

ABSTRACT

TEMPORAL CONTIGUITY AND VERBAL REDUNDANCY IN ONLINE INSTRUCTION: EXAMINING BOUNDARY CONDITIONS FOR MULTIMEDIA LEARNING

Hal Hinderliter, PhD,
Educational Technology, Research, and Assessment
Northern Illinois University, 2020
Dr. Ying Xie, Director

The Cognitive Theory of Multimedia Learning has enjoyed a prominent role in guiding the development of online instruction, but recent research has proposed boundary conditions for some of its key principles. Moreno and Mayer advanced the possibility of a reverse redundancy effect when narration and fully redundant on-screen text are presented in a temporally contiguous manner. The current study investigated the interaction and main effects of temporal contiguity and verbal redundancy by randomly assigning university students to one of four narrated text-only online lectures. Existing knowledge of the topic of instruction, information literacy, was assessed using a 15-question selected-response instrument developed for this research; the same instrument was later used to collect the posttreatment test data. Two-way ANCOVA analysis of the results revealed that when the pre test scores were accounted for, there was no significant difference in student achievement among the four treatment conditions. This finding held true even when the analysis was limited to low-knowledge students. While the current study's lack of significant results may be due to methodological issues, it indicates that text-only narrated presentations may not be subject to boundary conditions for the redundancy effect and that previously held beliefs in the negative impact of redundant text and narration are likely not applicable in such situations.

NORTHERN ILLINOIS UNIVERSITY
DE KALB, ILLINOIS

DECEMBER, 2020

TEMPORAL CONTIGUITY AND VERBAL REDUNDANCY IN ONLINE INSTRUCTION:
EXAMINING BOUNDARY CONDITIONS
FOR MULTIMEDIA LEARNING

BY

HAL HINDERLITER
©2020 Hal Hinderliter

A DISSERTATION SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE

DOCTOR OF PHILOSOPHY

DEPARTMENT OF EDUCATIONAL TECHNOLOGY, RESEARCH AND ASSESSMENT

Doctoral Director:
Xie Ying

ACKNOWLEDGEMENTS

I would like to acknowledge the wisdom, patience, and kindness of my dissertation committee chair, Dr. Ying Xie. In addition, I offer my gratitude to Dr. Todd Reeves for his expert guidance on statistical and research issues, as well as to Dr. Jason Rhode for his support and encouragement. Finally, I give thanks to my department chair, Dr. Wei Hung, and to Judy Puskar, ETRA's academic program advisor, for the faith they placed in me.

DEDICATION

This dissertation is dedicated to my wonderful wife, Mary, my son, Raymond, and to my father-in-law, Sylvester Schmid. Your support has meant the world to me!

TABLE OF CONTENTS	Page
LIST OF TABLES	vi
LIST OF FIGURES	vii
LIST OF APPENDICES	viii
CHAPTER 1: INTRODUCTION	1
Problem Statement	4
Purpose.....	5
Research Questions	5
Significance of the Study	6
Theoretical Framework and Constructs	7
Definition of Key Terms	7
CHAPTER 2: LITERATURE.....	10
Late Mid-Century Foundations	11
The Cognitive Era	13
The Age of Multimedia Learning	16
The Maturation of Online Instruction	21
Critics	23
Redefinitions	24
Relevant Empirical Studies	26
CHAPTER 3: METHOD	29
Null Hypotheses.....	30
Participants.....	31
Procedures.....	31
Phase 1: The Pretreatment Test.....	32
Phase 2: The Intervention	32
Phase 3: The Posttreatment Achievement Test.....	33

Ethical Concerns	33
Instrumentation	35
Pilot Testing and Instrument Validation	36
Analytic Approach	36
CHAPTER 4: RESULTS	38
Research Questions	39
Examination of Statistical Assumptions	39
Examination of Research Questions	44
Restricting Dataset to Low-Knowledge Learners	47
CHAPTER 5: DISCUSSION.....	48
Summary	49
Interpretation	50
Delimitations, Limitations, and Assumptions.....	53
Delimitations	53
Limitations	53
Assumptions.....	56
Implications.....	56
Recommendations for Further Research.....	58
Conclusions.....	59
REFERENCES	61
APPENDICES	66

LIST OF TABLES

Page

Table 1: Nine Multimedia Learning Effects (from Mayer, 2002, p. 133)	19
Table 2: Design of the Current Study	33
Table 3: Means and Standard Deviations of All Treatment Groups.....	40
Table 4: Homogeneity of Regression Slopes.....	42
Table 5: Levene's Test of Equality of Variances.....	43
Table 6: Tests of Normality	44
Table 7: Adjusted Means and Marginal Means from 2x2 ANCOVA Analysis	45
Table 8: ANCOVA Tests of Between-Subjects Effects	46
Table 9: Information Literacy Table of Specifications.....	69
Table 10: Item Fit Statistics for the 15-item Dataset	75

LIST OF FIGURES	Page
Figure 1: Concept Map of Related Cognitive Theories	3
Figure 2: Linearity of Covariate to Each Treatment	41
Figure 3: Homoscedasticity of Error Variances.....	43
Figure 4: Profile Plots of Estimated Marginal Means for the Dependent Variable	46
Figure 5: Item Characteristic Curves from the 15-item Dataset	74
Figure 6: Person-Item Map for the 15-Item Dataset	76

LIST OF APPENDICES	Page
APPENDIX A: DETAILS OF PILOT TESTING AND INSTRUMENT VALIDATION.....	66
APPENDIX B: DIRECTIONS TO SUBJECT MATTER EXPERTS	77
APPENDIX C: QUALTRICS FORM USED TO COLLECT SME FEEDBACK	79
APPENDIX D: RESPONSES FROM SUBJECT MATTER EXPERTS	82
APPENDIX E: FINAL ASSESSMENT QUESTIONS.....	89
APPENDIX F: TREATMENT ADMINISTRATION METHOD IN BLACKBOARD.....	95

CHAPTER 1: INTRODUCTION

Instructional design for online delivery has long taken a positive view towards the use of multimedia (Reiser & Gagné, 1983, p. 3). Multimedia enables the delivery of messages in multiple fashions, allowing learners to access a particular message in a new way, e.g., a classroom lecture delivered as a video or an e-book that incorporates animated diagrams and narration. According to Mayer's (2002) definition of a multimedia effect, delivering a message as both visual and auditory content will increase learners' opportunity to absorb and retain the instruction. This marked a break from the previously accepted notion that "combined audiovisual presentation is no better than auditory alone" (Penney, 1975, p. 69).

The multimedia effect is but one of a dozen principles that comprise Mayer's (2002) Cognitive Theory of Multimedia Learning (CTML). Recent developments in relevant theoretical frameworks for multimedia learning point to a reversal of many previously held assumptions regarding CTML's individual principles (Mayer, 2005; Mayer & Fiorella, 2014; van den Broek, Segers, & Verhoeven, 2014). Recent research that updates or redefines the components of CTML indicate the need for closer examination of the theoretical support for our current conceptions of multimedia design. These principles influence our practices and beliefs regarding the design of instructional content.

The widespread use and availability of video projectors in the classroom and computer monitors in the distance learning environment drive us towards communicating through rectangular presentation "slides." The use of presentation slides developed in specialized

software applications, e.g., Microsoft PowerPoint or Google Slides, is common practice (Orlikowski & Yates, 1994), but content creators are far from unified in their stylistic decisions regarding the visual design of these presentations. Some instructors are convinced that slides should be colorful to maintain student interest; others believe that every slide should contain at least one photograph. Among even casual users of PowerPoint, it is a commonly held convention that slides should not contain too many words if audio narration will also be present (Frish, Camerini, & Schulz, 2013; Savoji, Hassanabadi & Fasihipour, 2011). The conventional wisdom on this topic is to eschew redundant text, a guideline that is often declared as incontestable fact: “When a presenter is talking, the projection of a large amount of written text can overwhelm the limited capacity attention and language-based resources within working memory of the audience” (Garner, Alley, Gaudelli, & Zappe, 2009, p. 334).

Today, the rapid growth of distance learning compels us to examine the basis for such claims. As ever greater numbers of classes are converted to online delivery, it is important to revisit the experiments underlying this pedagogical position. The educational community must seek empirically based, theory-driven answers to the question, “How do changes in presentation style affect student outcomes?” Such research should consider online instruction as a distinct learning environment that differs substantially from face-to-face delivery in classrooms or lecture halls. The current research addresses this challenge by focusing on a basic form of online instruction often experienced by authentic online learners: narrated text-only slideshows.

As Sweller (2005, p. 159) noted, the phenomenon of interest to this research has been approached multiple times from multiple directions, with many authors failing to connect their efforts to previous similar endeavors. Past explorations of best practices for narrated

presentations have been based on the modality effect (the superiority of auditory presentation over visual presentation for items in short-term memory; Penney, 1975), dual coding (human cognition involves separate processing centers for language and nonverbal stimuli; Paivio, 1986), the redundancy principle (redundant material decreases the intelligibility of instruction; Chandler & Sweller, 1991), and the split-attention effect (forcing students to integrate two disparate sources of information imposes a cognitive load that decreases learning; Chandler & Sweller, 1992), among others (Figure 1). As the current study's literature review will reveal, the meanings of these terms have changed and even overlapped throughout recent decades; the one thread common to them all is the foregrounding of human cognitive processes.

