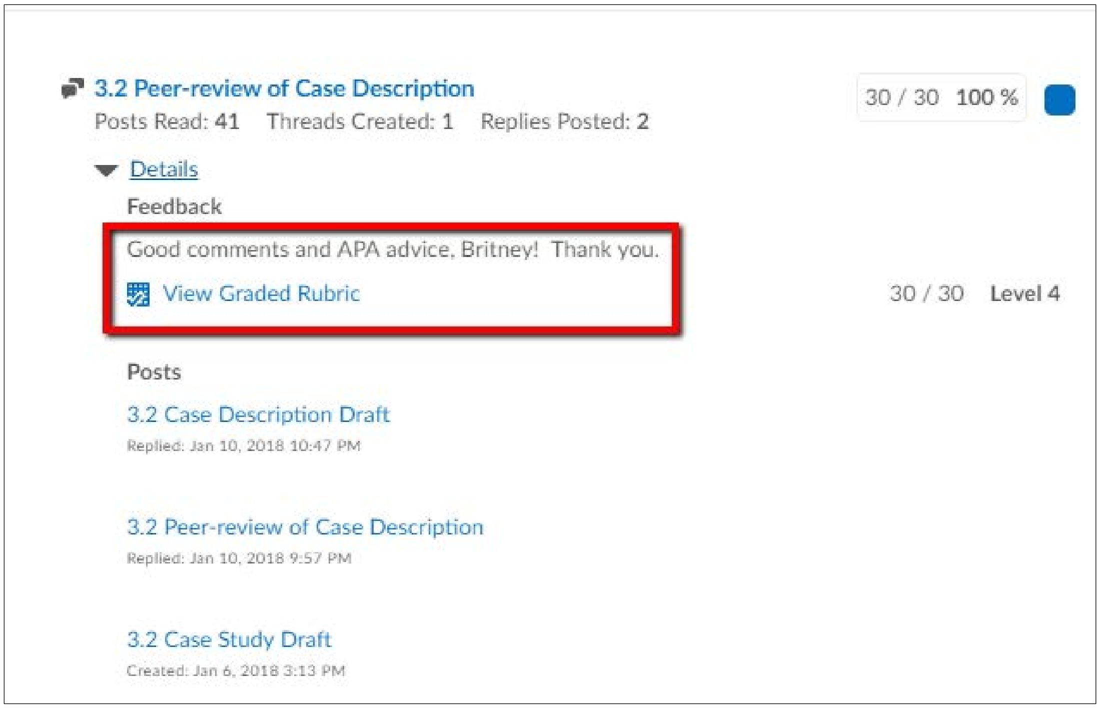


Figure 2. Screenshot of discussion progress in the LMS Class Progress screen.

In the design of most courses, rubrics were used to provide students with feedback. In the example provided in Figure 2, the instructor has used the grading rubric and provided a narrative comment: "Good comments and APA advice, Britney! Thank you." These LMS screens were the basis for collecting the data described in Table 3. The Rubric for Scoring Content and Individualization, Table 4, was used to assign a content and individualization score for each feedback instance. In instances where the instructor wrote something similar to "Please see my comments in the attached document," we also opened the document and scored the comments provided throughout the document. The above example would receive one point for both content and individualization. To score two points for content, the instructor would have to provide Britney with next steps for learning or with some specific aspect of her work that demonstrated the learning outcome well. To score two points for individualization, the instructor would have had to include a specific reference to Britney's work that could not be applied to many students' papers.

Table 4			
Rubric for Scoring Content and Individualization			
	0	**1**	**2**
Content- next steps for continued learning	The content of the feedback instance is purely motivational. (e.g., good job)	The content of the feedback instance could be perceived as either motivational or as a next step to extend or solidify learning. (e.g., Comma use is tricky.)	The content of the feedback clearly provides the student with next steps for learning or questions to extend their thinking. (e.g., Check out section 2.1 in the text and try this one again.)
Individualization- feels personal and connected to the individual learner	No individualization and the tone is impersonal. The instance could apply to any learner. (e.g., Check out section 2.1)	The feedback instance could be perceived as uniquely connected to the individual but it could also be applied to many students. (e.g., Based on your topic, you might enjoy this article)	The instance is clearly intended for the individual learner and no one else. (e.g., Is there a better way to align this project proposal with the career goal you and I talked about last week?)

The primary researcher and the research assistant compared data collection and categorization to ensure that we were capturing the LMS analytics similarly. Using the simple *content* and *individualization* rubric (Table 4), the two calibrated scoring until there was high inter-rater reliability (> 90%). To establish inter-rater reliability, the two scorers started by scoring comments on the same assignment. The researchers compared their ratings and discussed rationales for providing various scores. They continued scoring the same paper and comparing scores until their scores were calibrated- 90% of the scores of comments between the two scorers were the same. At that point, the two researchers scored assignments seaprately, working from one shared master spreadsheet.

The two researchers tallied information from the LMS to create a total count for each of the following fields for each learner: learner's course grade, number of grades received, number of assignment submissions (not discussions) with no written feedback from the course instructor or submissions that received a score of zero for content and individualization, number of instances of instructor feedback that were broad and not content related, number of instances of instructor feedback that were individualized and content related, percentage of content pages that the student visited, and number of days the student visited the course. Both researchers also tallied the following information from the LMS for each week in the course: the number of posts read, the number of threads created, and the number of replies posted as captured by the LMS system data. All of these analytics were later used to calculate the feedback variable scores F1-F4, described in Table 3.

End-of-course survey data: For the same course sections that were included in the LMS analytics data collection, the existing anonymous, end-of-course (EOC) survey data were collected. There is no way to identify an individual's responses to the EOC survey, but EOC data can be collected by course section. The mean score (Likert scale 1-5) was captured for each question in each course section for the study participants in question. Course section sizes ranged from five to 15 learners. Learners' responses to questions eight and nine were used as the measure of feedback timeliness. Using a Likert scale 1-5, *not at all* to *agree entirely*, students responded to the following two questions: "The instructor responded to student questions within 48 hours" and "The instructor returned graded work within seven days after the due date." The average score, median, and standard deviation were collected for each of these questions (Figure 3).

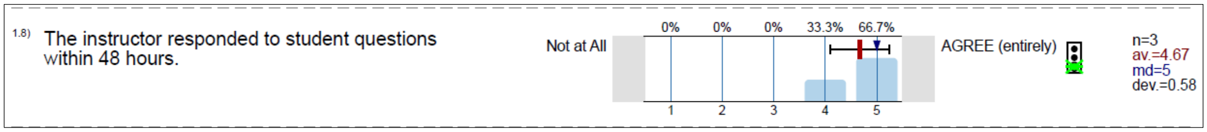


Figure 3. Example of end-of-course survey data from one question in one course section.

Research Question Two asks how students' perceptions of instructor feedback (student satisfaction) compares to the observable reality in the online course, so this comparison provides insight into the degree to which students' perceptions (satisfaction) are aligned with LMS analytics and manually captured feedback data. The leaders of the business school are interested in learning about any quantifiable indicators that could predict student satisfaction while also recognizing that satisfaction and learning are two separate constructs requiring two separate evaluation methods.

Achievement data: Peregrine Academic Services provides assessment and online educational support services to higher education institutions and other academic organizations, especially for the purpose of program evaluation and accreditation. The development, evaluation, and administration of Peregrine's test bank follows the principles and standards outlined by the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education (Peregrine, 2018). For each exam topic, Peregrine has developed a test bank of 300-500 questions, most of which are conceptual or application-based questions as opposed to definitional (Peregrine, 2018). Exams are administered online with 10 questions, displayed one at a time, per topic for a total of 100-120 questions per exam (Peregrine, 2018). Peregrine has put into place the following exam integrity measures: randomized questions from extensive question banks, randomized topic orders per exam, timed response periods for questions, a locked-down browser, and full restriction of the copy/paste functionality from the exam window (Peregrine, 2018). Inbound and outbound scores were included for the 206 students in the study. Although the assessments were embedded in the courses as required assignments, learners did not receive a course grade based on performance-only completion. It is likely that the level of effort invested in completing the assessments varies across the sample group.

Peregrine defines validity as “the extent to which the exam results are relevant and meaningful for the purpose of assessing the student retained knowledge of the selected program topics in order to assist university program managers with evaluating learning outcomes” (Peregrine, 2018, p. 3). In addition, Peregrine outlines an extensive list of procedures that were followed to improve the extent to which exam results can be used to make valid interpretations about the knowledge students have acquired during a program of study. Face and construct validity, as well as exam reliability measures, are fully described in Peregrine’s report which indicates that 400 institutions and 5,000 individuals have been involved in reviews of their products and services as of August 2015 (Peregrine, 2018).

Instruments

End-of-course survey. The end-of-course (EOC) survey (Appendix B) is the standard instrument delivered electronically to all students across the institution at the end of every course. It was developed by a faculty-lead task force at the institution approximately three years ago and has received just over 64,000 student responses. Responses are provided on a five-point Likert scale and are completely anonymous. The EOC provides an 85% response rate, and learners receive 1% of their grade as extra credit when they complete the survey.

Content and individualization rubric. The *Rubric for Scoring Content and Individualization* (Table 4) was used in the initial analysis of student/instructor interactions on two of the five dimensions of feedback being studied--content and individualization. The research assistant was trained to code feedback interactions in online courses using the rubric. The assistant and primary investigator followed a blind scoring protocol to learn to code feedback with >90% inter-rater reliability before proceeding with individual data analysis using the rubric. The rubric was used to score each comment that an instructor provided to a learner on written assignments.

Research Approvals

Institutional review board (IRB) approval was granted through Indiana Wesleyan University's IRB and then subsequently through Indiana University. The designated IRB study number is 1805351912, and signatories from both Indiana Wesleyan University and Indiana University provided approval on June 5th and 6th, respectively, 2018. Permission to collect data from a sample 250 students from the DeVoe School of Business was granted by the dean of the DeVoe School of Business.

There was one ethical risk described in the IRB proposal and review. The researcher's primary job responsibilities require that she share instances of faculty negligence or impropriety with the appropriate supervising dean. There was an ethical concern about disclosing these instructor behaviors if they were discovered during data collection. Would it be permissible for the researcher to share evidence of faculty negligence if it was discovered during research? To not do so would conflict with the researcher's professional responsibilities. If discovered during research, the IRB decided that it would be permissible for the researcher to share knowledge of faculty negligence, with the appropriate dean, just as it could occur in the midst of daily employment responsibilities. The research assistant, who also works daily with sensitive faculty and student information, signed a confidentiality agreement and was similarly permitted to disclose discoveries to the researcher or deans. No such concerns were observed or escalated as a result of data collection and analysis.

Procedures and Analysis

Research Question One. What is the relationship between the learner's feedback interactions in an online course and the learner's achievement in his/her disciplinary program of study? The variables to be analyzed in this first research question include four independent variables related to feedback interactions and one dependent variable- achievement.

To measure achievement, a paired samples t-test was used to compare the results of learners' pretest and posttest scores on the Peregrine standardized assessments described previously. The highest possible score on the test was 100. Pretests were taken by students during the first courses of their degree programs and posttests were taken during their capstone courses. The data met the assumptions necessary for a paired samples t-test. A boxplot showed five outliers on the bottom end and three outliers on the top end. None was determined to be extreme since none was more than 1.5 box-lengths away from the edge of the box. Therefore, they were kept in the analysis. Visual inspection of a Q-Q plot of the differences between students' pre and posttest scores reveals normal distribution, so the assumption of normality was not violated (Figure 4).

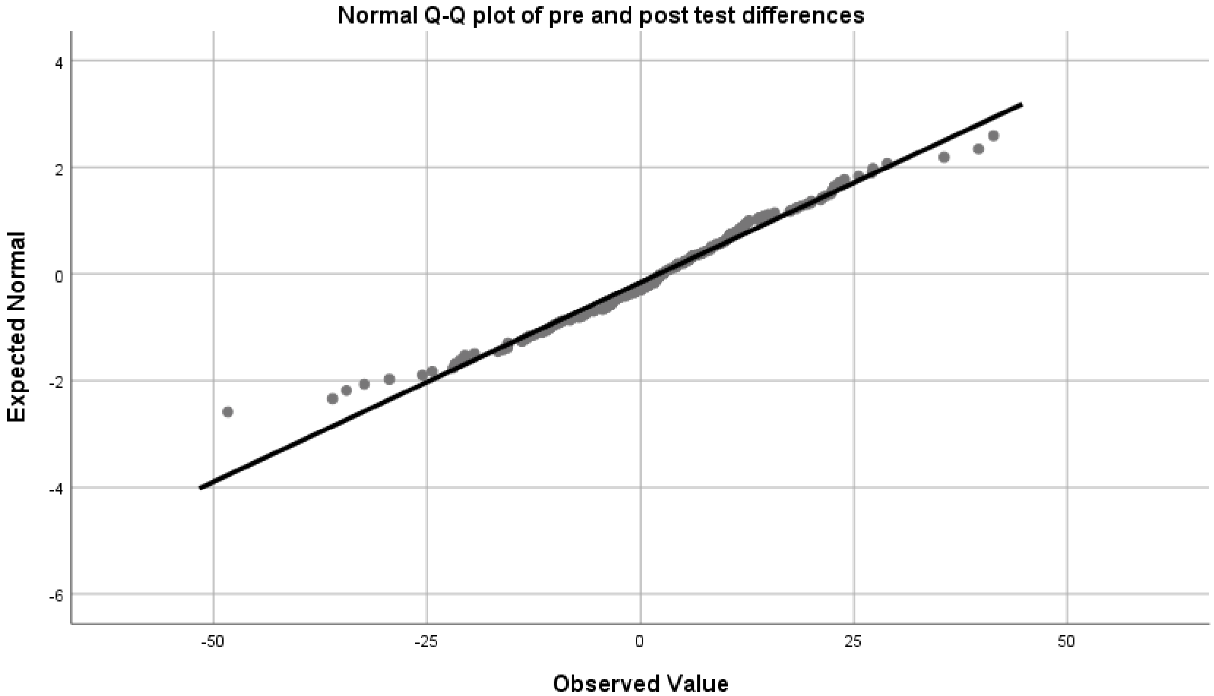


Figure 4. Q-Q plot of differences between pretest and posttest

Learners' scores on the pretest were (M=44.30, SD=11.10) and scores on the posttest were (M= 46.44, SD=12.98). Coursework in the business program elicited a mean increase of 2.14, 95% CI [0.30, 3.98] points on the end-of-program exam from pretest to posttest. A t-test was conducted to identify the difference between pretest and posttest scores. The t-test revealed a statistically significant difference at the $p < .05$ level, t (205) = 2.295, p = 0.023. The results showed that students significantly improved their scores on the posttest compared to the pretest.

On average, individual posttest scores were 12.98 points away from the mean score. Data indicate that our distribution is slightly positively skewed to the right and a -.220 for kurtosis indicates that our distribution is slightly lower than normal distribution. The range of scores varies from 17.8 to 83.75 on the standardized posttest. The histogram in Figure 5 provides a visual for the distribution.