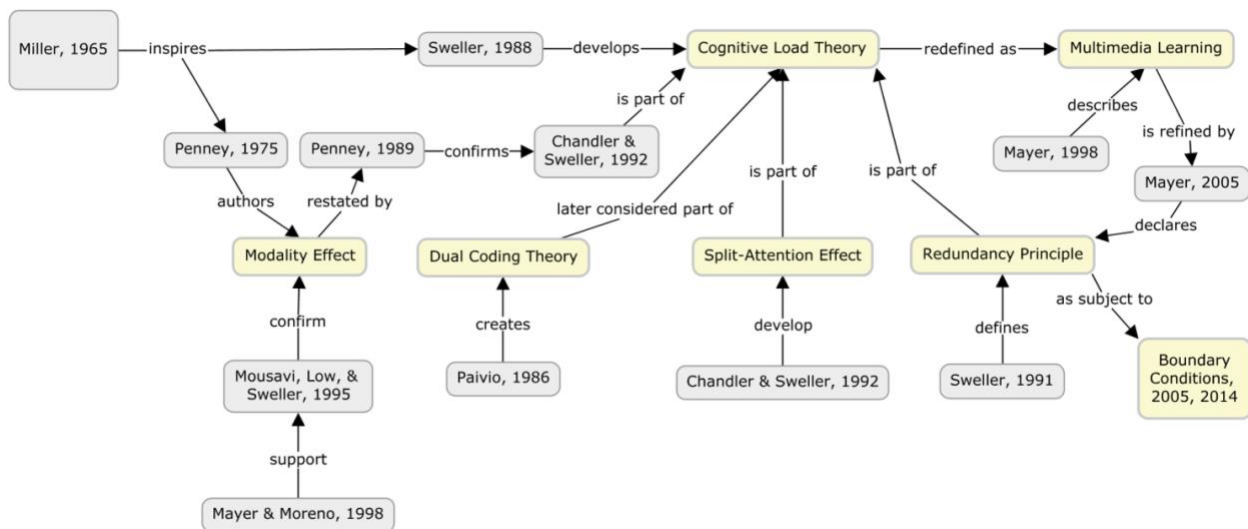


Figure 1

Concept Map of Related Cognitive Theories

Cognitive load theory (Sweller, 1988) made a substantive impact on the study of human learning processes. In this theory, the brain's production capabilities are tasked by the quantity of

items it is given; more tasks to perform equals a greater cognitive load. When the load is too great, the brain's limited cognitive capacity is exceeded. Mayer (2002) identified the redundancy effect, i.e., when identical material is presented in two modalities, as detrimental to the learning process due to its perceived role in increasing cognitive load. A few years later, new research (Mayer, 2005; Mayer & Fiorella, 2014; Mayer & Johnson, 2008) led to a reconceptualization of the redundancy principle in response to empirical data that revealed a beneficial effect from redundant text under specific circumstances. Mayer referred to these situations as *boundary conditions*. The current research is intended to investigate one of these boundary conditions.

Problem Statement

Aside from Mayer's own studies, scant empirical data exists to support this relatively new supposition of boundary conditions for the redundancy effect. This problem – a lack of targeted research into multimedia learning's potential asymptomatic behaviors – is marked by two especially problematic issues: the paucity of research on text-only narrated presentations and the inapplicability of much previous classroom-based research to today's online educational settings. This study seeks to address that gap in the literature while investigating Mayer's supposition of a boundary condition that arises from the combination of verbal redundancy, i.e., when audio narration corresponds directly to displayed text, and temporal contiguity, i.e., when displayed text is revealed in synchronization with the corresponding audio narration, with the goal of either supporting or failing to support such a boundary condition.

In the current research, an experimental quantitative design was used to collect data from a convenience sample of online students. Data were generated through a multiple-choice assessment, administered immediately after the subject had gone online to view one of four

randomly assigned narrated text-only presentations: a control treatment featuring sparse, non-temporally contiguous text (CTRL); a temporally contiguous treatment (TC); a verbally redundant treatment (VR); and a treatment that combined both verbally redundant and temporally contiguous attributes (BOTH). The same audio narration files and system-controlled slide timings accompanied all four versions; only the content and timing of visuals differed between treatments. This research design allowed analysis of the interaction between main effect A (temporal contiguity) and main effect B (verbal redundancy).

Purpose

The purpose of this study was to test for a relationship between the format of content presentation and student achievement. The independent variables for the experiment are *temporal contiguity* and *verbal redundancy*. The dependent variable is *achievement* related to a posttreatment test on the instructional content presented. *Existing knowledge* of the presented content, as assessed through a pretreatment assessment, is controlled for as a covariate.

Research Questions

The experiment is designed to address these research questions:

- Is there a statistically significant main effect of different temporal contiguity styles on students' achievement when controlling for pretest scores?
- Is there a statistically significant main effect of different verbal redundancy styles on students' achievement when controlling for pretest scores?
- Is there a statistically significant interaction effect of temporal contiguity and verbal redundancy on students' achievement when controlling for pretest scores?

As is standard for a study of this design, the null hypotheses can be expressed as:

$$H_0: \mu_{A_1} = \mu_{A_2}$$

$$H_0: \mu_{B_1} = \mu_{B_2}$$

$$H_0: \mu_{A_1 B_1} = \mu_{A_1 B_2} = \mu_{A_2 B_1} = \mu_{A_2 B_2}$$

Significance of the Study

Multiple concerns justify new research on this topic. Research conducted under up-to-date, real-world online learning conditions should be seen as an essential precursor to any recommendation of design principles for computer-based instructional materials. Advances in computer processing power, learning management systems, and interactive programming technologies render experiments from more than a decade ago unsuitable as proxies for the modern distance learning experience. The strongest justification, however, may be the lack of empirical data regarding best instructional design practices for *text-only* narrated presentations, despite their widespread use. As a result, studies whose variables included figures, diagrams, animations, or captions were excluded from this research.

This research benefits the instructional technology community, as well as the teachers and students it serves, by seeking to identify techniques that can advantage a broad swath of users – even when the instructional content consists only of narrated text. It also contributes to discussion around the applicability of established cognitive assumptions to the design of instruction for online environments, e.g., the widely held view that verbal redundancy is a waste of cognitive resources. As such, the results of this research are useful not only for designing more effective distance learning materials, but to enhance our existing views of human information processing.

Theoretical Framework and Constructs

This research adopts and extends Mayer's (2002) paradigmatic framework of multimedia learning. Of direct interest to the current research is Mayer's (2005) observation of *boundary conditions*, or situations in which theoretical principles do not produce their theorized effect. This research positioned text-only slides accompanied by matching narration, which they designated as examples of *verbal redundancy* (Mayer, 2005), as one such boundary condition, especially when the redundant text was incrementally revealed in a *temporally contiguous* manner (Mayer & Fiorella, 2014). Mayer's continuing work on boundary conditions has the potential to alter academia's current practices to the benefit of all online learners but questions persist regarding the authenticity of his observations. Some critics have gone so far as to challenge the legitimacy of multimedia learning theory due to the growing number of boundary conditions (de Jong, 2010), a call that warrants the current research.

Definition of Key Terms

To discuss this topic requires the delineation of two closely related concepts: media and mode. While informal use often treats these terms as synonyms, the fields of media studies, communication, and instructional design are among those who place a finer point on the issue. To these researchers, media can be defined as delivery methods, whereas mode can be thought of as the format of content. To quote Kress (2005), "I use the term mode for the culturally and socially produced resources for representation and medium as the term for the culturally produced means for distribution of these representations-as-meanings, that is, as messages. These technologies – those of representation, the modes and those of dissemination, the media – are always both independent of and interdependent with each other" (p. 6-7). In this manner, we can

think of *modes* as encompassing different artifacts of human communication, e.g., books, films, speeches, slideshows; *media* are the channels through which these representations are made available to audiences, e.g., the written word, photography, audio recordings, streaming video.

To further add to the confusion, educators often invoke slightly different meanings of the prefix “multi” when used with each term. When used with the word *modal*, the prefix *multi* often denotes a range of available options; e.g., in the humanities, composition teachers encourage multimodal projects in which self-expression can take the form of a painting, a photograph, or an interpretive dance. Preceding the term *media* with the prefix *multi*, however, usually indicates the simultaneous presence of two or more forms of media, e.g., an instructional multimedia presentation that incorporates text, images, and audio.

The use of multimedia is a hallmark of instructional content presented within distance learning courses, e.g., recorded lectures. For the purposes of this study, an *online lecture* is operationally defined as any multimedia educational content that is delivered digitally, either over the Internet or via storage media (e.g., on a CD, DVD, or flash drive), and features an audio component that approximates a traditional classroom lecture, supplies visuals that include explanatory text (which may be accompanied by some illustrations or photos), and presents information in a format that is mostly noninteractive. The pace of such online lectures can be user controlled, i.e., when learners are allowed to control the length of time each slide appears on screen, or system controlled, i.e., when each slide is advanced automatically without user intervention.

Given the minimalist prescriptions of mid-century pedagogies, it is unsurprising that later educational researchers looking for terms to describe the duplication of instructional content

were drawn to the word *redundancy* rather than “reinforcement.” Redundancy describes repetition but also implies that the duplicate item is unnecessary or potentially obstructive. Chandler and Sweller (1991) were among those who investigated multiple forms of redundancy, including when short summary paragraphs of text were provided alongside technical diagrams. Reese (1984) investigated verbal redundancy, i.e., when visible text is fully duplicated by spoken narration, and concluded that it enhanced learning; however, Kalyuga, Chandler, and Sweller (1999) concluded that verbal redundancy increases cognitive load and should be avoided.

Contiguity is the state of being contiguous, i.e., directly adjacent to something else. CTML references two types of contiguity: *spatial* contiguity, i.e., when two items are in close physical proximity, and *temporal* contiguity, i.e., when two items are timeline adjacent. When referring to verbally redundant presentations, the use of temporal contiguity means that lines of text are revealed “in sync,” i.e., they appear in synchronization with the audio narration.