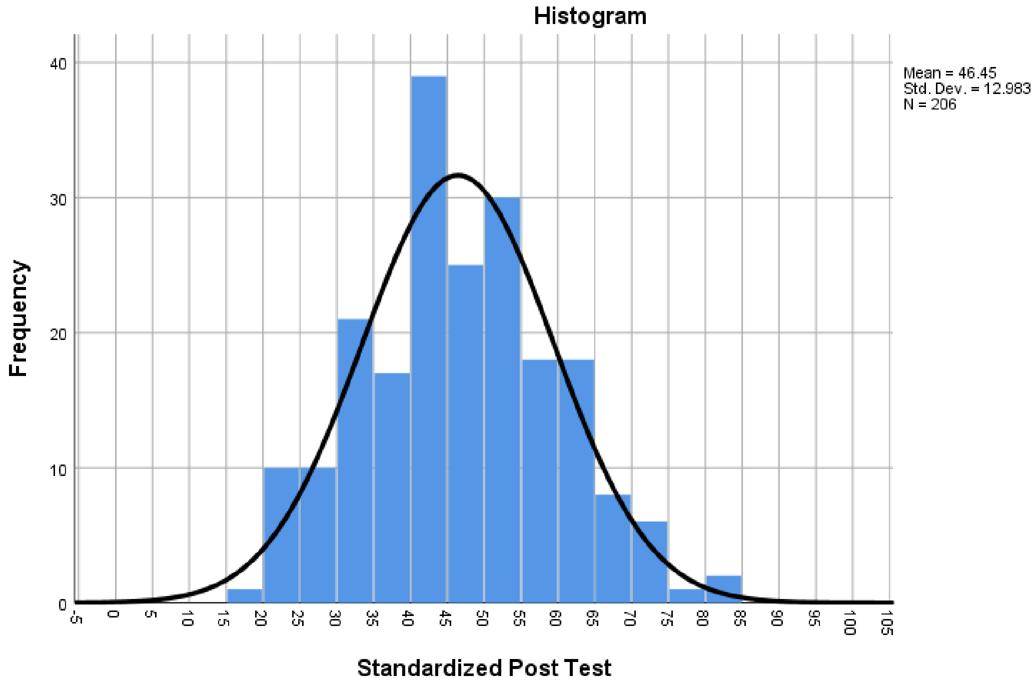


Figure 5. Histogram of standardized posttest scores

The next aspects of Research Question One to calculate were the LMS analytics data collected for each dimension of feedback (Table 3). How timely, frequent, distributed, individualized, and content-related were the instances of feedback that the learners received? Descriptions for the methods used to collect data for each of the dimensions are as follows.

F1-Frequency. Frequency, F1, was measured by counting (1) the number of posts a learner read according to LMS system data, (2) the number of discussion threads they created, and (3) the number of replies posted to their threads. Adding these numbers across all weeks in all three courses provides the overall frequency of feedback interactions with either peers or instructors in the online discussion forum.

Because courses differ in the number of opportunities provided for students to engage in feedback and dialogue with one another, frequency scores were converted to ratios for the purpose of comparison, similar to the methods used by Xing, Wadholm, Petakovic, and Goggins

(2015) in their mixed-methods, correlational study. For example, learner ID 1 had 45 opportunities to engage in dialogue with peers and instructors by either posting, reading, or replying in discussion forums. This learner engaged in 298 instances of posting, responding to, or reading feedback. The ratio of instances to opportunities is 6.622 (298/45). The higher the ratio number, the more feedback interactions the learner experienced per opportunity to engage in feedback. There would be unlimited opportunities for learners to interact for each opportunity. On average, students interacted 4.89 times per opportunity (Figure 6). The learner ID 1, described in the previous example, would be in the bar labeled with six and would join approximately 18 other learners in having a ratio of around six interactions per opportunity for dialogic feedback (Figure 6).

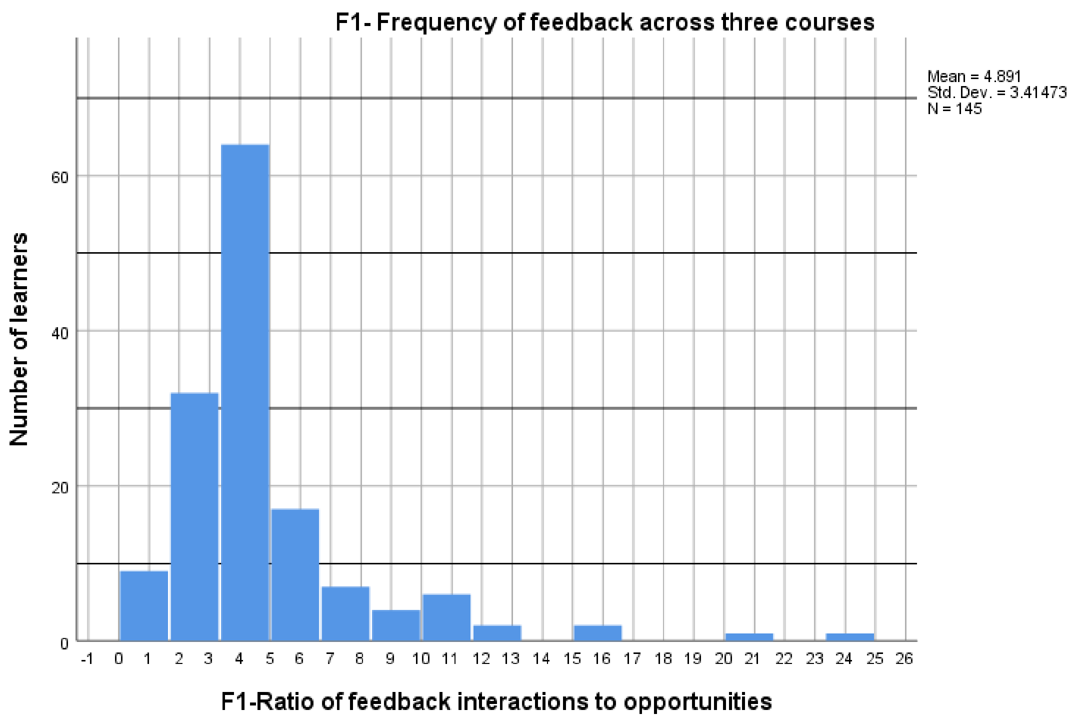


Figure 6. Feedback interactions per opportunity

F2-Distribution. For each learner, a number of interactions was calculated by counting interactions that occurred in each week of the course. Courses are five weeks long and data were collected across three sequential courses. For example, in one course, learner ID 1 had 14 interactions in the first week of the course, 21 in the second, 20 in the third, 16 in the fourth, and 18 in the last week of the course. These same data were collected for that same learner for two other courses and then totaled by week: Week One= 45, W2=49, W3=56, W4=61, and W5=87. Across three courses, this learner participated much more frequently in the final week than in the first. In theory, learners who exhibit consistent efforts over time are more successful than those who interact sporadically (distributed vs. massed practice). To create a metric that would reflect the degree to which the learner had distributed his/her practice across weeks of a course, a range was calculated by subtracting their lowest week from their highest. Higher scores on the F2-Distribution dimension indicate that the learner's participation is more highly dispersed across weeks than a learner whose F2-Distribution score is lower.

Results indicate that the average range of number of interactions is 14 meaning that learners' interactions across weeks of their courses differ by about 14 interactions. On the low end, about 19% of learners had no differences in the quantity of interactions across fifths of a course. Their interaction efforts were consistent over time. On the high end, one learner had a range of 78 interactions meaning that in one fifth of their coursework they participated 78 times more than in the lowest fifth.

Table 5

F2- Distribution of Feedback Interactions by Week

		W1_sum	W2_sum	W3_sum	W4_sum	W5_sum
N	Valid	192	192	192	192	192
	Missing	0	0	0	0	0
Mean		27.9271	30.7344	30.6406	28.7500	27.0469
Median		24.0000	26.0000	26.0000	24.0000	23.5000

Sum	5362.00	5901.00	5883.00	5520.00	5193.00

The summarized data in Table 5 show that, overall, the second fifth of the course (the second week in a 5-week course) had the highest average amount of interaction, and across all learners, the average amount of interaction from week to week varied by three interactions.

F3-Individualized Content. Two of the dimensions of feedback from the original rubric (Table 3) were combined into one dimension for analysis because measuring them separately violated the assumption of independence necessary for multiple regression analysis. When coding and scoring instructors' comments on students' papers, it became clear that when instructors provided content-specific feedback, those comments were most often individualized to the learner. It was rare to find individualized comments that were not content-specific or content-specific comments that were not also individualized. Instead of treating these dimensions as two separate variables as originally intended, they were analyzed as one variable: F3-Individualized Content (I_C).

Comments that were both content-specific and individualized counted as two instances of individualized content-specific feedback. If the comment was only content-specific OR individualized, it counted for one instance of this dimension of feedback. For example, "Chapter two provides good guidance on setting up financial reports" would count as one instance of I_C feedback because it is content-specific. The following comment would count as two instances of I_C feedback because it is both content-specific and individualized to the learner's response: "I'm glad you used tables, but check out the chapter two guidance on setting up financial reports to present your tables more clearly." Comments did not count as individualized or content-specific if they were broad or purely motivational (e.g., good work, try again, nice analysis).

Figure 7 shows the number of instances of individualized, content-specific feedback provided by instructors to students. For example, the tallest bar shows that 24 students received approximately 10 instances of individualized, content-specific feedback from instructors. To the far right of the figure, two students received as many as 42 instances of individualized, content-specific feedback, and to the far left, one student received as few as two instances.

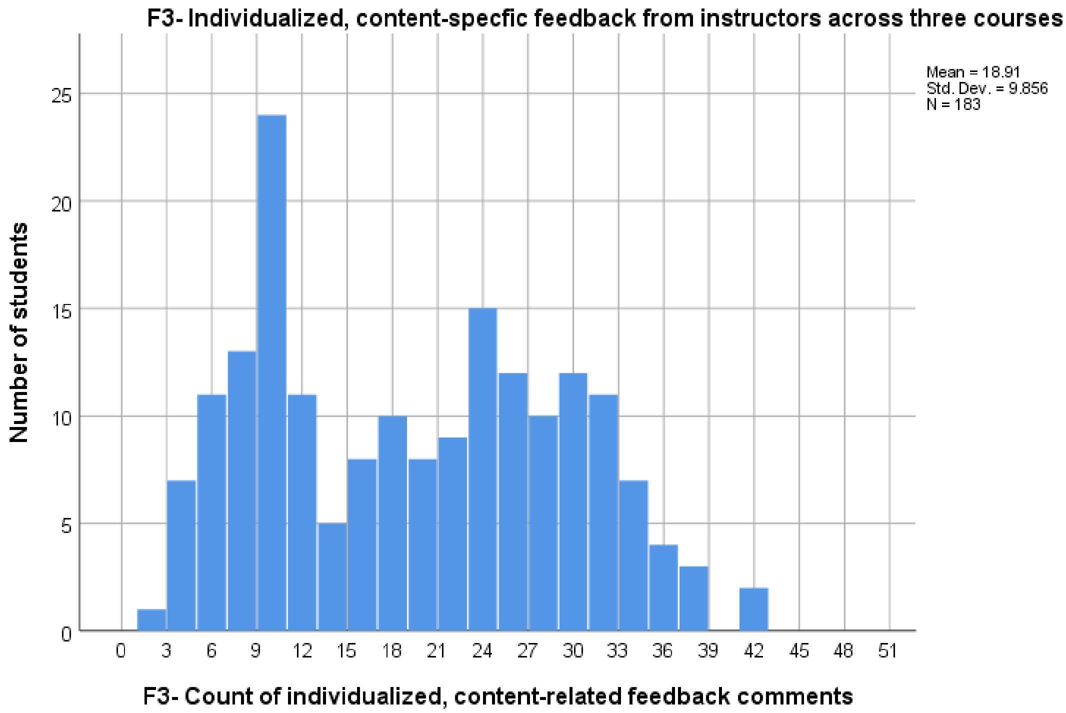


Figure 7. Frequency of individualized, content-specific comments

F4-Timeliness. The timeliness variable is the only variable that was calculated using learners' self-reported data rather than data available through the LMS. All students have the opportunity to respond to an end-of-course survey at the conclusion of every course. Two of the questions on the survey ask students to report on two aspects of timely feedback. The survey is delivered electronically, and students choose from five numbered options on a Likert scale with one being strongly disagree and five being strongly agree. Data were extracted from the survey results for questions eight and nine only. Item eight asked students to respond to the statement:

The instructor responded to student questions within 48 hours. Item nine asked students to respond to the statement: *The instructor returned graded work within seven days after the due date*. There were 174 valid responses to these two questions, an 84% response rate which matches the typical response rate for the institution. Mean scores, with standard deviations in parenthesis, for questions eight and nine respectively were 4.734 (0.292), and 4.738 (0.294). The students rated the items favorably showing that they are satisfied with the timeliness of the feedback they receive.

Feedback variables (F1-F4) summary. The intent of Research Question One was to learn more about the relationships between four dimensions of students' feedback experiences and those same students' posttest scores on a standardized exam. First, it was determined that there was a statistically significant difference between students' pretest scores and their posttest scores on the Peregrine business exams. The four dimensions of feedback measured were: F1-Frequency (the volume of feedback learners receive from their instructors and/or peers); F2-Distribution (the extent to which feedback interactions were dispersed evenly across the five weeks of a given course); F3-Individualized content (the number of comments from instructors that provide individualized and/or content related next steps); and F4-Timeliness (learners' self-report of timely instructor feedback). These four variables were independent variables for the regression analysis. The dependent variable was the Peregrine, posttest exam score.

Multiple regression analysis. A multiple linear regression analysis was used to compare learners' F1, F2, F3, and F4 independent variables with learners' final scores on the Peregrine exam to determine the extent to which any of the variables had an impact on learners' final exam scores. In general, multiple regression analyses would explain the variance in participants' posttest scores by describing the relevance of each of the independent variables as predictors of the total variance (Laerd, 2015). It is, therefore, the appropriate statistical analysis to use in answering Research Question One.

Of the 206 cases identified for this study, 139 provided data for all of the independent variables. Sixty-seven cases were removed from multiple regression statistical analysis because they did not satisfy all of the conditions necessary for inclusion: the learner took the particular course face-to-face instead of online, the learner dropped or joined the course mid-way, or the instructor delivered the course in a hybrid format and, as a result, some data were missing.

Assumptions. The multiple regression model meets the necessary assumptions and allows us to model the relationship between learners' posttest scores and their experiences with feedback during the final three courses of their business degree programs. Based on prior research described in the literature review, there should be a positive linear relationship between each of the feedback variables under analysis (frequency, distribution, content/individualization, and timeliness) and the posttest variable. The goal of this analysis is to determine how much of the variation in exam scores can be explained by the feedback variables, if any.