Boundary conditions can be thought of as areas of dispute within the interpretation of theory. Not long before Mayer began using this term as an alternative to the more simplistic “reverse redundancy effect,” scholars had spent a decade lauding Woolgar and Pawluch’s (1985) article, “Ontological Gerrymandering: The Anatomy of Social Problems Explanations.” Their widely read treatise declared that “...both theoretical statements and empirical case studies manipulate a boundary, making certain phenomena problematic while leaving others unproblematic” (p. 214). In the current study, the term *boundary conditions* refers to times when researchers have identified outcomes that appear to reverse previously accepted paradigms, e.g., when Mayer and Johnson (2008) found that verbally redundant presentations produced positive outcomes rather than the harm predicted by prior theorizations of a redundancy effect.

CHAPTER 2: LITERATURE

As formalized public education largely replaced vocational internships across the United States in the early 1800s, the economic devaluation of teachers' talents led to overcrowded classrooms with inexperienced instructors. Thanks to the printing press, these novice teachers were able to rely on primer books and other instructional aids as compensation for their lack of subject expertise. Despite crowded classrooms and static learning materials, students were expected to docilely absorb then reiterate the content to which they were exposed.

At the dawn of the twentieth century, progressive education reformers like John Dewey (1900) were voicing complaints that this form of education was not only less effective at generating creative thinkers, it was also discriminatory. He recommended that educational experiences be derived from the interaction of internal conditions, e.g., existing knowledge and attitudes, and objective conditions, e.g., “equipment, books, apparatus, toys, games ... the materials with which an individual interacts, and, most important of all, the total social set-up of the situations in which a person is engaged” (Dewey, 1938, p. 45). Despite Dewey's criticism, however, the belief that a standardized prepared lesson should elicit the same response from all learners, aside from those students who were "defective" in some way, remained the basis for virtually all public education in the United States until various reform groups gained popularity in the 1960s. These reformers separated their new ideas from those of the past by labeling the old ways as *behavioralism*: the belief that a specific stimulus should always produce the same behavior, regardless of the student's background and/or abilities.

One reform movement of fundamental interest to the current research is *cognitivism*: the belief that lessons should be tailored to align with the cognitive learning process of students. The cognitivists' approach, however, still positioned students as malleable lumps of clay that could be reshaped by the "correct" stimuli. The early 1980s saw the rise of a new school of thought: that although external stimuli were important, only students themselves could construct their own knowledge. In this approach, the role of instructors is to find techniques that best resonate with each student so as to assist with the knowledge-building process. Such constructivist theories, i.e., ideas about how students construct knowledge, are often derived from educational processes that pre date our modern public education system, e.g., apprenticeships in which students are taught to imitate the behavior and techniques of experts during lessons that are situated within the same environment where that knowledge will be used.

Although constructivists sometimes resist the conceptualization of students' minds as analogous to computers, many of the theories advanced by cognitivist thinkers remain embedded within modern instructional design practices. Perhaps the foremost of these is *cognitive load* theory (Sweller, 1988), which holds that the human brain has discrete, limited capacity for visual and auditory input.

Late Mid-Century Foundations

The study of multimedia's efficacy for instructional tasks predates the Internet, online learning, and the use of personal computers in classrooms, but its most relevant and applicable theories have been developed over the three most recent decades. Throughout Mayer's most influential research (see Mayer 2002, 2005; Mayer et. al, 1996; Mayer & Johnson, 2008), multiple aspects of multimedia learning are seen within a cognitive framework. This focus on a

cognitive science perspective for the examination of learning processes, i.e., how our brains process then store information, can be traced to Miller (1956). His seminal paper on the limits of short-term working memory put forth the idea that humans face “severe limitations on the amount of information that we are able to receive, process, and remember,” before postulating that these limitations could be reduced by “organizing the stimulus input simultaneously into several dimensions and successively into a sequence of chunks” (Miller, 1956, p. 96). By foregrounding the importance of human cognitive processes, Miller advanced the idea that human capacity to learn is restricted by finite working memory.

Miller’s ideas were expanded upon through Atkinson’s information processing theory (1968), Anderson’s schema theory (1977), and others (e.g., Fleming & Levie, 1978; Johnson-Laird, 1983). These theories supported Paivio’s (1986) dual coding theory, which proposed separate processing centers for language and nonverbal stimuli. Paivio theorized that presentation modalities, which he defined as verbal, i.e., language, or nonverbal, i.e., symbolic, have an important impact on students. He surmised that the human brain processes verbal and nonverbal information independently but that these systems could operate in parallel, with each input processing system able to stimulate activity in the other. Paivio believed that together these dual systems could build complex representational structures that aid in concept formation and information retrieval (Paivio, 1986, p. 54). His dual coding theory differed from prior behaviorist notions of knowledge acquisition, which held that modality was unimportant. Prior to the advent of cognitivism, only the concepts being communicated were considered important; the choice of delivery method was typically thought to be inconsequential.

Penney (1975) explored these principles in her analysis of empirical results from experiments published in the 1960s and early 1970s, leading her to conclude that short-term memory can benefit from the auditory presentation of information more than visual presentation. In keeping with the behaviorist practices of the time, however, the experiments she reviewed all focused on the rote memorization of short item lists. Penney emphasized that the advantages of audio narration over displayed text, a condition that she called the modality effect, were limited to short-term retention of information. Dissenting studies were mentioned (Kroll, Parkinson, & Parks, 1972; Marcer, 1967; Scarborough, 1972) in which visual presentations were found superior. Among these, Kroll et al. (1972) included a treatment that combined audio with visual stimuli and found that the combination was more effective than the individual components. In their experiment, subjects viewed a series of letters that were also spoken in verbally redundant audio narration and were required to shadow the stimuli, i.e., repeat each letter aloud immediately after hearing it, in order to increase their overall cognitive load. Their findings showed that subjects who experienced verbally redundant stimuli exhibited better recall than visual-only or verbal-only subjects, although the superiority of verbal redundancy was much less pronounced after longer recall intervals. Kroll et al.'s findings conflicted with what had already become the prevailing notion, that "combined audiovisual presentation is no better than auditory alone" (Penney, 1975, p. 69).

The Cognitive Era

The work of Miller and Penney was foundational to the research of Sweller (1988), who theorized that students' learning strategies could directly influence the *load*, i.e., stress, placed on their cognitive processes (p. 275). Treating the human mind's processing of cognitive problems

like a series of switch gates on a circuit board, he described a production model for human problem solving. Sweller's tremendously influential theory of cognitive load declared that the human brain has discrete, limited capacity for visual and auditory input and cautioned against overloading the brain's processing capacity so that problem-solving schemas may be acquired for transfer into long-term memory. Three forms of cognitive processing are summed to define a learner's total cognitive load: intrinsic load, i.e., the effort required to understand the primary learning task; extraneous load, i.e., the undesirable additional stress incurred by poorly formed instructions; and germane load, e.g., the effort involved in fixing new information within long-term memory. The resilience of cognitive load theory can be seen when instructional designers configure learning environments so as to avoid overwhelming learners' restricted audiovisual processing capacity, a situation known as *cognitive overload*. The popularity and broad acceptance of cognitive load theory has greatly influenced other related theories, e.g., multimedia learning theory, dual coding theory, the redundancy principle; indeed, some efforts that pre-date Sweller's (1988) pronouncement of cognitive load later claimed association with the concept, including Penney's (1989) revisiting of the modality effect.

Penney (1989) provided another literature review pertaining to the modality effect, this time positioning it within a more recognizably cognitive frame. The studies she reviewed showed the benefits of audio processing for receiving instruction, as learners were better able to retain the sequential characteristics of verbal input, e.g., remember lists. Of great interest to the current study is her observation that the internal articulation of visually presented text is a cognitive challenge that is easily interrupted (p. 400). As with the prior research conducted by Kroll et al. (1972), it points to the cognitive load created by reading text that differs from narration when

both stimuli occur simultaneously. Penney's work helped spark a larger movement to extend the implications of cognitive load theory. Her research made it possible for Chandler and Sweller (1992) to reconceptualize the modality effect as simply the inverse of her initial finding: if audio narration reduced cognitive load in understanding a visual diagram, then adding text to a presentation should increase cognitive load; as more text is added, Chandler and Sweller theorized, a greater cognitive load will be imposed when processing that modal input.

As technology made it easier to combine text and graphics during instruction, Chandler and Sweller (1991) proposed the redundancy principle, i.e., that redundant material decreases the intelligibility of instruction by overloading learners' processing capacity. This research featured six quantitative experiments, none of which involved an auditory component – their tests utilized printed booklets in which explanatory text and legends were positioned either adjacent to or separately from a series of diagrams. Significantly and despite their assertion, Chandler and Sweller emphasized that they “were not engaged in a theory validation exercise” and that “direct tests of, for example, attentional or cognitive load factors were not carried out” (p. 329). As such and as discussed in Chapter 5, the current research finds it important to caution against extending Chandler and Sweller's findings to any situations in which redundancy occurs across modalities, e.g., verbal redundancy of written text and audio narration.