The test for independence of observations produced a Durbin-Watson statistic of 2.06, which indicates independence of residuals. Visual inspection of scatterplots of the studentized residual by unstandardized predicted value indicated sufficient linearity. Visual inspections of individual scatterplots of each independent variable compared with the dependent variable also indicated sufficient linearity. There was sufficient homoscedasticity as assessed by visual inspection of the same scatterplots. Figure 8 provides an example of one of the scatterplots that was visually inspected. Points of data do not show any discernable patterns or groupings, so the assumptions described above were not violated.

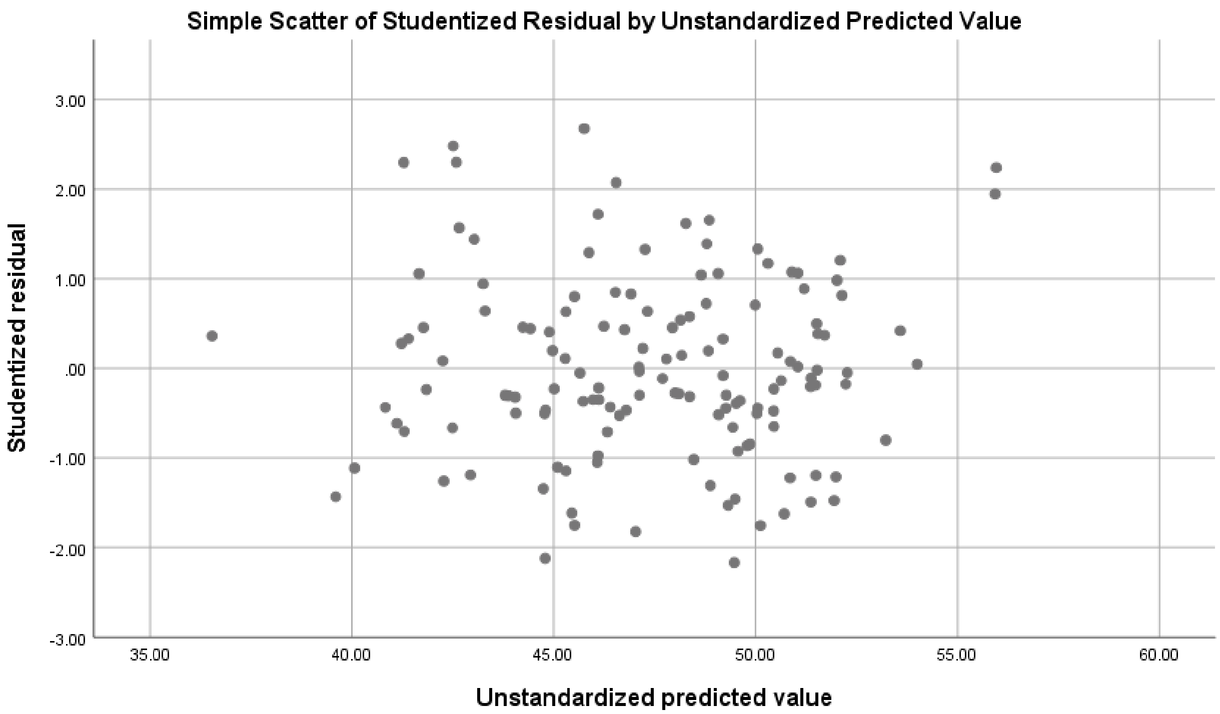


Figure 8. Testing assumption of homoscedasticity

According to Hair et al. (2014), when testing for collinearity, a Tolerance value less than 0.1 indicates a problem with collinearity. As indicated in Figure 9 in the two columns to the far right, all variables had Tolerance values greater than 0.1 and VIF values less than 10.

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	49.486	18.680		2.649	.009	12.537	86.435					
	F1-Freq	.230	.430	.061	.535	.594	-.620	1.080	.017	.046	.045	.546	1.832
	Range MAX-MIN	-.073	.094	-.088	-.772	.441	-.259	.114	-.074	-.067	-.065	.550	1.817
	F3- Calc of content related comments/ total # of graded assignments	-.133	.161	-.093	-.826	.410	-.451	.185	.045	-.071	-.070	.560	1.787
	F3 QofF FinCrs	.768	.365	.239	2.104	.037	.046	1.490	.189	.179	.178	.553	1.807
	F5-Timeliness	-1.294	3.969	-.029	-.326	.745	-9.145	6.556	.044	-.028	-.028	.875	1.142

a. Dependent Variable: Standardized Post Test

Figure 9. Testing Tolerance values assumption

Two cases were filtered out of the dataset because they failed the leverage point assumption test leaving 137 cases for analysis. Leverage values greater than 0.2 are considered risky (Huber, 1981), so these two cases were filtered out, and all former assumptions were checked again. Cook's values were checked, and there were no cases scoring above one. A P-P Plot of the post-test scores indicates that the assumption of normality had been met (Figure 10).

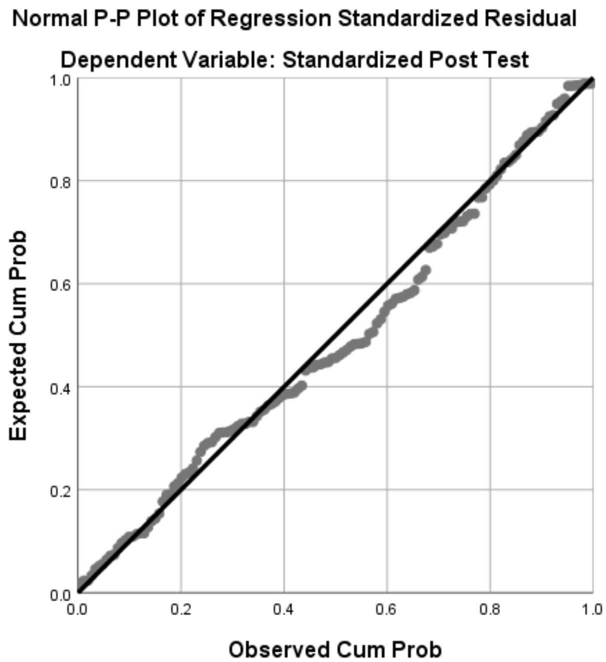


Figure 10. Test for assumption of normality using P-P plot

Research Question One results. The results of the multiple regression analysis show that the R2 for the overall model was 5.2% with an adjusted R2 of 1.6% which is a small size effect according to Cohen (2003). Additionally, there is no statistical significance to the overall model as indicated by 0.212 Sig. in the ANOVA for the model. Even with the addition of the independent variables, we do not have a model that is statistically significant at predicting students' exam scores, F (5, 131) = 1.446, p = 0.212 or p > .05. The measured dimensions of feedback (F1-frequency, F2-distribution, F3- individualized content, and F4-timeliness) were not statistically significant predictors of final exam scores. Table 6 presents data showing the standardized coefficients for each variable.

Table 6

Summary of Multiple Regression Analysis

Variable	B	SE_B	β
Intercept	51.719	18.804	
F1- Frequency	0.30	0.473	0.007
F2- Distribution	-0.74	0.94	-0.86
F3a- Content Comment Ratio	-0.135	0.164	-0.095
F3b- Content Comments Final Course Only	0.811	0.369	0.254
F5- Timeliness	-1.649	3.992	-0.038

Note. * $p < .05$; B = unstandardized regression coefficient; SE_B = Standard error of the coefficient; β = standardized coefficient

Research Question Two. The results of the second research question showed whether there were correlations between learners' satisfaction as captured by the end-of-course survey and feedback activities that were observed in the courses. This question was posed because the bias inherent in self-reported data caused administrative leaders to question the extent to which these data would be helpful for improving the designs of courses. If learners' perceptions of feedback were highly correlated with the number of feedback instances that were counted in the courses, the data from the end-of-course (EOC) surveys (an easily obtained data source) would be more seriously considered as a method for identifying courses where the design of the instruction should be revised.

From the original sample group of 206 cases, there were 174 cases that had EOC survey data available. This represents 84% of the total sample, which was close to the institutional average response rate of 85%. Learners responded to the nine-question, electronically delivered EOC survey using a five-point Likert scale where one equaled *not at all* and five equaled *agree entirely*. Q1 was omitted because it is irrelevant to this study. The EOC survey statements were:

Q2- The instructor respected me as an adult learner by demonstrating qualities such as patience and kindness.

Q3- The instructor demonstrated a willingness to assist students.

Q4- The instructor's knowledge of course content was evident in the instruction of the course.

Q6- The instructor's feedback provided direction or encouragement beneficial to my academic success.

Q7- The instructor's grading and feedback aligned with course rubrics, scoring guides, and written directives.

Q8- The instructor responded to student questions within 48 hours.

Q9- The instructor returned graded work within seven days after the due date.

Data were not reported for each individual student since surveys were taken anonymously; however, data could be analyzed by course section by using the average score for the section for each question. For example, Student ID One was taught in course section ASBO115, and data were available for each learner in course section ASBO115. Therefore, average scores from each question's results from ASBO115 were used as if they were student ID one's responses to the EOC survey. This method may have been less reliable than having an individual learner's score results, but it did eliminate biases and effects that can occur when students know that their data are being collected for research. These learner respondents were unaware of the research study; therefore, their data were not subject to response bias.

No more than seven learners were from any one course section, and the sample (n=174) represents 55 different course sections. Learners from the same course section tended to provide similar scores on their end-of-course surveys. There was evidence of this in the analysis in that standard deviations for each survey question are all less than 0.5 on a five point Likert scale.

Mean scores (with associated standard deviations in parentheses) for questions two (Q2) through nine (Q9) had ratings of: Q2- 4.77 (0.25), Q3- 4.72 (0.36), Q4- 4.69 (0.33), Q5- 4.66

(0.41), Q6- 4.62 (0.44), Q7- 4.66 (0.35), Q8- 4.72 (0.29), and Q9- 4.74 (0.29). All of the averages of the response scores were well above a level four rating showing that students mostly agree or strongly agree with positive statements about the quality of their courses. The highest overall scores were for questions eight and nine which were the two questions related to timely feedback. The average ratings of the sample group were very close (within +/- 0.2) to the average ratings of the larger institutional population. This indicated that the sample group was representative of the larger population with regard to their levels of satisfaction. Descriptive results showed that students were highly satisfied with the design of their courses.

Next, a Spearman's rho analysis was used to determine whether there was a relationship between learners' self-reported EOC data and the results of counting and scoring feedback interactions in the LMS. Spearman's rho was selected because the variables were interval/ratio data, and some variables exhibited non-linear associations with outliers. The hypothesis was that learners would rate many of the end-of-course survey items more highly if they received more instances of individualized, content-specific feedback.

The results indicated that there were significant positive correlations between the amount of individualized, content-related feedback that learners received and learners' responses to the EOC survey questions. Specifically, the positive correlation between learners' responses to questions two, eight, and nine and the amount of individualized, content-related feedback they received was statistically significant, $r_s(174) = 0.029$ (Q2), 0.031 (Q8), 0.043 (Q9), $p < 0.05$. The positive correlation between learners' responses to questions three through seven were strongly statistically significantly correlated with the amount of individualized, content-related feedback they received, $r_s(174) = 0.000$ (Q3), 0.000 (Q4), 0.009 (Q5), 0.000 (Q6), 0.000 (Q7), $p < 0.01$.

Correlation coefficients between EOC survey response data and feedback data are provided in Table 7.

Table 7									
Summary of Correlation Coefficients for Research Question Two									
Correlation Coefficient	.166*	.334**	.263**	.197**	.425**	.303**	.164*	.153*	1.000
Sig. (2-tailed)	.029	.000	.000	.009	.000	.000	.031	.043	
N	174	174	174	174	174	174	174	174	206
**. Correlation is significant at the 0.01 level (2-tailed). *. Correlation is significant at the 0.05 level (2-tailed).									

Further Analysis. Pearson correlation tests were conducted to determine whether there were any significant correlations between any of the relationships shown in Table 8. Results showed that one variable had a significant positive correlation with the posttest scores. When students received higher quantities of individualized and content-related feedback on the assignments in their capstone course, the final course in their programs, this correlated with higher scores on the posttest. None of the other variables had significant relationships with the posttest. These results indicate that one way to increase scores on the standardized posttest was to provide individualized and content specific feedback to students during the capstone course which occurred during the five weeks leading up to the date of the posttest.

Table 8		
Correlations Between Independent Variables and the Posttest		
F3- Individualized content feedback in the final course only.	Standardized Posttest	Pearson correlation= 0.17* Sig. (2-tailed)= 0.01 N=206, *$p < 0.05$
F3- Individualized content feedback on all assignments (3 courses)	Standardized Posttest	Pearson correlation= 0.13 Sig. (2-tailed)= 0.08 N=183

F1- Frequency of feedback	Standardized Posttest	Pearson correlation= 0.05 Sig. (2-tailed)= 0.52 N=165
F2- Distribution of feedback across weeks in a course	Standardized Posttest	Pearson correlation= 0.06 Sig. (2-tailed)= 0.39 N=192
F4- Timeliness	Standardized Posttest	Pearson correlation= 0.07 Sig. (2-tailed)= 0.37 N=174
AVG % age of online course pages visit (LMS analytics)	Standardized Posttest	Pearson correlation= 0.05 Sig. (2-tailed)= 0.45 N=192

Limitations

Similar to most correlational research, the primary limitation of this study was that it was not possible to rule out alternative explanations for any relationships that were observed between dependent and independent variables. For example, a learner's satisfaction with his learning experience could also have been influenced by the instructor's personality or by the learner's experience receiving personal attention. It is not possible to state definitively that it is individualized and content-related feedback exclusively that leads to increased student satisfaction.

A second limitation relates to the generalizability of the findings. The data collected represents diversity in many ways (e.g., age, gender, race/ethnicity, prior knowledge as evaluated by the disciplinary pre-test), but it does not represent diverse institutions or programs outside of the DeVoe School of Business. It also does not represent diverse delivery modalities since all of the study participants were working 100% online. Finally, beyond feedback interactions, there were many other variables that could contribute either positively or negatively to learner achievement. Quality of course content, relevance of assignments, media integration, technology integration, faculty experience, learner prior knowledge, and many other psychological factors (e.g., motivation, growth mindset, sense of belonging, etc.) all played a role in learner achievement. The findings in this study were limited to one school at one institution.

Learners from the DeVoe School of Business were selected as an intentional delimiter for the research. The prevalence of the pre/posttest used by the school provided a feasible method to test the design of the study. If the regression model was valid for its intended purpose, the sample could be expanded to include other disciplines where paired samples data do not exist.

A common concern in correlational studies is selection bias. Pre-test results for the participants in this study showed an almost identical level of prior knowledge to that of the larger population in the DeVoe School of Business. The normal Q-Q plot of pre and posttest differences (Figure 4) looked very similar to the same analysis of the larger school population. Also, the average score on the pre-test for the sample group N=206, M=44.30, SD=11.1, was very similar to the average scores and standard deviation for the entire population with data available N=652, M=45.2, SD=9.8. Response bias was not a limitation for this study since learners' data were collected after they had completed their academic programs without their knowledge.