The following year, Chandler and Sweller (1992) identified the closely related *split-attention effect* as the cognitive load created by switching between focal points. Published just before the twentieth century's chalkboards began to be replaced by classroom video projectors, their research described the results of two separate experiments: one that used a presentation with many diagrams and a second based on a similar presentation featuring very few diagrams. The

authors theorized that materials that required students to split their attention between two non-adjacent locations provided inferior results because they place too great a demand on students' cognitive processing capacity. Chandler and Sweller explained this phenomenon by saying, "Learners are often forced to split their attention between and mentally integrate disparate sources of information (e.g., text and diagrams) before the instructional material can be rendered intelligible. This preliminary process of mental integration, while an essential precursor to learning, is likely to impose a heavy extraneous cognitive load" (Chandler & Sweller, 1992, p. 233). Despite the lack of an audio component to either of the experiments detailed in this research, this untested connection between increased cognitive load and word-laden visual presentations continues to be seen as support for the status quo of minimal text within PowerPoint slide design (Savoji, Hassanabadi & Fasihipour, 2011).

The cognitive theorist Sweller has strongly advocated against the use of fully redundant text and narration as a waste of precious cognitive resources. Despite admitting that "information that is redundant for one person may be essential for another" (Sweller, 2005, p. 165), he remains adamant that "information should be presented in a single form only, i.e., with all other versions and all unnecessary explanation eliminated" (p. 167). The current study posits that such prescriptive treatments are discriminatory against populations that diverge from Sweller's assumed norms.

The Age of Multimedia Learning

By the mid-1990s, cognitive-load-related theories pertaining to how learners acquire information from presentations were multiple. With so many related ideas competing for

attention, it is unsurprising that Mayer seized the opportunity to coordinate them all beneath a single umbrella, the Cognitive Theory of Multimedia Learning (CTML).

Mayer's first exposition (Mayer, Bove, Bryman, Mars, & Topangco, 1996) of what would grow to become a highly influential theory focused on an instructional technique that combined visual and verbal stimuli in an annotated illustration called a *multimedia summary*. In their study, Mayer et al. laid out the multimedia summary's rationale for selecting, organizing, and combining words and images in ways that promote the retention and transfer of scientific information by reducing students' cognitive load. They attributed the success of this approach to the utilization of three principles: "conciseness, in that only a few illustrations and sentences were presented; coherence, in that the images and sentences were presented in cause-and-effect sequence; and coordination, in that the images were presented contiguously with their corresponding sentences" (p. 72).

Influential research from Mayer and Moreno (1998) was among the first studies to position questions of cognitive load and student instruction within a computer-mediated environment. This research focused on the question of what should accompany an illustration for learners to receive the most benefit; as Penney (1989) as well as Chandler and Sweller (1992) predicted, the findings favored audio narration over on-screen text explanations. Once again, however, the study contained no tests of narration *combined* with on-screen text. "The most important new practical implication of this study is that animations should be accompanied by auditory narration rather than by on-screen text," wrote Mayer and Moreno. "This implication is particularly important in light of the increasing use of animations and on-screen text both in courseware and on the World Wide Web. These results cast serious doubts on the implicit

assumption that the modality of words is irrelevant when designing multimedia lessons with pictures and words” (Mayer & Moreno, 1998, p. 318-319).

This research by Mayer and Moreno broadened the meaning of “split attention” to substantially overlap with Penney’s modality effect. Their research described two experiments where participants were shown animated diagrams on lightning formation and automotive braking systems. Half of the students viewed animations accompanied by text descriptions while the remainder viewed animations accompanied by audio narration. In both experiments, subjects who received audio explanations outperformed those who were shown text-based explanations. They summarized these results as, “Multimedia learners can integrate words and pictures more easily when the words are presented auditorily rather than visually” (Mayer & Moreno, 1998, p. 312). Similar to Penney’s 1989 research on the modality effect, Mayer and Moreno’s study was narrowly focused on how best to accompany a graphic illustration. Despite this limited scope, Mayer and Moreno’s 1998 research has been cited more than 1,600 times, often to advance the idea that presenting too much text will overload visual working memory, even when no graphics are present (e.g., Garner & Alley, 2011; Savoji, Hassanabadi & Fasihipour, 2011). Mayer’s future research would theorize exceptions to this proposition.

The following year, Moreno and Mayer (1999) continued to detail CTML, this time focusing on the effects of contiguity and modality. In this research, the concept of temporal contiguity, i.e., the synchronization of visual and spoken materials, is derived from previous conceptualizations of a split-attention effect (Chandler & Sweller, 1992; Sweller et al., 1990; Tarmizi & Sweller, 1988). Empirical data for Moreno and Mayer’s study was gathered from 81 subjects. The instructional treatment featured animations that depicted various meteorological

concepts. Results showed a large positive effect (.094) for modality, i.e., the use of audio narration to accompany the animations. Today, this concept could also be referred to as redundancy. Contiguity, as measured for the study, was found to have no main effect.

In a 2005 book for Cambridge University Press, Mayer published a fuller explanation of multimedia learning that detailed three primary assumptions: *dual channels*, i.e., that visual and auditory stimuli are processed separately; *limited capacity*, i.e., that the amount of information these channels can process is circumscribed; and *active processing*, i.e., that attending to, organizing, and integrating information lead to meaningful learning. In a literature review published a few years earlier, Mayer (2002) elaborated the first of multiple lists defining the key principles of multimedia learning. His research article in the journal *Psychology of Learning and Motivation* defined an initial nine key principles for CTML: the multimedia effect, the spatial contiguity effect, the temporal contiguity effect, the coherence effect, the modality effect, the redundancy effect, the pretraining effect, the signaling effect, and the personalization principle (Table 1).

Table 1

Nine Multimedia Learning Effects (from Mayer, 2002, p. 133)

Principle	Number of tests	Median effect size
<i>Multimedia effect:</i> Better transfer when a message contains words and pictures rather than words alone.	11 of 11	1.39
<i>Spatial contiguity effect:</i> Better transfer when printed words are placed near rather than far from corresponding pictures.	5 of 5	1.12
<i>Temporal contiguity effect:</i> Better transfer when corresponding narration and animation are presented simultaneously rather than successively.	8 of 8	1.30
<i>Coherence effect:</i> Better transfer when irrelevant words, pictures, and sounds are excluded rather than included.	11 of 11	1.11

<i>Modality effect</i> : Better transfer from animation and narration than from animation and on-screen text.	12 of 12	1.13
<i>Redundancy effect</i> : Better transfer from animation and narration than from animation, narration, and on-screen text.	3 of 3	0.84
<i>Pretraining effect</i> : Better transfer when training on components precedes rather than follows a message.	3 of 3	1.39
<i>Signaling effect</i> : Better transfer when narration is signaled rather than nonsignaled.	2 of 2	0.60
<i>Personalization principle</i> : Better transfer when words are in conversational style rather than formal style.	5 of 5	1.55

Of Mayer's original nine principles, the idea most relevant to the current research is the *redundancy effect*, which in recent times has come to mean the repetition of visually presented text with audio narration. In his 2002 literature review, Mayer identified redundancy as a factor that would increase cognitive load, thereby reducing the student's capacity for learning. It is also important to note that while every study referenced in Table 1 strongly supports Mayer's principles, those experiments preceded the advent of online learning and therefore may not be applicable to modern distance learning environments.

In that same year, Mayer's monolithic representation of verbal redundancy as an entirely negative practice was shattered. Mayer's realization that redundant narration may be of value in text-only situations can be traced to his 2002 research with Moreno, in which verbal redundancy in multimedia learning was tested in three experiments. Moreno and Mayer's (2002) findings included a three-fold increase in correct answers when learners were shown presentations in which the narration was an exact reflection of the on-screen text, i.e., verbally redundant. They found improved results across multiple forms of assessment, stating, "Students remembered significantly more when the verbal material was redundant than when it was not ... they

generated significantly more conceptual creative solutions on the transfer test when the verbal material was redundant than when it was not ... and they correctly matched more items when the verbal material was redundant than when it was not” (p. 162).

The Maturation of Online Instruction

A few years later, Mayer (2005) offered a more articulated sense of the redundancy effect in order to accommodate his own research findings of measurable benefits from specific forms of redundancy. The data considered for Mayer’s analysis was drawn from 16 previously published experiments, all of which featured a combination of graphics and text; in nine of these cases, audio narration was also involved. Although this work solidified much of Mayer’s earlier research, he also found it important to identify some isolated discrepancies regarding the redundancy principle, e.g., that “the redundancy effect can be eliminated or even reversed when the learners are experienced, the on-screen text is short, or the material lacks graphics” (Mayer, 2005, p. 299).

Soon after, Mayer and Johnson (2008) provided more detail on these “exceptions to the rule” which they called *boundary conditions*. Their findings posited text-only slides accompanied by redundant narration as one such occurrence of a boundary condition for redundancy. Previously seen as a purely deleterious effect, their research from 2005–2008 expanded to position redundancy as actually helpful for both narrated presentations that lacked graphics (i.e., “text-only” presentations) and when short text labels are adjacent to the graphics they describe. Mayer and Johnson describe boundary conditions as a “reverse redundancy effect” that can occur under several conditions: when the narration is complex or contains unfamiliar words, when the narration is not in the learner’s native language, when the audience is composed

of low-knowledge learners, or when the pace of presentation is slow or learner controlled. As will be detailed in Chapter 3, the current research used a system-controlled delivery to standardize temporal cognitive load across all treatments.

When they returned to this topic almost a decade later, Mayer and Fiorella (2014) additionally identified synchronization of visual text with narration, which they referred to as the *temporal contiguity* of presentation, as another component of the boundary condition for redundancy. That study stands as a rare example of influential research on the efficacy of text-only presentations. He and Fiorella described this aspect of the multimedia paradigm: “The redundancy effect can disappear when no graphics are presented. In this case, adding on-screen text does not create split attention ... because there is no other material to process in the visual channel” (Mayer & Fiorella, 2014, p. 299). Their 2014 work continued to revise previous lines of thought on this interaction by positing that a temporally contiguous and verbally redundant presentation style can benefit rather than impair cognitive function, ostensibly through a reduction in cognitive load.