There were some unanticipated challenges during data collection that reduced the sample size of usable data from 206 to 137 usable cases for the regression analysis. System data were not available for 67 learners because of an LMS transition and because learners had been switching in and out of online programs to or from face-to-face programs, and two other cases were removed because they failed the assumption of Tolerance test. Data were only available in the online courses, and sometimes, even though a learner was classified as "online" and therefore included in the participant group, he/she could have taken a face-to-face course either before or after the date of classification. This issue caused the majority of the 67 unusable cases.

A second challenge revolved around the Peregrine posttest exam results. For the date range selected for this study, Peregrine end-of-program exams were new to the DeVoe School of Business faculty and administrators. In fact, the learners in this study were some of the first to take the exam both at the beginning and end of their academic programs. During data collection, it became clear that the difference between pre and posttest exam scores, although significant, was quite small. On average, only three to five points of growth were achieved by learners between their pretests and their posttests. Some learners had taken 30 credit hours of discipline specific coursework, so faculty were expecting much greater gains. The sensitivity of the posttest exam was not sufficient to capture the learning that had occurred for learners throughout their coursework.

For Research Question Two, the institutional end-of-course survey was chosen as the data collection instrument. It should be noted that this instrument was collaboratively developed by the institution and has not undergone the rigorous testing that often accompanies normed or validated survey instruments. A common limitation for self-report data is the risk for respondents to present themselves or others favorably to make a positive impression (Worthen, White, Fan, & Sudweeks, 1998). Nevertheless, in this study, much of the risk for social desirability bias (Kopcha & Sullivan, 2006) was mitigated because respondents were not aware that their responses would be used for research purposes. Responses were anonymous to the instructor to increase respondent honesty.

Overall, the scope of this study was primarily focused on one school within one university. The purpose was to determine whether instructional design efforts to focus on feedback could be quantified and correlated with achievement and satisfaction. If it could be shown that certain levels of feedback interaction were directly related to learner exam scores and

EOC survey results, instructional design teams would be better equipped to identify courses in need of revision. They would also be able to measure the results of feedback-related instructional design improvements made to the courses. The lack of alignment between the standardized posttest and instructional objectives was a significant limitation to the usefulness of the results. Caution should be exercised in generalizing the results of this study into other contexts.

Chapter 4: Findings

Introduction

This chapter provides a summary of the results that were reported in the previous chapter. The overall purpose for the research questions are reviewed, followed by a summary of an issue with the dependent variable—the posttest exam data. Next, alignment is addressed with a summary of the primary ways in which the research results align with the problem that prompted the development of the research questions. The chapter concludes with a summary of the most actionable findings.

The purpose of this correlational study was to determine whether data captured from several dimensions of a learner's feedback experience could be used to predict learner achievement and satisfaction. Part of the goal was to determine whether there is merit in considering LMS analytics data as proxy measures or "lead measures" for quality (McChesney, Covey, & Huling, 2016, p. 9). Some studies indicate that there is merit in collecting and analyzing certain LMS learning analytics data as predictors of student success (Calvert, 2014). Participants for this research were selected from a population of adult, online learners at a university in the Midwest. Sources of data included LMS analytics data, observations of online feedback interactions, end-of-course survey data, and pretest and posttest exam scores. A multiple regression analysis revealed that the following variables: (1) the frequency of feedback, (2) the distribution of feedback, (3) the individualized and content-specific nature of feedback, and (4) the timeliness of feedback were not statistically significant predictors of achievement on the standardized, posttest exam taken by learners at the end of their academic degree program (i.e., associate, bachelor's, and master's). Nonetheless, there was a statistically significant ($p<.05$) positive correlation between the number of instances of individualized and content-specific feedback a learner received and his/her satisfaction with courses. There was also a statistically significant ($p<.05$) positive correlation between the amount of individualized and content-specific feedback learners received in their final capstone courses and their posttest exam scores.

Alignment to the Problem

The problem that initiated this research revealed itself when a group of instructional designers and faculty subject matter experts were incorporating adaptive learning technology into an online course. Even though final exam scores increased, learners were less satisfied with the experience. To ameliorate the situation for future course design projects, the team developed a rubric (Table 1) that, they theorized, would help them balance human and artificial sources of feedback when designing courses. Adaptive learning tools could deliver timely and frequent feedback, for example, but they are not effective at providing personal feedback to humanize the online learning experience. The team wanted to balance both learner achievement and learner satisfaction. They hypothesized that a course that scores highly on the rubric (Table 1) could provide both benefits. It was not feasible to address all aspects of the rubric (Table 1) with this research study. To narrow the scope of the study, two research questions were asked to learn more about the validity of the rubric for evaluating courses. The two research questions asked were: (1) what is the relationship between the learner's feedback interactions in an online course and the learner's satisfaction and achievement in his/her disciplinary program of study, and (2) how do learners' perceptions of instructor feedback (their satisfaction) compare to the observable reality in the online course?

The results of this research indicated that the more individualized, content-specific feedback learners received during their final capstone courses, the higher they would score on the posttest exam occurring directly after the feedback. Similarly, the more individualized and content-specific feedback comments they received over three courses, the more likely they were to be satisfied with all of their online course experiences. This finding informed the root problem because it provided evidence that the design of the original adaptive learning course could have been improved by also providing opportunities for instructor-student individualized feedback interactions.

Posttest exam issues. As described in Chapter Three, standardized, posttest exam data were used as the dependent variable for RQ1, measuring learner achievement across multiple courses. The results of the multiple regression analysis indicated that achievement, as measured by the standardized posttest exam, could not be predicted by the LMS analytics feedback variables captured for this research. An assumption of the research was that the posttest exam topics were aligned with the disciplinary content taught in learners' courses. It became apparent, both through this research and through the work of the faculty leaders in the DeVoe School of Business, that learners were not meeting established benchmarks on the posttest exam. On average, learners were only gaining three to five points (out of 100), between their pretests at the entrance of the program and posttests at the end of the program

One year after the posttest data were first collected and discussed with administrative leaders, faculty course leaders made adjustments to the curriculum. They analyzed the exam topics where minimal growth had been achieved by learners, and then adjusted the topics covered by the exam or the design of the courses accordingly to improve alignment. Posttest exam topics were much better aligned with the learning outcomes taught throughout the program. Such alignment led to higher gains for students. While some topics displayed marginal growth of just two to three points (business leadership), topics like operations/production management showed an average increase of 14 points between the pretest and posttest.

In the current study, the Peregrine posttest assessment that had been implemented at the time of data collection was not optimally aligned with the course curriculum. There was evidence that posttest scores immediately started to improve once faculty aligned the posttest exam topics with the program learning outcomes and curriculum map. It is conceivable that if the posttest had been a more sensitive measure of student learning, the results of the regression analysis might have differed. There was evidence that poor posttest-to-outcome alignment was one confounding variable in this study. Overcoming this limitation was not within the scope of the current study. This limitation primarily affected Research Question One which relied on the posttest data as the dependent variable.

Descriptive Results for Research Question One

The average score on the posttest was 46.44 points (of a total possible 100 points), while the average pretest score was 44.30 points. While the improvement was statistically significant ($p<.05$), it was not sensitive enough to capture the range of learning that occurred amongst 206 learners. Additionally, because the highest exam score was 83.75, there were clearly content items within the assessment that were unfamiliar to all learners.

The independent variables that were captured and compared with the posttest scores were either calculated using analytics data from the LMS or from students' self-reported data, and all of them related to a different dimension of feedback. LMS data related specifically to dimensions of feedback were chosen to avoid what other researchers have determined to be a trivial focus on inconsequential learning analytics data (i.e., logins or page views) (Gasevic, Dawson, &

Siemens, 2015). The following sections provide a description of the findings for each feedback variable.

Frequency. First, when capturing the frequency of students' feedback dialogues with one another and the instructor, the results of this study revealed that four interactions per opportunity was most common. An opportunity for feedback could have been a discussion forum question, a quiz with automated responses, or an eLearning tutorial. Administrators were pleased with this finding because it showed that students' activities in the LMS matched or exceeded the expectation in most course discussion forums. The forums generally require an initial response from the learner plus two replies throughout the week. If every learner only interacted to the extent required, the average would be three, but the results for the (F1) frequency dimension showed that the highest number of learners, over 60, interacted an average of four times for any given opportunity. In reality, these data showed that the expectations of the course designers (i.e., three posts) were being fulfilled or exceeded. A similar finding occurred in research from Angeli, Valanides, and Bonk (2003) where learners posted an average of 5.6 responses when the required number was five.

Distribution. The second dimension of feedback that was quantified was distribution. The principles of "spacing or interleaving" have been well-documented as methods to improve cognitive processing and retention in education (Roediger & Pyc, 2012, p. 243). Learners who are experiencing effective distribution of content and practice across a learning experience are also receiving feedback at regular intervals. The timing of feedback interactions among learners was captured using LMS system data. Results showed a wide range of behaviors related to sporadic or consistent learner interaction. Almost 20% of learners followed a very consistent distribution routine. They interacted in every fifth of a course (across three courses) equivalently and consistently. There was little to no difference between the amount that they interacted in the first week of any course and the amount that they interacted at the end of every course.

For the most consistent (*well-distributed* related to feedback interaction behaviors) half of the learners, the difference between their week with the most feedback and their week with the least feedback was 14 or fewer interactions. Conversely, for the least consistently distributed (most cramming behaviors or *undistributed*) half of the learners, the differences between their least interactive and most interactive weeks was up to 78 interactions.

Stated another way, consider grouping the 206 learners into two halves- the *well-distributed* half and the *undistributed* half. Learners in the *well-distributed* half interacted (provided and received feedback) to almost the same degree in every week of their courses. For example, if they interacted 20 times in week one of their courses, they also interacted 20 times in weeks two, three, four, etc. The feedback interaction behaviors of this *well-distributed* half of the students showed a variance of 14 or less across the six weeks of a course.

In the other group, the *undistributed* half of learners in this research, their feedback interactions varied drastically from one week to the next. They may have interacted five times in week one of a course but 25 times in week five. For the fifty percent of the learners in the *undistributed* feedback half, the distribution of their feedback interactions varied from 15 to 78 meaning that on the very high end, some learners had 75+/- more interactions in some weeks of a course than in other weeks. Even though some students lacked distributed interactions, the analysis found that these students were no less likely to perform well on the posttest exam, and their satisfaction levels were also equivalent to their peers who chose to distribute their interactions more consistently across the five weeks of a course.

Individualized and content-related. The next dimension of feedback that was quantified was the number of individualized and content-related comments provided by instructors to students. The results showed that the most common number of comments students received, across three courses, was ten comments. This result seemed low to the academic leaders who reviewed the results. Once they realized that motivational comments (e.g., “good,” “great idea,” “really?”) did not qualify, ten comments across three courses seemed more feasible but still not ideal. Because the provision of individualized, content-related comments was the only variable that produced significance for both student satisfaction and achievement, leaders wanted to find new ways to prioritize faculty time around these efforts.

Timeliness. Finally, learners were satisfied with the timeliness of the feedback that they were receiving from faculty. There was not enough variance to adequately measure the impact of timely feedback on achievement and satisfaction. Differences among responses were measured to the hundredth of a point on a five point scale. To truly test this variable, a researcher would need a wider variety of timely feedback performance levels against which to compare student achievement and satisfaction.

Significant findings. None of the results described here were correlated significantly with posttest exam scores when they were analyzed as part of the multiple regression model. However, there were two statistically significant correlations when variables were compared using a Pearson correlation test. First, learners’ grades in the capstone course were significantly positively correlated with the posttest exam ($p<.05$). Within the same five-week timeframe, learners took the online course and also completed the external, standardized posttest exam. Students received a completion grade for taking the exam, but instructors did not see the exam scores until after their final grades were due. Students were also completing a capstone assignment for the course which was assessed by the instructor and comprised the bulk of the course grade earned by the student. The correlation between the end-of-course grade and the posttest exam scores was a good finding for the dean of the DeVoe School of Business. It showed that the course grades assigned by faculty were measuring some of the same outcomes as the externally validated posttest exam even if the gains were small, compared to the pretest.

The second significant relationship that was found was between the individualized, content-related (I_C) feedback students received from instructors in their capstone class and their scores on the posttest exam. Learners who received more content-related feedback in their final course also performed significantly better on the posttest exam ($p<.05$) than students who received fewer instances of I_C feedback. This result did not hold true when data from all three of students' final courses were added together; there was no significant correlation between the number of I_C feedback interactions a student had cumulatively across three courses and their posttest exam scores. The correlation was only significant within the final course. This result leads to questions about transfer of learning as well as students' motivation to apply the feedback since they were approaching an exam.

Overall, the results for Research Question One showed that the dimensions of feedback--timeliness, distribution, frequency, and individualized-content--together were not predictive of student achievement across three courses (9 credit hours). Because these variables aligned to the original feedback rubric (Table 1), there was little evidence to indicate that a course that scores well on the original feedback rubric as a whole would also produce higher student achievement. However, there were some results within this larger dataset that should continue to inform the design of online courses, and especially online courses that incorporate learning technologies.

The results indicate that individualized and content-related comments from instructors to students have a significant impact on learner achievement when the comments are provided in close proximity (i.e., within five weeks) of the learner's exam performance. There are instructional design implications in that designers should ensure that other course requirements do not prevent instructors from having the time to provide meaningful, individualized feedback. Also, instructional designers can provide job aides to facilitate content-specific feedback. For example, creating a bank of comments that instructors can copy, paste and personalize for a learner could improve their efficiency in providing feedback.

Descriptive Results for Research Question Two

Student evaluations of teaching (SETs) are sometimes met with skepticism (Zabaleta, 2007), so part of the intent for Research Question Two was to determine whether there was evidence that learners' responses about the curriculum and instruction could be validated by the system data collected. For example, if the instructor was indeed providing individualized and content-related (I_C) feedback on the student's papers, a dean would expect to see the student rate the instructor highly on the end-of-course (EOC) survey statement "The instructor's knowledge of course content was evident in the instruction of the course." The results showed that this was the case. There were significant correlations between the responses for every EOC question and the number of I_C comments students received from instructors. If a learner received more instances of I_C feedback, he/she was more satisfied with the course than a student who received fewer I_C comments. A rationale for this result is described in the next chapter.

Analytics data. As a method of collecting data for the research questions, it was also the intent of this study to leverage the readily available analytics that exist within the learning management system (LMS). LMS data were used to collect information on the timeliness, frequency, and distribution of students' feedback interactions. These data were not precise measurements of the constructs under analysis, but they were chosen intentionally, nonetheless, because of their ease of use and availability. The rationale was that the results of the research could be more practically utilized if the data sources were readily available for further, future use.

The lack of correlation between several of the feedback variables chosen for this research and learner achievement does not indicate that these dimensions of feedback are inconsequential. For example, one should not assume that because the regression analysis model was not statistically significant, timely delivery of feedback does not have a significant impact on achievement. Much research has been conducted to show that timeliness has a significant impact on achievement. Instead, the results of this study indicate that the methods used for quantifying the variables and the regression model were not sufficient to capture the impact of the feedback dimensions.