Following Mayer’s 2005 identification of boundary conditions, other researchers have proposed additional boundary conditions for CTML’s principles. The CTML tenets impacted by these studies include the personalization effect, e.g., that when formal language was replaced by conversational language, results for higher knowledge learners did not improve (McLaren, DeLeeuw, & Mayer, 2011; Wang et al., 2008). Some theorists might be wary of expanding a framework to include a dozen principles, but Mayer (2017) recently added three more principles to CTML: *segmenting*, i.e., breaking complex instruction into manageable chunks; *voice*, i.e., that human voices should be natural rather than computer generated; and *embodiment*, i.e., the

inclusion of human gestures from on-screen agents. As CTML expanded from nine principles in 2002 to twelve principles in 2017, each with the possibility of boundary conditions that might cause a reversal of the stated effect, its complexity has become a matter of concern. Even prior to CTML's expansion, de Jong (2010) challenged the legitimacy of multimedia learning theory due to the growing number of proposed boundary conditions. While each new extension to CTML demands attention, this added complexity also increases the need for assurance of its utility. To that end, numerous questions can be posed to help us rethink specific aspects of our broadly accepted guidelines for the use of multimedia.

Critics

Tabbers, Martens, and van Merriënboer (2004) asked if broadly accepted findings regarding the modality effect might not be generalizable, due to the unique conditions and content of the previous research. Tabbers et al. surveyed the previous research in this area, then questioned if these experiments were adequately reflective of real educational environments (p. 74). The authors expressed concern that many of the landmark studies in this area were conducted under laboratory conditions and involved only brief instruction focused solely on technical domains; such research includes Jeung et al. (1997), Kalyuga et al. (1999), Mayer and Moreno (1998), Moreno and Mayer (1999), Mousavi et al. (1995), and Tindall-Ford et al. (1997).

Even Tabbers and colleagues' experiment was based on group exercises using Internet-connected computers in a college laboratory, as opposed to students working from the isolation of their homes or other extra-classroom settings. This potential inapplicability of previous research to today's online educational settings exacerbates the lack of targeted research into multimedia learning's potential asymptomatic behaviors. Since the publication of Tabbers et al.'s

provocative research, the use of computer-based experiment design has grown, but not all previous studies have been thoroughly replicated using adequate sample sizes within modern online instructional environments.

Unconventionally, Tabbers et al.'s (2004) research showed that the use of visual learning material was superior to audio in terms of student transfer and retention: "Replacing visual text with spoken text even had a negative effect on learning, contrary to what both cognitive load theory and Mayer's theory of multimedia learning would predict" (p. 80). Tabbers et al. (2004) were not the only researchers who found it difficult to confirm the superiority of audio narration over written instruction. Along with their own results, they cite Jeung, Chandler, and Sweller (1997) and Kalyuga, Chandler, and Sweller (1999) as empirical research that found the use of narration instead of text was not always beneficial, especially when the speed of instruction was system determined rather than user controlled.

Redefinitions

In their article, "Cognitive Load Theory, Modality of Presentation and the Transient Information Effect," Leahy and Sweller (2011) presented two experiments with primary school students that the authors claimed reaffirmed the latter's earlier work with Chandler while simultaneously redefining the modality effect. Their research is mentioned here because the conditions of Leahy and Sweller's experiments further blurred the distinction between modality and verbal redundancy. Empirical data for Leahy and Sweller's (2011) research was garnered from two small-scale studies conducted on 11- and 12-year-old students. The results of the first study show that students who only heard concepts described performed better than those who both heard and read the same information simultaneously; however, the second study produced

the opposite result. The authors claimed that these results, although conflicting, supported the idea that verbal lectures should not be accompanied by a fully corresponding presentation of text. Subsequent researchers have largely accepted that assertion, rather than follow Leahy and Sweller's call to "repeat our experiments using adults with independent measures of cognitive load" (Leahy & Sweller, 2011, p. 950).

Following after Chandler and Sweller's 1992 anti-text redefinition of Penney's 1989 audio-centric findings, this revised definition of the modality effect now referred to improvements achieved by sharing content presentation between visual and auditory delivery channels rather than duplicating all content across both channels – that the modality effect is evidenced "when instructional material that is presented in a dual, audio/visual format is superior to a visual only presentation" (Leahy & Sweller, 2011, p. 943). Perhaps most importantly, their findings were brought into alignment with the previous findings of Chandler and Sweller (1992) when Leahy and Sweller declared that a beneficial modality effect can only occur when detailed verbal narration is accompanied by a sparse amount of text. Although their study contained no empirical data that would support such a claim for text-only presentations, Leahy and Sweller opposed having the narrator repeat the slide's text verbatim: "If one source of information merely recapitulates another source of information in a different modality, the effect will not be obtained" (Leahy & Sweller, 2011, p. 943). The authors cast the mixed results from their own experiments as an affirmation of their theory, but Leahy and Sweller were forced to spend one-half of the article's first page discussing possible explanations for the reverse modality effect observed in their second experiment.

Relevant Empirical Studies

Searching the databases ERIC by EBSCO, JSTOR, and Google Scholar for empirical research related to “multimedia learning,” “redundancy principle,” “split-attention effect,” or “modality effect” in conjunction with “cognitive load” provided several dozen applicable papers. The relevant studies discussed here were all conducted within the last two decades. It is also worth noting what was discarded during this review. Some journals have published non-empirical papers on this topic; many are similar to Garner and Alley (2011), in which the authors recommend “specific ways to overcome the problems with the default settings of PowerPoint” (p. 101). However, the authors present no empirical data to support their criticisms of conventional PowerPoint design. Like others encountered during the preliminary review of literature, the authors make reference to earlier statements by Richard Mayer, John Sweller, and especially Edward Tufte as they criticize text-heavy PowerPoint for conflicting with cognitive load theory; however, these presumptions were not supported by empirical research data. This approach is in keeping with Tufte’s popular 2003 book, *The Cognitive Style of PowerPoint*, which gained notoriety for laying blame for the space shuttle Challenger’s explosion on miscommunication engendered by NASA’s heavy reliance on PowerPoint. Despite the massive influence this text achieved through the mainstream press, it contains no quantitative experiments on the topic.

Toh, Munassar, and Yahaya (2010) presented the results of their experiment with non-native English speakers, which showed that learners exposed to temporally contiguous, verbally redundant instruction performed significantly better than those who experienced only audio narration; in both treatments, the slides also contained static images. This behavior corroborated

Mayer's supposition of a boundary condition for redundancy, although Toh et al. identified the phenomenon as a "reverse redundancy effect." Their study also contrasted the conceptual nature of the given instruction with the classic studies in this area, which predominately used instruction on technical and scientific concepts. After analyzing the results from 209 subjects, Toh et al. found that "redundant synchronized on-screen text did not impede learning; rather it reduced the cognitive load and thereby enhanced learning" (p. 988).

Savoji, Hassanabadi, and Fasihipour (2011) revealed they found no evidence to support the traditional conception of a modality effect for multimedia instruction. Four treatment groups (low interactivity groups passively watched a fully automatic presentation while high interactivity groups needed to click items to move forward, across both text-based and audio-based presentations) found no significant difference in students' perceptions of difficulty, but the high-interactivity group expended more mental effort while the low-interactivity group had better comprehension scores. More pertinently, differences were not significant between text-based and audio-narration-based presentations. Regarding the modality effect's presumption that "information in narration form instead of on-screen text will improve learning," Savoji, Hassanabadi, and Fasihipour (2011) declared that "the results did not support the hypothesis" (p. 1491).

More recently, van den Broek, Segers, and Verhoeven (2014) studied 84 university students and also found evidence of a *reversed* modality effect (i.e., boundary condition) rather than the expected support for Sweller's interpretation of this theory. While initial test results were not significantly different when comparing students who viewed visual-only educational materials versus those who experienced a narrated version of the same content, students who

were exposed to visual-only material (diagram plus text) performed significantly better than the audio-visual (diagram with spoken narration) group one day after learning on three of four performance measures. The researchers concluded that their experiment's "results contradict common multimedia design recommendations and instead suggest that learner-paced presentations should include on-screen text" (van den Broek, Segers, & Verhoeven, 2014, p. 438).

CHAPTER 3: METHOD

An experimental quantitative design was used to address the current study's research questions while maintaining continuity with previous studies in this area. As Shadish, Cook, and Campbell (2001) noted, "The strength of experimentation is its ability to illuminate causal inference. The weakness of experimentation is doubt about the extent to which that causal relationship generalizes" (p. 18). This research, like those past efforts with which it hopes to correspond, embraces a postpositivist realist perspective. As Smith (1983) delineated, "From the quantitative perspective the overall purpose of educational research is to explain, and by extension to be able to predict, the relationship between or the invariant succession of educational objects and events" (p. 11). By carefully isolating the dependence of specific variables within the context of an observable phenomenon, the experimental method can be said to provide an objective mechanism for pursuing epistemic goals.

This study features an experimental quantitative research design, which enabled the collection and interpretation of data concerning learners' ability to recall and identify information presented during an online instructional presentation. The experiment was designed specifically to test two attributes of multimedia learning theory's expanded definition of redundancy by looking for evidence of Mayer's proposed boundary conditions derived from verbal redundancy (Mayer, 2005) and temporal contiguity (Mayer & Fiorella, 2014). The current research also included a combined treatment to assess the value of interactions between verbal redundancy and temporal contiguity. Due to the modest sample size for each treatment, a pretest was used to assess participants' existing knowledge of the assessment topic. Two-way

ANCOVA testing was used to determine the differences between adjusted population means for the independent variables (verbal redundancy and temporal contiguity), the covariate (the pretreatment test), and the dependent variable (the posttreatment achievement test).