Chapter 5: Discussion

Introduction

There were two research questions that drove this research, both pertaining to the relationships between several dimensions of online feedback and learners' resulting achievement and satisfaction. Regression analysis provided an opportunity to measure whether it was the timeliness of feedback or the distributed nature of the feedback that had the most impact on student achievement. The multiple regression analysis revealed that frequency, distribution, individualization, and timeliness of feedback were not statistically significant predictors of achievement. Nevertheless, there were sub-findings that could contribute to future refinement and models. The findings to be discussed in this chapter include the following: (a) the impact of *personalization* in online instruction as measured by individualized, content-related comments delivered from instructors to learners, (b) evidence in support of *meaningful grading* practices among instructors in the research, (c) suggested improvements for LMS analytics data, and (d) course design improvement implications.

Interpretation of the Findings

Personalization. For this research, initially, data for five dimensions of feedback were captured using LMS analytics data and a survey. Analytics data are operationally defined as the data that are automatically collected by the LMS and provided to instructors via a tool dashboard or reports. Some data were also collected using manual approaches (opening online course pages and documents, counting, scoring, and recording information on a spreadsheet). Of the five dimensions under analysis, two dimensions were captured using LMS analytics data (frequency and distribution), one dimension was captured using self-report data from learners (timeliness), and data for the final two dimensions (individualization and content-specific comments) were manually captured and combined into one variable because they often occurred simultaneously. Even though the multiple regression analysis of these variables was not statistically significant, there was one statistically significant relationship from within the model. When students received individualized and content-related (I_C) comments on their papers from instructors within their capstone courses, it had a statistically significant positive effect on their achievement as captured by the posttest. It may not be coincidental that the I_C variable was the only dimension of feedback, in this research study, that was exclusively delivered by human instructors to learners. Connectivists would resonate with the idea that individualized and content-specific connections between instructors and learners in a course are essential (Goldie, 2016; Siemens, 2005).

Learners who received more I_C feedback instances were more likely to perform better on the posttest exam, and they were also more likely to be satisfied with the course. They rated both the curriculum and the instruction higher on the end-of-course (EOC) survey than students who experienced fewer instances of I_C feedback.

A growing body of research exists around the premise that online environments require some element of social presence, humanization, and critical interaction among people for effective learning to occur (Garrison & Cleveland-Innes, 2005). An asynchronous course can easily become a “robotic set of instructions” rather than a “dynamic learning environment” (Bickle & Rucker, 2018). Without social interaction and/or internal reflective dialogue, one might question how learning could occur at all. In their meta-analysis of the research on social presence as it relates to online learning satisfaction, Richardson, Maeda, and Caskuru (2017) found that there are strong positive relationships between social presence, satisfaction, and learners’ perceptions of their learning. Authors remarked at the multi-dimensional nature of social presence and at the variety of differences that exist for defining the construct, let alone quantifying it (Richardson et al., 2017).

For the purpose of facilitating future research, the operationalized definition of *social presence* that best informs the implications of this research comes from Gunawardena and Zittle (1997) where they assert that social presence is “the degree to which a person is perceived as ‘real’ in mediated communication” (p. 8). All of the communication in this research was occurring asynchronously using computer mediated systems. Comments provided by instructors, on students’ essays and projects counted for the calculation of the I_C variable if they were either: (a) individually personal to the learner (i.e., related to their prior work, current interests, personal strengths and needs, or future goals), and/or (b) directly related to the content or topic. Comments could be *content-related* but not *individualized* and vice versa. If comments were both *content-related* and *individualized*, they were counted twice. One litmus test that was used to further define individualization was that if the comment could be applied to any student, it was not *individualized*.

Social presence was not a construct intended for study by this research, but it seems potentially important that the only exclusively human dimension of the variables chosen for this research was the one dimension that produced a significant impact on both student achievement and satisfaction. Was it the individualized nature of the feedback that produced the impact, or was it the origin of the feedback--an expert instructor? This question remains unaddressed by this study.

Meaningful grades. College faculty and administrators are well aware of issues with grade inflation and disproportionate grade distributions. Especially in upper-level courses, capstone courses, and graduate courses, learners seem to expect an A grade if they complete the assignments on time. In the experience of the researcher, when administrators are reviewing assessment and evaluation data for academic program reviews, it is not uncommon to hear questions around grade distribution. Felton and Koper begin their analysis of the issues around grades by stating that “grade inflation is a longstanding problem whose seriousness is demonstrated by a wide variety of studies of grade distributions” (2005, p. 561).

The DeVoe School of Business (DSB) is no different in that the most common grade for the capstone courses is an A. Program directors and administrators are left to wonder what the grade actually represents. Additionally, part of the purpose of this research was to inform the leaders of the DSB regarding the quality of learners’ educational experiences. By quantifying and reporting the results of the data collection around feedback variables, school leaders gained a clearer picture of what learners were experiencing (i.e., the timeliness, distribution, frequency, etc. of the feedback they receive). Leaders learned that the end-of-course grades earned by learners in their capstone courses did have a significant positive correlation with the posttest exam scores. Grade inflation may still be a concern, but these data provided administrators with some level of reassurance that faculty instructors were grading in ways that mirror the externally validated assessment. Perhaps faculty who teach the capstone course have an awareness of the topics students will see on the posttest and therefore assess coursework with those criteria in mind. Although it was not a central part of the research questions posited originally, this finding was informative for the participants.

Usefulness of LMS data. The multiple regression model that was designed for this research was not at all predictive of student achievement and satisfaction. Why? Theoretically, the dimensions of feedback chosen for inclusion in the model have a strong base of research that would indicate that each feedback variable chosen for this study should, at least to some degree, have been partially responsible for aspects of learner achievement and satisfaction.

A very similar study did produce a statistically significant regression model (Jo, Yu, Lee, & Kim, 2015). In Jo et. al.'s research, learning analytics data were used as predictors to explain student achievement using course grades (2015). They found that they were able to predict over 90% of a student's course grade using variables like login frequency and the time spent in the online course.

It was discussed earlier, as a limitation, that the posttest exam used to measure learner achievement (the dependent variable) was not sufficiently aligned to course content. Moreover, even if aligned, it still may not have been sufficiently sensitive to capture student learning. The researcher believes that this was the primary reason that the regression model was not significant. Other explanations, however, do exist.

A confounding variable to consider is the degree to which LMS analytics data can be relied upon as a valid source of data to quantify the independent variables related to feedback. LMS analytics data were chosen intentionally in full awareness of the lack of data integrity that may accompany them. They are readily available at the instructor level, so it would be ideal to find ways to make valid interpretations from these data. The results of this research only serve to further illuminate the immaturity of the LMS analytics currently available in popular LMSs (e.g., Canvas, Blackboard, Brightspace, and Moodle). It is as though learning technology providers have presented analytics that are available to them as data scientists rather than analytics that are genuinely meaningful for teaching and learning.

Firat (2016) defines learning analytics as a focus on "reaching patterns or tendencies via data sets related to students… to maintain the development of supplementary and personalized higher education systems” (p. 76). Part of the aim of this research was to discover patterns in the LMS learning analytics data related to students’ dialogic feedback behaviors in the discussion forum tool. The LMS (D2L’s Brightspace) provides metrics such as the number of replies posted, the number of posts read, the percentage of course pages viewed, the amount of time spent within the course each day, the number of days each week or each month that the learner visited the course, etc. If it could be shown that certain metrics or benchmarks are highly correlated with certain levels of student performance, instructors could make use of the analytics by intervening with students. Likewise, instructional designers could structure course environments to increase the likelihood of the desired behaviors and decrease occurrences of the undesired situations. LMS providers and researchers alike have touted the myriad benefits and potential of such analytics data (Firat, 2016).

While learning analytics data from popular LMSs have been used successfully by many learning analytics researchers (as is evidenced by perusing the session titles at the Learning Analytics and Knowledge conference each year), at least at the institution involved in this research, online instructors have yet to realize the benefits of having these data just a click away.

The easily accessible metrics such as *time spent in the LMS* appear to have no impact on overall learner achievement (Castano-Munoz, Duart, & Sancho-Vinuesa, 2014). A predictive model, similar to the one designed for this research, was carried out in a research study involving 41 undergraduate students (Jo, Yu, Lee, & Kim, 2015). It revealed that variance in students’ course grades could be explained with learning analytics variables (i.e., total logins, login intervals, total assignments submitted, etc.). However, as reported earlier, course grades alone are not the best indicators of student achievement; grade inflation issues challenge our notions of what grades actually represent. In essence, the learning analytics in Jo et al.’s (2015) research were being employed to track the degree to which the path charted for the learner was being followed, virtually guaranteeing a positive correlation between the analytics variables and the course grade. The actual learning that occurred is unknown.

A robust use of learning analytics, in the future, would help faculty and instructional design teams gain insight into the effectiveness of the design of a course learning environment. Data would reflect situational variables that are known to improve learning. Analytics could provide evidence of learners' engagement in thinking routines (Ritchart, Church, & Morrison, 2011). For example, what if questions could be programmed as 'pop-ups' to ask learners about course design features? Learners would have to answer before continuing, and questions could be programmed to ask things like, "Did you find the feedback for this question helpful?" or "Did you have enough prior knowledge to understand the article you just read?" Collecting hundreds of these bits of information and then comparing them with other variables (like time on the page, average number of pages visited, etc.) could help instructional designers make more meaningful changes to learning experiences. Research of the future might ask: is there a correlation between students who say "no, the feedback was not helpful" and students who perform poorly on assessments? Unfortunately, at the present time, LMS learning analytics are in their nascent stages and in need of enhancement and refinement.

With regard to the specific analytics chosen for this study, by using the LMS data, researchers had to make assumptions and use data in ways that pushed the boundaries for its intended uses. For example, the researchers counted "number of posts read" as dialogic feedback interactions (frequency) even though students could click "read all posts" and automatically skew this number. Also, there was no way to sort or differentiate student-student feedback from instructor-student feedback in the discussion forums without manually opening each discussion thread, so differentiating by source was not possible. The researcher also lacked a reliable method, using the analytics, to see the timeliness of feedback, and therefore had to rely on students' self-reports. All of these issues likely lead to imprecise measures of the independent variables identified for the research.

Implications for Theory and Practice

A new paradigm of education recognizes that learner-centered instruction is a worthy 21st century goal (Reigeluth, Beatty, & Myers, 2016). Intentionality around the ways learners receive and use feedback is essential in learner-centered instructional design (Dawson et al., 2018). In this research study, the comments written by instructors on students' papers proved to be of most significance, providing positive correlations with both student achievement and satisfaction ($p<.05$). Even though the predictive model that was tested did not produce a statistically significant impact on achievement, this could have been caused by issues with the posttest exam as the dependent variable rather than a lack of relationship among the variables. Other research has certainly established the value of timely, frequent, and distributed feedback, so course designers and instructors should continue to incorporate all of the dimensions of feedback.

Recommendations for faculty instructors. In particular, special emphasis should be placed on providing opportunities for learners to receive individualized and content-specific feedback from the instructor. The most common challenge associated with the delivery of these kinds of feedback is that it is time consuming for the instructor, and some question the degree to which students read or use the feedback. Instructional designers and faculty should leverage automation, peer feedback, and/or adaptive learning technology to provide formative feedback during key timeframes in the course when instructors will be heavily engaged in providing written or video feedback that is individualized and moves the learning forward. Faculty instructors who are struggling with low end-of-course survey evaluations should implement strategies to provide feedback to learners that connects directly to them as individuals and serves to move their learning forward. Furthermore, the design of the course or series of courses should ask students to use the feedback they received. The following list of starters was adapted from *Making Thinking Visible* and may prove useful for faculty instructors who wish to provide individualized and content-related comments on student work (Ritchhart, Church, & Morrison, 2011).

1. Can you identify any patterns from your classmates' submissions for this project? OR
   I'm noticing a pattern in your work…
2. What generalization could you make from the specific examples you provided? OR
   Thanks for supporting your generalization about …. with…
3. What alternative possibilities could there be to explain the result you identified? OR
   A more relevant alternative would be…
4. From the evidence you presented, what is the most convincing piece? OR
   Try supporting your argument with more reliable evidence.
5. What is your plan to continue to grow in this area? OR
   A protocol is a series of steps you follow to improve consistency. What protocol could you develop for yourself to improve your writing skills?
6. What are the primary knowledge claims you are presenting? OR
   There might be an implicit bias or assumption that is impacting your thinking here.
7. From what you have identified, what are the priorities and how do you know? OR
   From what you wrote, it appears that your conditions for knowing something are …. Is that accurate?

Strategies to provide individualized comments include collecting knowledge about the learner. Any understanding about students' personal goals, present circumstances, or future dreams can prove valuable as instructors work these connections into their feedback comments. If the prevailing instructional theories of the 21st century continue to espouse socially constructed learning as well as connectivist learning, instructional designers and faculty will need to become more adept at capitalizing on the efficiencies that technology offers so that humans can do what we do best: connect with one another.

Generalizations. All of the data for this research study were collected from one institution in the Midwest. The demographics were reflective of the Midwestern region of the United States. Participants were 206 randomly selected adult, online learners studying for business degrees at the associate, bachelor's, or master's level. Courses were designed by teams of instructional designers and faculty subject matter experts. Then, the collaboratively developed master courses were copied and delivered by faculty instructors. Courses were taken by students one-at-a-time, and each course was 5-6 weeks in length. Caution should be exercised when generalizing results to other institutions, especially where the instructional design model differs.

Limitations

Similar to most correlational research, one limitation of this study is that there are a variety of explanations for any finding. A relationship was observed between I_C feedback and student achievement and satisfaction, but other variables like the length of the assignment submitted could have confounded the results. If a student writes more, will they necessarily receive more feedback comments? And, therefore, is it the length of the assignment or the feedback that produced that improved learning and satisfaction? Are students who produce longer assignments just more likely to be satisfied and to learn more? These questions remain unanswered.

A second limitation relates to the diversity of the delivery modality. All of the courses analyzed were delivered online in a time-based, structured course design setting. Competency-based education that includes self-pacing, hybrid design that includes face-to-face instruction, or a course design that includes more unstructured learning could all produce significantly different results.

Response bias was not a limitation for this study since learners' data were collected after they had completed their academic programs without their knowledge. The major challenge revolved around the Peregrine posttest exam results. Because the exam was not well-aligned to coursework topics, it was not a sufficient dependent variable.

Recommendations for Future Research

With the continual increases in online learning (including hybrid and blended) educators of all kinds, K-12 and higher education will need to devise ways to understand more about learners' strengths and needs before the learner is held accountable. In higher education, because of the many millions of dollars invested in financial aid for college learners, the U.S. Department of Education needs effective methods to prevent bad actors from taking students' funds without providing an education.

Researchers and data scientists need to work in collaboration to design learning analytics that reflect behaviors that accurately predict success. Research is needed to discover the patterns in online learning environments that typically lead to student success and the patterns that should cause instructors and advisors to intervene. These patterns may include dimensions of feedback, but they may also include students' self-reflections or inventory assessments. Regardless, the data currently presented to instructors through LMS analytics dashboards do not appear to provide much predictive insight into attainment of learning outcomes.

In addition to such studies of student success patterns, design studies are needed to further understand the impact and implications of individualized, content-related feedback on student achievement and satisfaction. Why was the effect of I_C feedback apparently not cumulative in this study? It was only significant in the final course. What aspects of I_C feedback are generalized/transferred by students into new contexts? Are there particular categories or elements of I_C feedback that are more effective than others? To what extent does it matter that the comments originate with a human author? If students thought that the comments came from artificial intelligence, would they be perceived or acted upon differently?

Third, there is a need for a series of design studies to further understand the usefulness of readily available LMS analytics data. Because this model was not predictive, is there a model that could be? What if additional variables were included like a measure of student motivation or prior GPA? It would be worth an investment of funds for institutions to find the invisible online student who disappears. Cost of acquisition for each online learner is currently around $1,700 at the institution in the study. Keeping more of those students by monitoring and intervening effectively creates a fast return on investment in the monitoring and intervention systems.

Finally, design studies might be undertaken with more precise measures of the distribution dimension of feedback to determine whether there is an impact to distributed interaction for the adult, online learner. This research presents some initial evidence that there was no correlation between students who did all of their work in the final week of a five week course and students who worked consistently across all five weeks. This finding is valuable because it indicates that current distance education policies around regular and substantive interaction may be either ill-informed or misunderstood in implementation.

Regular and Substantive Interaction. Currently, at the institution involved in the research, a huge amount of effort is expended to hold students accountable to contributing substantially every single week during an accelerated, online course. Perhaps this is wasted effort. Perhaps this level of consistency for an adult, online learner (average age- 36) is not necessary. Perhaps adult learners are more capable than children or young adults of understanding when and where they need to consistently acquire knowledge and skill and when they can take longer breaks and then engage in deep ways over shorter periods of time. These institutional policies around regular and substantive interaction exist less because of instructional design choices of the people closest to the teaching and learning and more because of federal policy around delivery of distance education.

Federal government policymakers have long struggled to satisfy the divergent needs of institutions and employers in the postsecondary market. Quality assurance is needed to keep "bad actors" out of the higher education marketplace, but employers see an ever-widening skills gap, and they need a skilled workforce (Jaschik, 2015). In brief, financial aid is disbursed for distance education only to institutions that can prove regular and substantive contact with students. Regular and substantive contact means that students have consistent discipline-focused interaction with a qualified faculty member (qualified according to the regional or specialized accrediting body regulations), that is initiated by the faculty member. The exact frequency is not defined beyond "regular" (Laitinen, 2012).

When the rules for regular and substantive interaction were written, there were no multi-million dollar cognitive tutoring educational technology companies (Laitinen, 2012). Pressures to lower the cost of higher education to fill the workforce skills gaps produce reliance on technology to offer some of a student's feedback needs. Leveraging technology could reduce cost without sacrificing quality (Prensky, 2016; Reigeluth, Myers & Lee 2016). In higher education in the United States, perhaps federal distance education policies related to regular and substantive interaction limit the degree to which innovations can impact learning designs.

New methods of measuring the quality of a learner's higher education experience are needed. Policymakers have relied on contact hours and units of time as proxies for measures of learning for far too long (Laitinen, 2012). As long as the regular and substantive rules remain in force, and as long as the Office of the Inspector General continues to recommend fines to institutions like Western Governor's University ("Western Governors University," 2017), for neglecting regular and substantive interaction, innovations that include AI, adaptive learning technologies, and intelligent tutoring will remain in their experimental and nascent stages.

In short, employers are dissatisfied with college graduates' skills but lack a better alternative (Jaschik, 2015); colleges are dissatisfied with outdated regulations that mandate time as a proxy for learning; faculty are dissatisfied with scattered and burdensome assessment requirements; and students are dissatisfied with loads of debt and uncertainty about whether or not their degree will help them achieve their goals. Federal regulators are trying to protect students while also making provision for the innovation that they know is necessary. Scholars, philanthropists, public policymakers, and learners themselves are calling for change (Selingo, 2013; McGee, 2015; Carey, 2016; Engle, 2016; Gallagher, 2016).

An Emerging Solution. This research certainly does not propose an answer to the vastly complicated web of challenges described in the prior section, but it does propose an emerging solution to consider- an idea worthy of investigation and further research. First, measuring each dimension of a learner's feedback experience is a better proxy for quality than measuring regular and substantive contact. Second, measuring and communicating the dimensions of the learner feedback experience could provide educational technology companies with a competitive edge when other providers don't offer the same level of implementation support. The dimensions described herein work in concert with one another to produce learner-centered instructional experiences. Third, there is some evidence that requiring adult, online learners to interact in an online space every three or four days does not produce measurable differences in achievement or satisfaction compared with learners who choose to interact on a less consistent basis. If capturing regular and substantive interaction data, as a metric, does not lead to success, designers should instead capture the frequency and timeliness of the individualized and content-specific feedback a learner receives along with learners' satisfaction around their ability to make use of that feedback.

Conclusion

As learning management systems and learning analytics tools become easier to integrate, more essential to managing online learning, and more widely implemented, users will need to avoid temptations to make assumptions about course quality based on easily accessible metrics. For example, it might be easy for an instructor to assume that a student who has contributed 14 times to the discussion forum is necessarily learning more than a student who contributes only the requisite three times. The results of this research do not support that assumption. Second, the impact of individualization in online instruction should not be under-estimated. As instructional designers partner with subject-matter experts to create learning experiences, leveraging adaptive learning tools that provide immediate feedback should be done with an intent to free up instructor time so that instructors have ample opportunity to provide individualized and content-specific feedback to learners.

The explanation for the rationale behind this research began with a description of the iron triangle of instructional design as proposed by Honebein and Honebein (2015). It was discussed that instructional designers experience a microcosm of the same tensions being played out on a national stage in the United States—the tension between effectiveness, efficiency and appeal. Many questions remain, but the results of this study reinforce the importance of high-quality feedback for effective and appealing learning experiences.

References

Adkins, S. (2016). *2015 International learning technology investment patterns*. Ambient Insight. Monroe, WA.

Amory, A. (2010). Education technology and hidden ideological contradictions. *Educational Technology & Society*, *13*(1), 69–79. Retrieved from https://www.j-ets.net/ETS/journals/13_1/8.pdf

Anderson, L. W. (1979). Adaptive education. *Educational Leadership*, *37*(2), 140. Retrieved from http://www.ascd.org/ASCD/pdf/journals/ed_lead/el_197911_anderson.pdf

Angeli, C., Valanides, N., & Bonk, C. J. (2003). Communication in a web-based conferencing system: The quality of computer-mediated interactions. *British Journal of Educational Technology*, *34*(1), 31–43. https://doi-org.proxyiub.uits.iu.edu/10.1111/1467-8535.00302

Barr, R. B., & Tagg, J. (1995). From teaching to learning: A new paradigm for under-graduate education. *Change*, *27*(6), 13-25.

Bickle, M. C., & Rucker, R. (2018). Student-to-student interaction: Humanizing the online classroom using technology and group assignments. *Quarterly Review of Distance Education*, *19*(1), 1–11. Retrieved from http://proxyiub.uits.iu.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ1190065&site=eds-live&scope=site

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7–74. https://doi.org/10.1080/0969595980050102

Boud, D., & Molloy, E. (2013). Rethinking models of feedback for learning: The challenge of design. *Assessment & Evaluation in Higher Education*, *38*(6), 698–712. https://doi.org/10.1080/02602938.2012.691462

Brown, J., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, *18(1)*, 32–42.

Calvert, C. E. (2014). Developing a model and applications for probabilities of student success: A case study of predictive analytics. *Open Learning: The Journal of Open, Distance and e-Learning*, *29*(2), 160–173. https://doi.org/10.1080/02680513.2014.931805

Carey, K. (2016). *The end of college: Creating the future of learning and the University of Everywhere*. (First Riverhead trade paperback edition). New York, NY: Riverhead Books.

Carless, D. (2018). Feedback loops and the longer-term: Towards feedback spirals. *Assessment & Evaluation in Higher Education*, https://doi.org/10.1080/02602938.2018.1531108

Carless, D., Salter, D., Yang, M., & Lam, J. (2011). Developing sustainable feedback practices. *Studies in Higher Education*, *36*(4), 395. Retrieved from https://srhe.tandfonline.com/doi/abs/10.1080/03075071003642449

Carroll, J. B., & Spearritt, D. (1967). A study of a “Model of school learning”: Monography number 4, 18. Retrieved from http://eric.ed.gov/?id=ED045477

CAST. (2011). *Enhance capacity for monitoring progress - UDL guidelines examples* [Education]. Retrieved from https://sites.google.com/site/udlguidelinesexamples/home/provide-multiple-means-of-action-and-expression/provide-options-for-executive-functions/enhance-capacity-for-monitoring-progress

Castaño-Muñoz, J., Duart, J. M., & Sancho-Vinuesa, T. (2014). The Internet in face-to-face higher education: Can interactive learning improve academic achievement? *British Journal of Educational Technology*, *45*(1), 149-159.

Chevalier, R. D. (2011). When did ADDIE become addie? *Performance Improvement*, *50*(6), 10–14. https://doi.org/10.1002/pfi.20221

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

A Conversation with Benjamin Bloom. (1979). *Educational Leadership*, *37*(2), 157. Retrieved from http://proxyiub.uits.iu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=7731326&site=eds-live&scope=site

Crisp, E. A., & Bonk, C. J. (2018). Defining the learner feedback experience. *TechTrends, 62*(6*),* 585-593. https://doi.org/10.1007/s11528-018-0264-y

Dawson, P., Henderson, M., Mahoney, P., Phillips, M., Ryan, T., Boud, D., & Molloy, E. (2018). What makes for effective feedback: Staff and student perspectives. *Assessment and Evaluation in Higher Education, 43*(4), 1–12. https://doi.org/10.1080/02602938.2018.1467877

Dennen, V. P., Darabi, A. A., & Smith, L. J. (2007). Instructor-learner interaction in online courses: The relative perceived importance of particular instructor actions on performance and satisfaction. *Distance Education*, *28*(1), 65–79. https://doi.org/10.1080/01587910701305319

Dirlam, D. K. (2017). *Teachers, learners, modes of practice: Theory and methodology for identifying knowledge development*. UK: Routledge.

Engle, J. (2016). *Answering the call: Institutions and states lead the way toward better measures of postsecondary performance* (p. 31). Bill & Melinda Gates Foundation.

Felton, J., & Koper, P. T. (2005). Nominal GPA and real GPA: A simple adjustment that compensates for grade inflation. *Assessment and Evaluation in Higher Education, 30*(6), 561–569. Retrieved from http://proxyiub.uits.iu.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ718849&site=eds-live&scope=site

Firat, M. (2016). Determining the effects of LMS learning behaviors on academic achievement in a learning analytic perspective. *Journal of Information Technology Education*: *Research*, *15*, 75-87. Retrieved from http://www.jite.org/documents/Vol15/JITEv15ResearchP075-087Firat1928.pdf

Francom, G. M. (2016). Principles for task-centered instruction. In C. M. Reigeluth, B. J. Beatty, & R. D. Myers (Eds.), *Instructional-design theories and models* (Vols. 1–4, pp. 65–92). Hillsdale, NJ: Routledge.

Fraser, B. J., Walberg, H. J., Welch, W. W., & Hattie, J. A. (1987). Syntheses of educational productivity research. *International Journal of Educational Research*, *11*(2), 147–252.

Gallagher, S. R. (2016). *The future of university credentials: new developments at the intersection of higher education and hiring*. Cambridge, MA: Harvard Education Press.

Garrison, D. R., & Cleveland-Innes, M. (2005). Facilitating cognitive presence in online learning: Interaction is not enough. *American Journal of Distance Education*, *19*(3), 133-148. https://doi:10.1407/s15389286ajde1903_2

Gašević, D., Dawson, S., & Siemens, G. (2015). Let’s not forget: Learning analytics are about learning. *TechTrends*, *59*(1), 64–71. https://doi.org/10.1007/s11528-014-0822-x

Gibbs, G., & Simpson, C. (2004). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education*, (1). Retrieved from https://www.open.ac.uk/fast/pdfs/Gibbs%20and%20Simpson%202004-05.pdf

Ginder, S.A., Kelly-Reid, J.E., & Mann, F.B. (2017). *Enrollment and employees in postsecondary institutions, Fall 2016; and financial statistics and academic libraries, fiscal year 2016: First look* (NCES 2018- 002). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved from http://nces.ed.gov/pubsearch.

Goldie, J. G. S. (2016) Connectivism: A knowledge learning theory for the digital age? *Medical Teacher, 38(*10), 1064-1069. https://doi:10.3109/0142159X.2016.1173661

Gunawardena, C. N., & Zittle, F. J. (1997). Social presence as a predictor of satisfaction within a computer-mediated conferencing environment. *The American Journal of Distance Education*, *11*(3), 8-26. https://doi.10.1080/08923649709526970

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis* (7th ed.). Essex, UK: Pearson.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Herbert, M. (2006). Staying the course: A study in online student satisfaction and retention. *Online Journal of Distance Learning Administration*, *9*(4). Retrieved from https://www.westga.edu/~distance/ojdla/.

Honebein, P. C., & Honebein, C. H. (2015). Effectiveness, efficiency, and appeal: Pick any two? The influence of learning domains and learning outcomes on designer judgments of useful instructional methods. *Educational Technology Research and Development*. https://doi.org/10.1007/s11423-015-9396-3

Hounsell, D. (1987). Essay writing and the quality of feedback. In J. Richardson, M. W. Eysenck, & D. W. Piper (Eds.), *Student learning: Research in education and cognitive psychology* (pp. 109–119). Milton Keynes, UK: Open University Press.

Huber, P. J. (1981). *Robust statistics*. New York, NY: John Wiley & Sons.

Integrated Postsecondary Education Data System. (n.d.). *IPEDS Access Databases.* Washington, DC: National Center for Education Statistics.

Jaehnig, W., & Miller, M. (2007). Feedback types in programmed instruction: A systematic review. *The Psychological Record*, *57*, 219–232.

Jarrett, J. (2013). Bigfoot, goldilocks, and moonshots: A report from the frontiers of personalized learning. *EDUCAUSE Review*, *48*(2), 30-34, 36-40. Retrieved from http://er.educause.edu/~/media/files/article-downloads/erm1323.pdf

Jaschik, S. (2015, January 20). Well-prepared in their own eyes. *Inside Higher Education*.

Jo, I. H., Yu, T., Lee, H., & Kim, Y. (2015). Relations between student online learning behavior and academic achievement in higher education: A learning analytics approach. In G. Chen, V. Kumar, Kinshuk, R. Huang, & S. C. Kong (Eds.), *Emerging issues in smart learning,* (pp. 275-287). Berlin Heidelberg: Springer.

Johnson, G. M. (2004). Constructivist remediation: Correction in context. *International Journal of Special Education*, *19*, 72-88.

Johnson, G. M. (2014). The ecology of interactive learning environments: Situating traditional theory. *Interactive Learning Environments*, *22*(3), 298–308. https://doi.org/10.1080/10494820.2011.649768

Jonassen, D. H. (1991). Objectivism versus constructivism: Do we need a new philosophical paradigm? *Journal of Educational Research*, *39*(3), 5–14.