The data collection process involved three phases. In the first, dichotomous data was collected using a pretest assessment administered sans any related instruction; in the second, participants were instructed on the topic of information literacy via an online instructional presentation; the third phase was the repetition of the same assessment immediately following the instructional intervention. All phases were available asynchronously from the course's Blackboard website and were configured using the Sharable Content Object Reference Model (SCORM) protocol in order to report the assessment data to Blackboard. The first phase (pretest) was made available via Blackboard for a period of seven days. After seven days, Phases 2 and 3 became available. All phases were completed within the Spring 2020 semester.

Null Hypotheses

- Compared to narrated slides featuring sparse, unsynchronized text, the use of a verbally redundant presentation style will not significantly affect assessment results.
- Compared to narrated slides featuring sparse, unsynchronized text, the use of a temporally contiguous presentation style will not significantly affect assessment results.
- Compared to narrated slides featuring sparse, unsynchronized text, the combination of verbally redundant and temporally contiguous presentation styles will not significantly affect assessment results.

Tests of these hypotheses were evaluated at an alpha level of .05.

Participants

Subjects were drawn from a convenience sample of students participating in a popular multisection online class at a midwestern university ($n = 113$; age $m = 21$; gender $f = 53$; $m = 59$; $ESL = 9$). Two online sections of the same first-year general education class were assigned the online instructional module through its appearance on a course module page within Blackboard, the research site's learning management system (LMS). Students across both sections were randomly assigned to one of four treatment groups; those assignments were transparent to the participants.

Procedures

The present study collected data in first and last of three phases: a pretreatment assessment, followed by the intervention, then the posttreatment assessment.

The instructional content and assessments were delivered as an HTML5 multimedia SCORM package, developed using the Quizmaker tool from the iSpring Suite plug-in for Microsoft PowerPoint. These SCORM packages relayed learners' final scores to the LMS but also provided a more detailed report of each individual submission via email. To take advantage of this additional detail, the current research is based on an analysis of the emailed response data.

The time to complete each assessment was limited to one hour, and participants were forced to complete the assessment once it began. Participants were given seven days to complete the pretest, which became available at 12:01 AM on April 7. One week later at 12:01 AM on April 14, the posttreatment test was made available. This led to a wide range of time differentials between the completion of the two rounds of online presentation/assessment packages.

Phase 1: The Pretreatment Test

The pretreatment test (also known as the pretest) was administered to a convenience sample of 129 undergraduate students enrolled in two sections of a popular online class. All participants received the same pretest, consisting of multiple-choice questions on the topic of information literacy. The pretest enabled the analysis to control for participants' existing knowledge of the instructional topic; it also provided statistical data used to evaluate the validity of the novel instrument. This phase ended with the submission of the participant's responses.

Phase 2: The Intervention

In this phase, each participant was randomly assigned to one of four treatment groups (CTRL, VR, TC, BOTH). Members of each group viewed a similar 22-minute system-controlled instructional presentation, sharing the same audio narration but with each treatment featuring a different variation of the independent variable:

- The control group (CTRL) established the baseline as a traditional PowerPoint-style presentation with detailed narration but very few words on each slide.
- The temporal contiguity (TC) group received the same audio narration and PowerPoint slides as in the control group, but with the appearance of individual words and phrases timed to appear in synchronization.
- The verbal redundancy (VR) group received the same audio, but with full correspondence of the displayed text and the words spoken in the narration.
- The combined verbal redundancy and temporal contiguity (BOTH) group received the same audio as well as fully redundant text that appeared in synchronization with the narration.

The 2x2 research design is illustrated in Table 2. Participants engaged with the online instruction in authentic distance learning environments; e.g., at home, in the library, or at a coffeeshop.

Table 2

Design of the Current Study

Independent Variables		Cells of the Study Design
Verbal Redundancy	Temporal Contiguity	
No	No	CTRL: No temporal contiguity, no verbal redundancy
	Yes	TC: With temporal contiguity, no verbal redundancy
Yes	No	VR: No temporal contiguity, with verbal redundancy
	Yes	BOTH: With temporal contiguity, with verbal redundancy

Phase 3: The Posttreatment Achievement Test

Immediately following the online instruction, each participant completed the same assessment questions presented in the pretest. While the sequence of questions remained constant for both administrations, the order of the possible responses to each question was randomized. As with the pretest, this phase also ended with the submission of the participant's responses.

Ethical Concerns

The only substantive potential threat this study posed for participants was a breach of confidentiality. In order to minimize the potential for harm, data handling procedures were designed to maintain subject anonymity. Student names and campus ID numbers were the only identifying information collected. All test scores were de-identified prior to analysis with IBM

SPSS and R, and all identifiable data were permanently deleted in accordance with a data management plan. This involved editing the raw data in Microsoft Excel, where student names and campus ID numbers were replaced with a generic case number prior to further analysis in SPSS. The data plan for this research called for the anonymized data to be stored for a period of three years, to allow time for prepublication review. Data containing student names and campus IDs were not stored.

A waiver of informed consent from the university's Institutional Review Board allowed participants to view the experiment as a normal class exercise. This step was necessary to ensure the ecological validity of the experiment, due to the risk that participants' knowledge of the experiment could negatively impact the accuracy and generalizability of the findings. If students were aware in advance that they would receive full credit for completing the assessment, regardless of how many correct answers were given, a considerable number of students would be likely to take the assessment less seriously, e.g., clicking the first answer on every question in order to complete the process more rapidly. Historically, this lopsided ratio of minimal risk to widespread benefits has led generations of education researchers to make this same waiver request, especially when the participants are above the age of 18. The "deception" involved in this approach is simple omission, where the experiment is unannounced and appears to be part of the normal learning routine. Only by making the experiment unannounced could the risk of damage to this study's ecological validity be minimized.

Concern over confidentiality is joined by an additional concern: that the nature of the experiment means some groups of students were exposed to instruction that may have been less effective than the instruction received by other participants. The potential harms of this concern

come in two forms: (a) that students' grades in the host course may be negatively impacted compared to classmates who received the more effective training, and (b) those same students may have a reduced understanding of the material being covered. These issues were discussed in this study's IRB application and are common to any classroom-based experiment. Justification to exempt the experiment from IRB review was made on the basis that (a) once the experiment had ended, all participants would receive full credit for the assignment, and (b) all formats of instruction offered in this experiment were equivalent or superior to the online presentations offered during previous semesters of this course. The IRB granted the current study's request to waive informed consent on the condition that all subjects would be debriefed at the semester's end and given the opportunity to withdraw their data from the study; none of the participants opted to have their data withdrawn.

Instrumentation

The instrument used in the current research has been developed specifically for this application and validated through multiple procedures, which are outlined in Appendix A. As part of pilot testing, the assessment items contained on this instrument were subjected to Rasch testing for goodness-of-fit at the global and item levels. After one of the draft questions was eliminated due to item fit concerns, the instrument was finalized as 15 selected-response (multiple-choice) questions, each with one correct answer and three distractors. Participants were provided with four possible answers for each question, with the correct answer valued at one (correct) or zero (incorrect) then weighted equally at 6.6667 points to produce a maximum overall score of 100. The same 15 assessment questions were used during both the pretest and the experiment.

Pilot Testing and Instrument Validation

An initial experiment design using an established instrument from Dwyer and Lamberski (1977) was conducted during spring and summer 2019, but after evaluation, this instrument was deemed inappropriate for use with text-only narrated instructional content. After searching for a more suitable existing instrument, I decided that the creation of a new instrument was warranted. Following an initial experiment from fall 2019, the new instrument was substantially revised in January 2020, then pilot tested in March 2020. The results of this pilot testing are offered as evidence of this instrument's validity. Selecting from among the evidence-gathering procedures outlined by Reeves and Marbach-Ad (2016), three methods of evaluation were used (Appendix A). Through the construction of a table of specifications, an expert panel review, and an analysis of pilot testing results ($n = 51$), the current research offers evidence to support the validity of this instrument through multiple forms of evidence. The final instrument consisted of 15 questions. The scale had a satisfactory level of internal consistency, as determined by a Cronbach's alpha of 0.745 for the posttreatment assessment and 0.688 for the pretreatment assessment.

Analytic Approach

Quantitative results from the pretest and posttest were collected using the same instrument. Data was collected from the emails as text, generating an initial list of 129 participants. The data were formatted in Microsoft Excel, where case-wise deletion was used to eliminate partial cases, e.g., students who quit their web browser without completing the exercise or did not complete both tests ($n = 14$) and those who completed the exercises in the wrong sequence ($n = 1$). After analysis in SPSS, one case was determined to be an outlier and removed. The remaining 113 records were subjected to further analysis in SPSS.

The results of each treatment were assessed through two-way ANCOVA analysis. Multiple assumptions of ANCOVA (continuous DV, independent categorical IVs, continuous covariate, and independence of observations) were verified prior to the selection of a 2x2 ANCOVA as the current study's statistical analysis method. This statistical operation was used to detect how the independent variables, temporal contiguity and verbal redundancy, influenced the dependent variable of achievement. An interaction effect should be seen as supporting the conclusion that the redundancy principle is subject to a boundary condition under this combination of factors.