Judge, T. A., Piccolo, R. R., Podsakoff, N. P., Shaw, J. C., & Rich, B. L. (2010). The relationship between pay and job satisfaction: A meta-analysis of the literature. *Journal of Vocational Behavior*, *77*, 157–167. https://doi.org/10.1016/j.jvb.2010.04.002

Kalaitzidis, T. J., Litts, B., & Halverson, E. R. (2016). Designing collaborative production of digital media. In *Instructional-design theories and models* (Vol. 4, pp. 173–204). Hillsdale, NJ: Routledge.

Karagiorgi, Y., & Symeou, L. (2005). Translating constructivism into instructional design: Potential and limitations. *Educational Technology and Society*, *8*(1), 17–27. Retrieved from http://www.jstor.org/stable/jeductechsoci.8.1.17

Keller, F. S. (1974). Ten years of personalized instruction. *Teaching of Psychology*, *1*(1), 4–9. https://doi.org/10.1177/009862837400100102

Keller, F. S. (1985). Lightning strikes twice. *Teaching of Psychology*, *12*(1), 4–8. Retrieved from http://proxyiub.uits.iu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ316017&site=eds-live&scope=site

Kluger, A. (1989). Negative and positive effects of feedback on motivation: The role of control and individual differences. Presented at Society for Industrial and Organizational Psychology, Boston, 1989. Boston, MA.

Kluger, A., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*(2), 254–284. https://doi.org/0033-2909/96

Koedinger, K., & Corbett, A. (2006). Cognitive tutors: Technology bringing learning science to the classroom. In R. K. Sawyer (Ed.) *The Cambridge handbook of: The learning sciences*. New York, NY: Cambridge University Press. Retrieved from http://psycnet.apa.org/record/2006-07157-005

Konert, J., Gobel, S., & Steinmetz, R. (2012). Toward social serious games. In *Proceedings of the 6th European conference on games based learning* (p. 653). Waterford Institute of Technology, Ireland: Thomson ISI. Retrieved from: ftp://ftp.kom.tu-darmstadt.de/papers/KGS12-1.pdf

Kopcha, T. H. & Sullivan, H. (2006). Self-presentation bias in surveys of teachers' educational technology practices. *Educational Technology Research and Development, 55,* 627-646. https://doi.org/10.1007/s11423-006-9011-8

Kreighbaum, A. (2018). More pressure on accreditors in changing higher ed landscape. *Inside Higher Ed*. Retrieved from https://www.insidehighered.com/news/2018/08/15/accreditors-pressured-apply-more-scrutiny-profits-look-change-tax-status

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77*(6), 1121–1134. https://doi.org/10.1037/0022-3514.77.6.1121

Kulhavy, R. W. (1977). Feedback in written instruction. *Review of Educational Research*, *47*(2), 211–232. https://doi.org/10.2307/1170128

Kulik, C. C, Kulik, J. A., & Bangert-Drowns, R. L. (1990). Effectiveness of mastery learning programs: A meta-analysis. *Review of Educational Research*, *60*(2), 265–299.

Laerd Statistics. (2015). Paired-samples t-test using SPSS statistics. *Statistical tutorials and software guides*. Retrieved from https://statistics.laerd.com

Laitinen, A. (2012). *Cracking the credit hour* (Policy Paper) (p. 12). Washington, DC: New America Foundation.

Lea, M. R., & Street, B. (2000). Student writing and staff feedback in higher education: An academic literacies approach. In M. R. Lea & B. Stierer (Eds.), *Student writing in higher education: New contexts*. Buckingham, UK: Open University Press.

Lieberman, M. (2018). Furor over blended and active learning. Retrieved from https://www.insidehighered.com/digital-learning/article/2018/09/21/blended-learning-model-university-central-florida-draws-business

Marsh, E. J., Lozito, J. P., Umanath, S., Bjork, E. L., & Bjork, R. A. (2012). Using verification feedback to correct errors made on a multiple-choice test. *Memory*, *20*(6), 645–653. https://doi.org/10.1080/09658211.2012.684882

Mayer, R. E. (2009). Constructivism as a theory of learning versus constructivism as a prescription for instruction. In S. Tobias & T. M. Duffy (Eds.), *Constructivist instruction: Success or failure?* (pp. 184-200). New York, NY, US: Routledge/Taylor & Francis Group.

McChesney, C., Covey, S., & Huling, J. (2016). *The 4 disciplines of execution: Achieving your wildly important goals*. New York, NY: Free Press.

McGee, J. (2015). *Breakpoint: The changing marketplace for higher education*. Baltimore, MD: Johns Hopkins University Press.

McGuire, B. F. (2016). Integrating the intangibles into asynchronous online instruction: Strategies for improving interaction and social presence. *Journal of Effective Teaching*, *16*(3), 62–75. Retrieved from https://eric.ed.gov/?id=EJ1125809

McLeod, S. (2018, March 10). Edward Thorndike - law of effect: Simply psychology [Educational]. Retrieved from https://www.simplypsychology.org/edward-thorndike.html

Merrill, M. D. (2013). *First principles of instruction: Identifying and designing effective, efficient and engaging instruction*. San Francisco, CA: Pfeiffer.

Mikulincer, M. (1994). *Human learned helplessness: A coping perspective*. New York, NY: Plenum Press.

Mohammadi, H. (2015). Investigating users’ perspectives on e-learning: An integration of TAM and IS success model. *Computers in Human Behavior*, *45*, 359–374. https://doi.org/10.1016/j.chb.2014.07.044

Moore, J. W., & Smith, W. I. (1964). Role of knowledge of results in programmed instruction. *Psychological Reports*, *14*, 407–423.

Moore, M. G. (1989). The types of interactions. *The American Journal of Distance Education*, *3*(2), 1–6.

Murray, M. C., & Perez, J. (2015). Informing and performing: A study comparing adaptive learning to traditional learning. *Informing Science: The International Journal of an Emerging Transdiscipline*, *18*(1), 111-125.

Nicol, D., Thomson, A., & Breslin, C. (2014). Rethinking feedback practices in higher education: A peer review perspective. *Assessment and Evaluation in Higher Education*, *39*(1), 102–122. https://doi.org/10.1080/02602938.2013.795518

Pennebaker, J. W., Gosling, S. D., & Ferrell, J. D. (2013). Daily online testing in large classes: Boosting college performance while reducing achievement gaps. *PLOS ONE*, *8*(11). https://doi.org/10.1371/journal.pone.0079774

Peregrine Academic Services. (2018). *Assessment service validity and reliability* (Report No. 2018-1). Gillette, WY.

Prensky, M. (2016). *Education to better their world: Unleashing the power of 21st-century kids*. Teachers College Press.

Pressey, S. L. (1950). Development and appraisal of devices providing immediate automatic scoring of objective tests and concomitant self-instruction. *The Journal of Psychology*, *29*(2), 417–447. Retrieved from https://doi-org.proxyiub.uits.iu.edu/10.1080/00223980.1950.9916043

Pysh, F., Blank, S., & Lambert, R. (1969). The effects of step size response mode and KOR upon achievement in programmed instruction. *Canadian Psychologist*, *10*(1), 49–64.

Reigeluth, C. M., Beatty, B., & Myers, R. (2016). *Instructional-design theories and models* (Vol. IV). Hillsdale, NJ: Lawrence Erlbaum Associates.

Reigeluth, C. M., Myers, R. D., & Lee, D. (2016). The learner-centered paradigm of education. In C. M. Reigeluth, B. J. Beatty, & R. D. Myers (Eds.), *Instructional-design theories and models* (Vol. 4, pp. 5–32). Hillsdale, NJ: Routledge.

Reupert, A., Maybery, D., Patrick, K., & Chittleborough, P. (2009). The importance of being human: Instructors' personal presence in distance programs. *International Journal of Teaching and Learning in Higher Education*, *21*(1), 47–56. Retrieved from http://eric.ed.gov/?id=EJ896241

Rha, I. (1988). *The role of CAI feedback in acquiring declarative and hierarchical knowledge in the context of social science concept learning* (Doctoral dissertation). Retrieved from ProQuest. (303570064)

Richardson, J. C., Maeda, Y., Lv, J., & Caskurlu, S. (2017). Social presence in relation to students' satisfaction and learning in the online environment: A meta-analysis. *Computers in Human Behavior, 71*, 402-417. Retrieved from https://doi.org/10.1016/j.chb.2017.02.001

Ritchhart, R., Church, M., & Morrison, K. (2011). *Making thinking visible: How to promote engagement, understanding, and independence for all learners* (1st ed.). San Francisco, CA: Jossey-Bass.

Roediger, H. L., & Pyc, M. A. (2012). Inexpensive techniques to improve education: Applying cognitive psychology to enhance educational practice. *Journal of Applied Research in Memory and Cognition*, *1*(4), 242–248. https://doi.org/10.1016/j.jarmac.2012.09.002

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*(2), 119–44.

Selingo, J. J. (2013). *College (un)bound: The future of higher education and what it means for students*. Las Vegas: Amazon Publishing.

Siemens, G. (2005). Connectivism: A learning theory for the digital age. *International Journal of Instructional Technology and Distance Learning, 2*(1). Retrieved from http://www.itdl.org/Journal/Jan_05/article01.htm

Simon, H. A. (1995). The information-processing theory of mind. *American Psychologist*, *50*(7), 507–508. https://doi.org/10.1037/0003-066X.50.7.507

Smith, P. L., & Dillon, C. L. (1999). Comparing distance learning and classroom learning: Conceptual considerations. *American Journal of Distance Education*, *13*(2), 6–23. https://doi.org/10.1080/08923649909527020

Soo, K. S., & Bonk, C. J. (1998). *Interaction: What does it mean in online distance education?* Freiburg, Germany: Ed-Media/Ed-Telecom 98 World Conference on Educational Multimedia and Hypermedia. Retrieved from https://files.eric.ed.gov/fulltext/ED428724.pdf

Spiller, D. (2013, August). Feedback for learning. *TDU Talk*, (2), 26. Retrieved from http://www.waikato.ac.nz/tdu/pdf/TDU%20Talk/TDU%20Talk%20issue%202%20-%20FINAL%20VERSION.pdf

Straumsheim, C. (2015, July 24). Investments in ed-tech companies reach new high in first half of 2015. Retrieved from https://www.insidehighered.com/news/2015/07/24/investments-ed-tech-companies-reach-new-high-first-half-2015

Surma, T., Vanhoyweghen, K., Camp, G., & Kirschner, P. A. (2018). The coverage of distributed practice and retrieval practice in Flemish and Dutch teacher education textbooks. *Teaching and Teacher Education*, *74*, 229–237. https://doi.org/10.1016/j.tate.2018.05.007

Svila, J. (2014). Advances in design-based research. *Frontline Learning Research*, *6*, 35-45. Retrieved from https://journals.sfu.ca/flr/index.php/journal/article/view/114/211

Szpunar, K. K., Moulton, S. T., & Schacter, D. L. (2013). Mind wandering and education: from the classroom to online learning. *Frontiers in Psychology*, 4. https://doi.org/10.3389/fpsyg.2013.00495

Thorndike, E. L. (1913). *The elements of psychology*. New York, NY: A. G. Seiler. Retrieved from http://www.gutenberg.org/ebooks/39300

U.S. Department of Education, National Center for Education Statistics. (2018). *Digest of Education Statistics*, *2016* (NCES 2017-094), Table 311.15. Retrieved from https://nces.ed.gov/fastfacts/display.asp?id=80

Vianna, E., & Stetsenko, A. (2006). Embracing history through transforming it: Contrasting Piagetian versus Vygotskian (activity) theories of learning and development to expand constructivism within a dialectical view of history. *Theory and Psychology*, *16*(1), 81–108. https://doi.org/10.1177/0959354306060108

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Voorhees, R. A., & Voorhees, A. B. (2016). Principles for competency based education. In C. Reigeluth, B. Beaty & R. Myers (Eds.), *Instructional-design theories and models* (Vol. 4). Hillsdale, NJ: Routledge.

Watson, W. R., & Watson, S. L. (2016). Principles for personalized instruction. In C. Reigeluth, B. Beaty & R. Myers (Eds.), *Instructional-design theories and models* (Vol. 4, pp. 93–120). Hillsdale, NJ: Routledge.

Western Governors University was not eligible to participate in the title IV programs. (2017). (Office of Inspector General) (p. 93). Washington, DC. Retrieved from https://www2.ed.gov/about/offices/list/oig/auditreports/fy2017/a05m0009.pdf

What Works Clearinghouse. (n.d.). Retrieved July 25, 2016, from http://ies.ed.gov/ncee/wwc/

Wiliam, D. (2011). *Embedded formative assessment*. Bloomington, IN: Solution Tree Press.

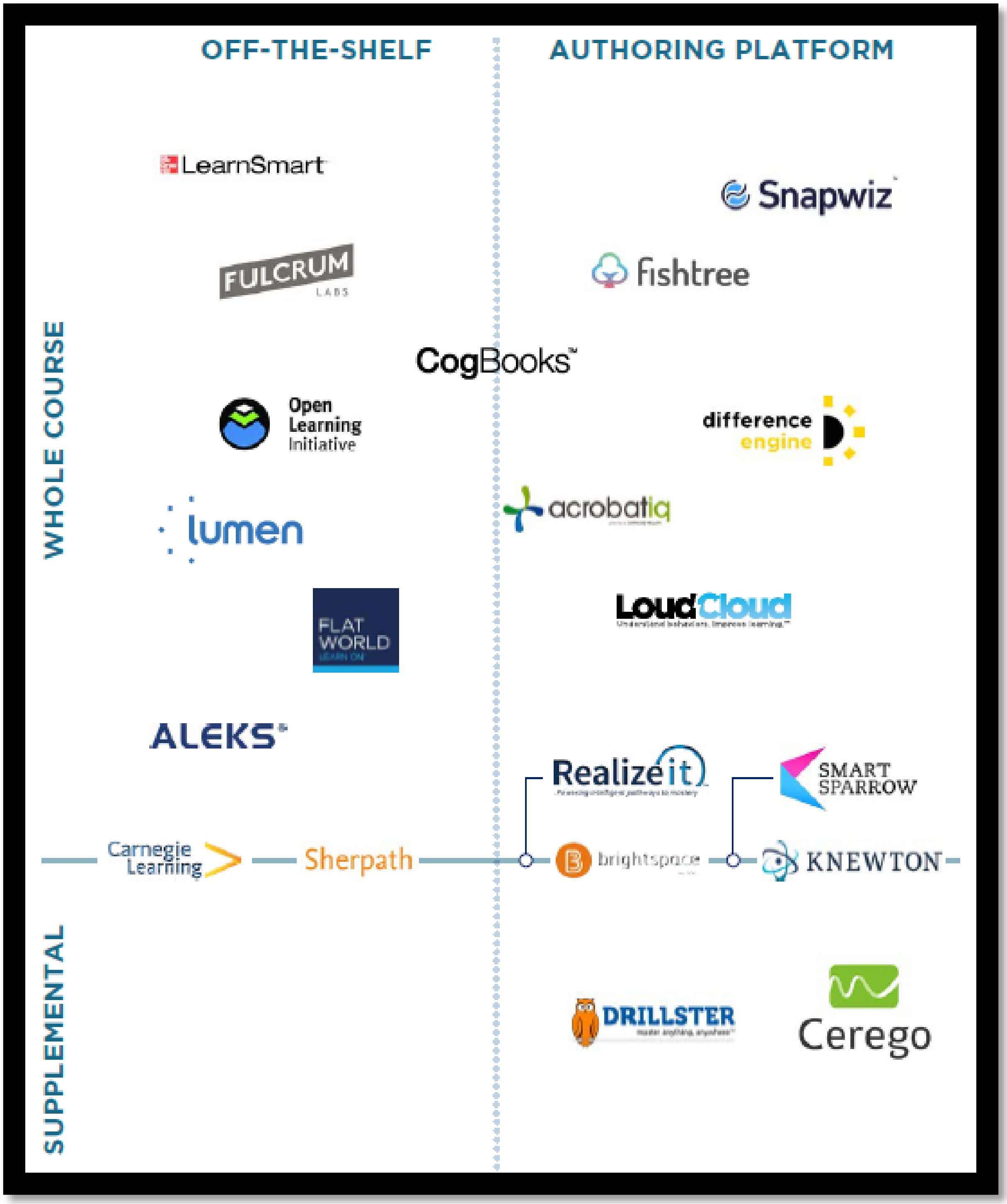
Worthen, B. R., White, K. R., Fan, X., & Sudweeks, R. (1998). *Measurement and assessment in schools.* Needham, MA: Allyn and Bacon.