CHAPTER 4: RESULTS

The intent of this study was to look for confirmatory evidence of Mayer's (2005) proposition of boundary conditions for the redundancy principle. This proposition alters the widely accepted premise of a redundancy principle, i.e., that the simultaneous presentation of redundant narration and on-screen text would lead to an increase in extraneous cognitive load, thereby inhibiting the learning process. The proposed boundary condition states that when text-only presentations featuring audio narration are accompanied by fully redundant text that is incrementally revealed, i.e., conditions that are both verbally redundant and temporally contiguous, extraneous cognitive load is actually decreased and learning enhanced (Mayer & Fiorella, 2014). To investigate this issue, a quantitative experiment was designed in which undergraduate students enrolled in two sections of an online class were randomly assigned to one of four treatment conditions, featuring a narrated text-only online presentation including either temporally contiguous text, verbally redundant text, both, or neither. The speed of presentation was system controlled across all treatments to avoid "time spent on slide" emerging as a confounding variable. The dependent variable, student achievement, was measured via a 15-question selected-response instrument assessed both prior to and after the treatment. Tests of the statistical significance of main effects and interactions were evaluated at a .05 significance level.

Research Questions

The current study was designed to address these research questions:

1. Is there a statistically significant main effect of different temporal contiguity styles on students' achievement when controlling for pretest scores?
2. Is there a statistically significant main effect of different verbal redundancy styles on students' achievement when controlling for pretest scores?
3. Is there a statistically significant interaction effect of temporal contiguity and verbal redundancy on students' achievement when controlling for pretest scores?

Examination of Statistical Assumptions

To examine the dependent variable of student achievement relative to the topic of instruction (information literacy), a two-way ANCOVA was conducted featuring two levels of verbal redundancy and temporal contiguity as independent variables. This research design allowed analysis of any interaction between main effect A (temporal contiguity) and main effect B (verbal redundancy). The difficulty of the assessment was sufficient to prevent a ceiling effect, i.e., a positive skew resulting from multiple subjects achieving the maximum score. The dependent variable of "achievement" was aggregated from dichotomous individual scores to generate a continuous scale. Descriptive statistics including the mean scores and standard deviations for each of the four treatment groups were calculated (Table 3). As can be seen from the table, the means were highly similar across treatments.

Table 3*Means and Standard Deviations of All Treatment Groups*

	<i>Pretreatment</i>			<i>Posttreatment</i>		
	M	SD	n	M	SD	n
<i>Control Group</i>	50.57	26.52	29	75.40	20.48	29
<i>Temporal Contiguity</i>	49.89	18.91	29	69.66	21.79	29
<i>Verbal Redundancy</i>	50.34	16.39	29	74.71	15.67	29
<i>Both</i>	51.03	21.66	26	75.13	20.09	26

Multiple assumptions for ANCOVA analysis were evaluated in SPSS:

- A scatterplot was used to visually assess whether there is a linear relationship between the covariate, i.e., the pretest scores, and the dependent variable, i.e., the posttest scores. Plots were produced both for each combination of levels for the independent variables (Figure 2). Loess lines fit to 80% of data points aided in the judgment of linearity, leading to the conclusion that there was a linear relationship between pretest scores and posttest scores concentration for each treatment, as assessed by visual inspection of a scatterplot.

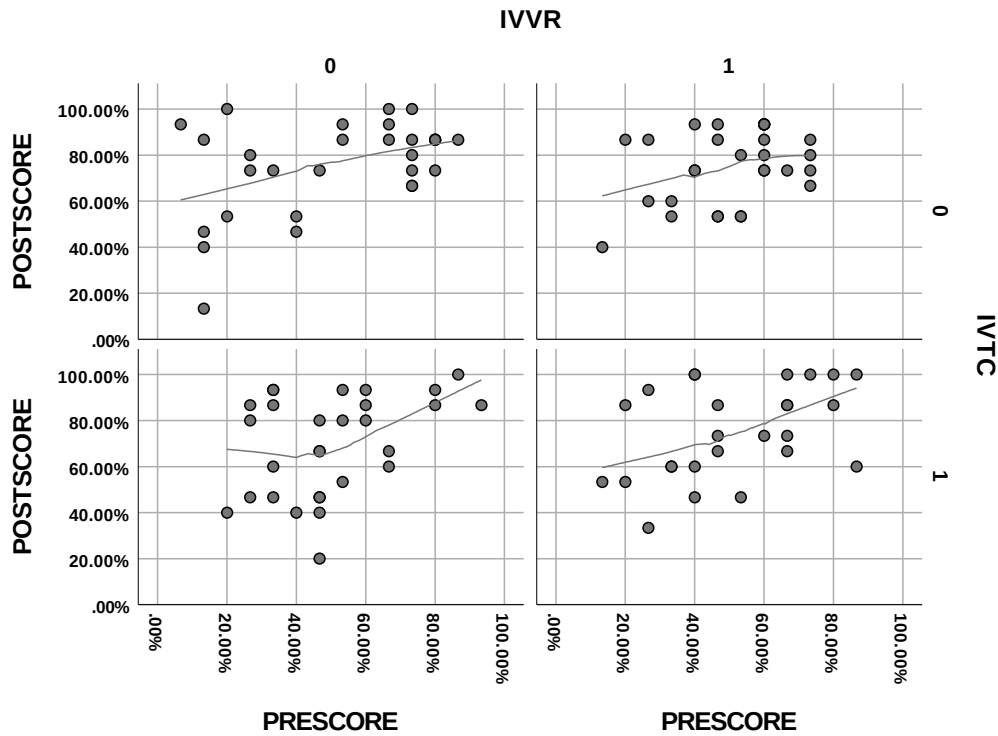


Figure 2

Linearity of Covariate to Each Treatment

- To verify the assumption that the slope of the relationship between the covariate and the dependent variable is similar for each treatment, i.e., homogeneity of regression slopes, the grouping variable “Treatment” was used to evaluate the differences between a two-way ANCOVA and a two-way ANCOVA with interaction terms as described in a test of between-subjects effects. The results were not significant (Table 4). There was homogeneity of regression slopes as determined by a comparison between the two-way ANCOVA model with and without interaction terms, $F(3, 105) = 0.092, p = .964$.

Table 4*Homogeneity of Regression Slopes*

Tests of Between-Subjects Effects					
Dependent Variable: POSTSCORE					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	7437.243^a	7	1062.463	3.168	.004
Intercept	44696.568	1	44696.568	133.270	.000
TREAT	282.447	3	94.149	.281	.839
PRESCORE	6030.326	1	6030.326	17.980	.000
TREAT * PRESCORE	92.764	3	30.921	.092	.964
Error	35215.264	105	335.383		
Total	656222.222	113			
Corrected Total	42652.507	112			

- To verify homoscedasticity of error variances within each combination of levels for the independent variables, a group of scatterplots was generated from SPSS (Figure 3). While one of the four individual plots did indicate some potential issues as evidenced by a tendency towards a decreasing funnel shape, the evidence was not deemed substantial enough to invalidate the assumption of linearity. Therefore, it was concluded that there was homoscedasticity within each combination of groups of the two independent variables, as assessed by visual inspection of the standardized residuals plotted against the predicted values for each group.

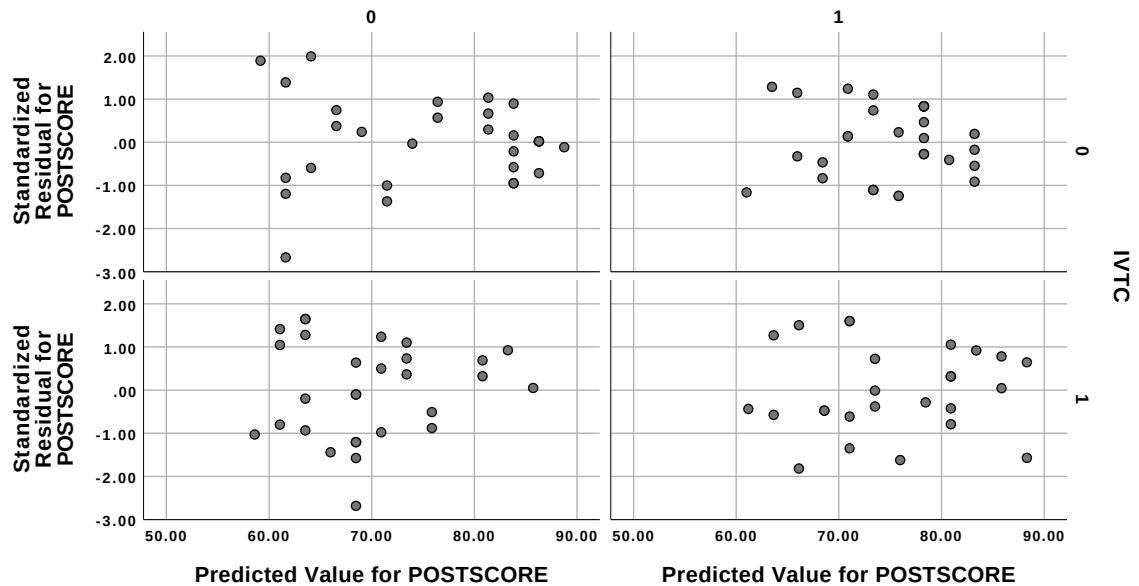


Figure 3

Homoscedasticity of Error Variances

- Testing for homogeneity of variances was conducted within SPSS by generating Levene's test of equality of variances (Table 5). The results were not significant. There was homogeneity of variances, as assessed by Levene's test of homogeneity of variance ($p = .659$).

Table 5

Levene's Test of Equality of Variances

Dependent Variable: POSTSCORE			
F	df1	df2	Sig.
.996	3	109	.397

- Testing for unusual data points involved inspection of the dependent variable's Studentized residuals. The Studentized versions of the residuals were used because they are quantified in standard deviation units, which simplifies the identification of outliers.