Xie, K. (2013). What do the numbers say? The influence of motivation and peer feedback on students' behaviour in online discussions. *British Journal of Educational Technology*, *44*(2), 288–301. Retrieved from https://onlinelibrary.wiley.com/journal/14678535

Xing, W., Wadholm, R., Petakovic, E., & Goggins, S. (2015). Group learning assessment: Developing a theory-informed analytics. *Journal of Educational Technology and Society*, *18*(2), 110. Retrieved from https://www.j-ets.net/ETS/index.html.

Zabaleta, F. (2007). The use and misuse of student evaluations of teaching. *Teaching in Higher Education*, *12*(1), 55–76. https://doi-org.proxyiub.uits.iu.edu/10.1080/13562510601102131

Appendix A



Pearson CITE conference, 2016, unknown author/publisher

Appendix B

End-of-course survey questions

2. The instructor **respected** me as an adult learner by demonstrating qualities such as patience and kindness.

3. The instructor demonstrated a **willingness to assist** students.

4. The instructor's **knowledge of course content** was evident in the instruction of the course.

5. The instructor was **active in discussions**.

6. The instructor's **feedback** provided direction or encouragement beneficial to my academic success.

7. The instructor's **grading** and feedback aligned with course rubrics, scoring guides, and written directives.

8. The instructor **responded to student questions** within 48 hours.

9. The instructor **returned graded work** within seven days after the due date.

10. The course was **without significant errors** in the syllabus, workshop documents, tests/quizzes, etc.

11. **Assignments and activities** in this course expanded my knowledge of the subject(s).

12. The **instructional and reference materials** provided were relevant to the course.

13. My **workload** was appropriately distributed throughout the course.

CURRICULUM VITAE- ERIN A. CRISP

EDUCATION

Indiana University, Bloomington, IN
Ed.D. Instructional Systems Technology,
Minor: Learning Sciences July 2019

Towson University, Towson, MD
M.S. in Instructional Technology: Design and Development
Honors: Selected as the graduate student speaker for Commencement Dec. 2011

Indiana Wesleyan University, Marion, IN
B.A. in Education: English
Minor: Special Education: Learning disabilities K-12 April 2000

CAREER EXPERIENCE

Indiana Wesleyan University, Marion, IN
Executive Director of the Center for Learning and Innovation
April 2019- Present

- Provide visionary and strategic leadership to three teams- strategic program launch, instructional design and faculty enrichment serving 500 faculty and 900 online courses across 90+ degree programs.
- Consult with university committees and councils to provide instructional technology expertise.
- Facilitate faculty scholarship of teaching and learning (SOTL) by partnering with faculty and discovering avenues for research, publication and presentation.
- Contribute to and operationalize the university strategic plan.
- Advise on matters of curriculum and instruction as a member of executive cabinet.

Director of Academic Assessment and Evaluation-
Aug. 2016- Oct. 2018

- Led university efforts to adopt assessment standards and a gap analyses process for each academic unit. Elected by peers as the chair of the University Assessment Council, 2017-18.
- Condensed large data sets into accessible reports for specific intended audiences.
- Directed the RFP process for a data dashboard development project.
- Designed and implemented a continuous improvement process for 30+ online and face-to-face degree programs, across three schools, serving 10,000 adult learners annually.
- Developed a team responsible for delivering timely and accurate data to academic leaders.

- Invited to represent CAPS on the steering committee for the selection and implementation of a new learning management system. Our team successfully transitioned 800 online courses in 6 months.
- Invited to create and facilitate numerous assessment workshops, seminars, and webinars.
- Developed assessment related online training modules for faculty and administrators.
- Assisted in accreditation efforts for successful recognition from the Higher Learning Commission, social work (CSWE), business (ACBSP), and education (CAEP) accrediting organizations.
- Selected to participate in the Association for Continuing Higher Education's emerging leaders institute as one of 12 emerging higher education leaders (2017). Served on the planning committee for the 2018 cohort.
- Designed a strategic plan for the assessment of general education learning outcomes across the curriculum in the College of Adult and Professional Studies.
- Provided continuous improvement consulting services to academic units across the institution.

Director of Instructional Design- Center for Learning and Innovation
Jan. 2015- Aug. 2016

- Led a team of instructional designers and developers through a restructuring of team responsibilities.
- Created an eLearning Development Specialist position and initiated a strategy to provide eLearning development services using Adobe Captivate and Articulate Storyline 2.
- Managed an annual workflow for course creation and revision that resulted in the development of 60+/- courses annually.
- Managed a $500k budget and strategically prioritized budget dollars to offer development opportunities for my team in addition to exceeding prior years' productivity numbers.
- Increased course quality as measured by Quality Matters evaluation standards, through the development of a curriculum handbook and online training to document procedures.
- Designated as the primary institutional resource for Competency Based Education (CBE) and Adaptive Learning innovations. Selected to participate in the C-BEN Network's collaboratory in November, 2017.

Instructional Designer- Center for Learning and Innovation
Dec. 2013- Jan. 2015

- Migrated course content for 800 courses from document based Blackboard into HTML in eCollege.
- Learned to author HTML using Adobe Dreamweaver, troubleshooting formatting and style sheet errors.

- Collaborated with subject matter expert faculty members in the development or revision of online/hybrid courses in a wide variety of fields.
- Developed and delivered training related to best practices in instructional design.
- Increased team efficiencies, including an online ticketing system for course development or revision, by successfully implementing a new project management software (Wrike).

Adjunct Faculty-
ECED 210: Early Childhood Literacy and Assessment 2014-Present
UNV 104: Introduction to Integrative Learning 2018- Present

- Developed online courses for students in the early childhood AS degree and integrative studies programs.
- Provided formative feedback to learners, extra materials, and facilitated final projects.

Northwest Evaluation Association, Portland, OR (Remote employee)
Professional Development Facilitator (Consultant 2 years, Full time 2 years)
June 2010 – 2014

- Facilitated networking and professional development with 300+ educators, conducting 50+ full-day events annually.
- Developed and facilitated workshops focused on interpreting and applying data for instructional improvement as well as the development of collaborative professional learning communities.
- Engaged administrative leaders in strategic planning for school and district continuous improvement.
- Facilitated online webinars for groups of 25-30 teacher leaders monthly.

Stevenson University, Owings Mills, MD
Instructional Designer Dec 2011 – June 2012

- Designed 7-10 courses for online graduate programs in the Schools of Business and Law
- Earned certification as a Quality Matters™ course developer
- Facilitated over 20 workshops for faculty members on effective use of technology in instruction

Towson University, Towson, MD
Adjunct Instructor- ISTC 201: Introduction to Research Methods Fall 2011
Graduate/ Research Assistant in Education, Technology and Literacy Aug 2010- Jan 2012

- Used Adobe Captivate, Blackboard and many Web 2.0 tools,
- Designed and developed online course components.

- Trained faculty members in online course and instructional technology use, and provided guest lectures on technology in education in college classes.

Bluegrass Writing Project, Lexington, KY
Technology Liaison
May 2007- May 2009

- Trained a group of teacher leaders to design, gather and post website content
- Designed, developed and facilitated instructional technology workshops for teachers

West Jessamine High School, Nicholasville, KY
Literacy Coach / Reading Specialist
2006-2009

- Designed and facilitated workshops for educators within the district, presented at state level conferences, and coordinated assessment implementation at the school and district level.

Teacher, New Palestine, IN & Fishers, IN
2000-2006

- Taught Language Arts to middle and high school students with a focus on integrating relevant technologies
- Implemented a writing workshop approach to measurably improve student engagement.

SELECTED PUBLICATIONS AND PRESENTATIONS

Title	Type/Audience	Date/Year
Student Success in the Online Classroom: Assessing Cognitive, Professional, and Character Domains	60 minute presentation: Assessment Institute, IUPUI, co-presented with full-time faculty member, Dr. Theresa Veach.	October, 2018
Global Learning and the Academic Program Review Process: A Sustainable System for Continuous Improvement	20 minute presentation: Assessment Institute, IUPUI, co-presented with full-time faculty member, Dr. Theresa Veach.	October, 2018
The Program Health Index: A new Approach to Program evaluation	Poster presentation: Association for the Assessment of Learning in Higher Education	June, 2018
Breathing New Life into Program Review	Concurrent Session presentation at LEAP Conference, IUPUI, and Follow-up roundtable	March 2nd, 2018

	facilitator. Co-presented with DeVoe School of Business Program Director, Dedra Daehn.	
Defining The Learner Feedback Experience	Juried, academic journal article in *TechTrends*-(10th ranked journal in technology and education). Co-authored with Dr. Curtis Bonk.	March 2018
Using the Outcomes Alignment Template for Academic Programs	Workshop presented for the Association for the Assessment of Learning in Higher Education	June, 2017
Instructional Design Iron Triangle: Effectiveness, Efficiency and Appeal	Invited guest presenter for faculty professional development workshop, School of Service and Leadership	May, 2017
Are you seeing what I'm seeing?	Juried journal article publication *Intersection, 1(3), 7–14.*	Winter, 2017
Instructional Systems Design Research to Practice	Invited panel presenter: Graduate instructional systems technology conference (GIST), Indiana University	March, 2017
Student and Faculty Perceptions of Adaptive Learning	ACHE/ICCHE Midwest Conference	March, 2017
Adaptive Learning: A study comparing two tools	Award Winning Paper: Awarded best graduate student presentation by Online Learning Consortium. Co-authored with faculty member, Dr. Jodi Mills	November, 2016
Data Conversations: A Protocol for Growing a Healthy Culture of Data Use	Workshop presented for the Association for the Assessment of Learning in Higher Education	June 2016
Education 3.0: Individualized, Authentic and Connected	Society for Information Technology & Teacher Education International Conference 2016 (pp. 203-210). Chesapeake, VA: Association for the Advancement of Computing in Education (AACE)	June 2016
Evidence of Student Success	Paper presented at the Society for Information Technology and Teacher Education	June 2015
Rubrics: Current Research, Best Practices, and creating them in the Real World	Paper presented at Pearson Cite 2014 Online Learning Conference	May 2014
Developing Effective Professional Development Spaces, Real and Virtual, for College of Education Faculty	Paper presented at the Society for Information Technology and Teacher Education	June 2012
ActivExpressions Responders for Assessment	Paper presented at the Towson Faculty Professional Development Day	2011
Wikis and blogs in the classroom	Paper presented at the Kentucky Society for Technology in Education	2009

INDEPENDENT PAID CONSULTING

Curriculum development for the DeVoe School of Business's Doctorate in Business Administration degree program designed using problem-based learning.

Aug. 2018- April 2019

Reviewed DeVoe School of Business self-study documentation submitted to ACBSP for accreditation.

December 2017

Provided a problem-based learning workshop for faculty who were working to design a new degree program using a PBL approach.

November 2017

Led curriculum mapping and framework design for the Occupational Therapy doctoral program in the School of Health Sciences at Indiana Wesleyan University, resulting in successful ACOTE accreditation.

January to May 2017

Led curriculum mapping and framework design for the Masters in Public Health program in the School of Health Sciences at Indiana Wesleyan University.

March-August 2017

Collaborated on a $500,000 grant application for the Center for Women in Ministry at Wesley Seminary. (Funded)

January 2017

CREATIVE WRITING

Contracted to publish a new edition of *Why God Favors Women in Ministry* by Ken Schenck, (Paid)
Digitally published 10+ pieces for *Annesley Writers Forum* (Unpaid)
Alive, Conference follow-up, hard copy distribution to 600 readers, digital distribution to thousands, (Paid)
Contributor and Editor for *Unmasked*, a contracted publication for Wesleyan Publishing House, (Paid)
Contributor for *Light from the Word*, Wesleyan Publishing House, (Paid)
Editor for three editions of the Asbury Theological Seminary *Reader*, 2010, 2011, 2012 (Paid)

AWARDS

Best Graduate Student Presentation Award, Online Learning Consortium (Sloan-C)	2016
Student Speaker for Commencement Ceremony, Towson University	2012
Research Assistantship, Towson University	2011 – 2012
Graduate Assistant, Towson University	2010 – 2011
Secondary Educator of the Year Award- Indiana Wesleyan University	2000
Malcolm and Nadine Evans English Scholars Award- $4,000	1999
Departmental Scholar Award for Educators- $2,000	1999

LANGUAGES

English- Native language
French- Speak conversationally; read and write fluently

MEMBERSHIPS

Association for Continuing Higher Education, 2017- Present
Association for the Assessment of Learning in Higher Education, 2014-Present
Editor of Annesley Writer's Forum, 2014-15
Product Advisory Committee: Educate Online, 2014-2016
National Council of Teachers of English, 2000-2008
Council for Exceptional Children, 2000-2002
Association for School Curriculum and Development, 2000-2006
Society of Informational Technology and Teacher Education, 2006-2014

REFERENCES

Dr. Brock Reiman, brock.reiman@indwes.edu , Vice-president for Academic Affairs, Indiana Wesleyan University

Dr. Lorne Oke, lorne.oke@indwes.edu, Executive Director of the Center for Learning and Innovation, Indiana Wesleyan University

Dr. Jeff Kenton, jeff.kenton@towson.edu, Associate Dean, School of Education, Towson University

Barbara Mullins, barb.mullins@nwea.org, Sr. Curriculum and Instruction Developer, Northwest Evaluation Association

Alissa Harrington, alissaharrington@hotmail.com, Instructional Designer, Johns Hopkins University