There were no outliers remaining in the data, as indicated by no cases greater than ± 3 standard deviations. Inspection of the uncentered leverage values indicated no cases of high-leverage data points. Review of the Cook's distance values determined that there were no influential points.

- Testing for normality for each combination of levels for the two independent variables (IVTC = temporal contiguity; IVVR = verbal redundancy) was conducted via the Shapiro-Wilk test (Table 6). Standardized residuals were normally distributed, as assessed by Shapiro-Wilk test ($p > .05$).

Table 6

Tests of Normality

IVTC	IVVR		Kolmogorov-Smirnov ^a			Shapiro-Wilk	
			Statistic	df	Sig.	Statistic	df
0	0	Standardized Residual for POSTSCORE	.075	29	.200 [*]	.980	29
	1	Standardized Residual for POSTSCORE	.126	29	.200 [*]	.937	29
1	0	Standardized Residual for POSTSCORE	.107	29	.200 [*]	.958	29
	1	Standardized Residual for POSTSCORE	.111	26	.200 [*]	.957	26

Examination of Research Questions

Research Question 1 asked whether temporal contiguity can be associated with a significant main effect on students' achievement when controlling for pretest scores. The current research sought to answer that question by comparing results for the TC and CTRL groups; this comparison found no significant difference in participants' achievement between narrated online presentations that did and did not reveal on-screen text in a temporally contiguous manner. As seen in Table 7, the adjusted marginal mean for achievement in the temporally contiguous group

(69.86%) was lower than the control group (75.35%) but did not represent a statistically significant difference, $F(1, 108) = .612, p = .436$.

Table 7

Adjusted Means and Marginal Means from 2x2 ANCOVA Analysis

		Verbal Redundancy		
		No	Yes	
Temporal Contiguity	No	Cell Mean 1 75.35 (n = 29)	Cell Mean 2 74.75 (n = 29)	Marginal Mean 1 = 75.05
	Yes	Cell Mean 3 69.86 (n = 29)	Cell Mean 4 74.91 (n = 26)	Marginal Mean 2 = 72.39
		Marginal Mean 3 = 72.61	Marginal Mean 4 = 74.83	

Research Question 2 asked whether verbal redundancy can be associated with a significant main effect on students' achievement when controlling for pretest scores. The current research sought to answer that question by comparing results for the VR and CTRL groups; this comparison found no significant difference in participants' achievement between narrated online presentations with and without text that was fully redundant across the audio and visual channels. Adjusted marginal mean for achievement in the verbally redundant group (74.75%) was lower than the control group (75.35%) but did not represent a statistically significant difference, $F(1, 108) = .426, p = .515$.

Research Question 3 asked whether there was an interaction effect of verbal redundancy and temporal contiguity on students' achievement when controlling for pretest scores. Conclusions were based on an evaluation of the two-way ANCOVA results (Table 8) as shown in the tests of between-subjects effects. There was no statistically significant two-way interaction

between temporal contiguity and verbal redundancy on achievement scores while controlling for pretest scores, $F(1, 108) = .689, p = .408$.

Table 8

ANCOVA Tests of Between-Subjects Effects

Dependent Variable: POSTSCORE						
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	7344.479 ^a	4	1836.120	5.616	.000	.172
Intercept	49779.855	1	49779.855	152.266	.000	.585
PRESCORE	6703.230	1	6703.230	20.504	.000	.160
IVTC	200.062	1	200.062	.612	.436	.006
IVVR	139.298	1	139.298	.426	.515	.004
IVTC * IVVR	225.329	1	225.329	.689	.408	.006
Error	35308.028	108	326.926			
Total	656222.222	113				
Corrected Total	42652.507	112				

a. R Squared = .172 (Adjusted R Squared = .142)

Although profile plots of the estimated marginal means for each independent variable indicated some interaction between temporal contiguity and verbal redundancy (Figure 4), this interaction was not significant.

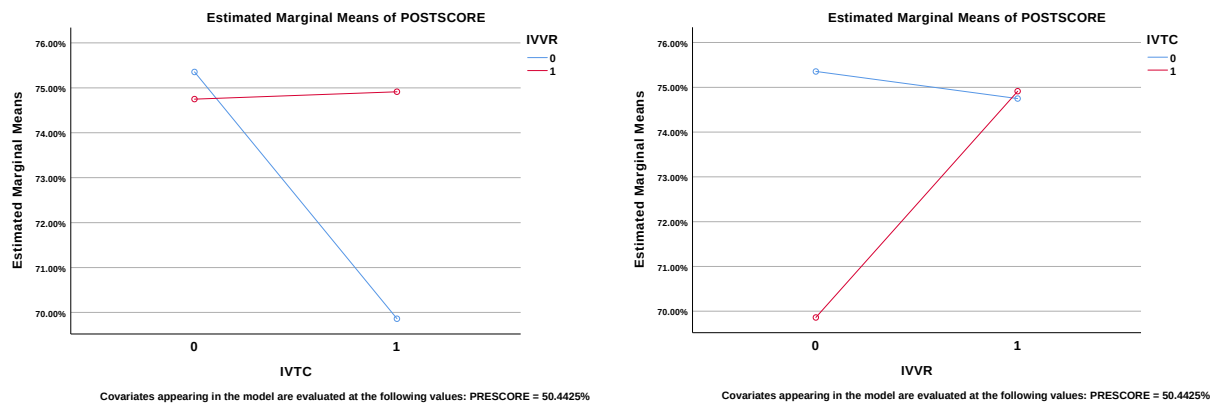


Figure 4

Profile Plots of Estimated Marginal Means for the Dependent Variable

Restricting Dataset to Low-Knowledge Learners

These results indicate that the treatments did not produce a significant effect across the sample as a whole, but Mayer's statements regarding a boundary condition for the redundancy principle have often included caveats that such a condition would be strongest among low-knowledge learners. To investigate whether the current study's outcomes support this supposition, the current study's research questions were evaluated against a subset of responses that featured only those participants whose pretest scores were below the mean of 50.57 ($n = 57$). As with the full set of data, examining the results generated by SPSS continued to show that the results of the experiment's treatments remained insignificant. Among low-knowledge participants, the ANCOVA for temporal contiguity versus verbal redundancy on posttest achievement scores when controlling for pretest scores was found to have no statistically significant main effect for temporal contiguity, $F(1, 52) = .158, p = .692$; verbal redundancy, $F(1, 52) = 1.354, p = .250$; or interactions, $F(1, 52) = .087, p = .769$.

CHAPTER 5: DISCUSSION

As the online delivery of instructional content continues to grow, the need for theoretically based and empirically proven guidance on the design of online presentations has become paramount. The current study addressed the origins of popular conventions surrounding presentation design practices (Chandler & Sweller, 1992; Mayer, 2002; Mayer & Moreno, 1998; Mayer & Moreno, 1999; Penney, 1975; Sweller, 1988), then questioned those assumptions based on an examination of recent theoretical developments in multimedia theory.

After a series of experiments, Mayer (2005), and later Mayer with Fiorella (2014), proposed conditions under which the redundancy principle, i.e., the theory that redundant material decreases the intelligibility of instruction (Chandler & Sweller, 1991), might be reversed. Specifically, the proposed boundary condition investigated by the current study is the idea that presentations which are both temporally contiguous and verbally redundant benefit, rather than impair, the cognitive processes of learners (Mayer & Fiorella, 2014). Yet despite their statement that “the redundancy effect can disappear when no graphics are presented” (Mayer & Fiorella, 2014, p. 299), there remains a gap in the literature when it comes to empirical results from research on text-only narrated slideshows. Inspired by the abundance of text-only presentations in academia, the current research was conducted as a pretest/posttest quantitative experiment to explore Mayer and Fiorella’s proposed boundary conditions for the redundancy effect in conjunction with the use of text-only presentations in which verbal redundancy and temporal contiguity were employed.

Summary

Through a carefully designed quantitative experiment, the interaction and main effects of temporal contiguity and verbal redundancy were assessed within two sections of an existing online class for undergraduate students at a small midwestern university. Without prior disclosure, these students were randomly assigned to one of four treatment conditions: a narrated text-only online presentation using a typical design of minimal words that were fully visible at the start of each slide's narration (the CTRL treatment); slides with the same audio narration and minimal text, but with the words timed to appear in synchronization with the audio (the TC treatment); a similar set of slides where the on-screen text was fully redundant to the narration (the VR treatment); and slides in which the fully redundant text was progressively disclosed in synchronization with the audio (the BOTH treatment). Approximately one week before the treatments were administered, each participant's existing knowledge of the topic of instruction (information literacy) was assessed using a 15-question selected-response instrument developed specially for this research. Later, after each participant watched their assigned treatment, the same instrument was used to collect the posttreatment test data through an online assessment. While the sequence of the questions remained the same for the pretest and posttest, the order of the responses was randomized. As this study was specifically designed to test the two independent variables individually as well as in combination, the findings include the comparability of either verbally redundant text or temporally contiguous text as well as the combination of these two conditions. Two-way ANCOVA analysis of the results revealed that when the pretest scores were controlled for, there was no significant difference in student

achievement among the four treatment conditions. This finding held true even when the analysis was limited to low-knowledge students.

Interpretation

While it is important to acknowledge that the current study's lack of significant results may be due to methodological issues, findings from the current study align with previous research (Jeung, Chandler, & Sweller, 1997; Kalyuga, Chandler, & Sweller, 1999; Savoji, Hassanabadi, & Fasihipour, 2011; Tabbers, Martens, & van Merriënboer, 2004) by concluding that Mayer's proposal of boundary conditions for the redundancy principle is not supported by the current study's empirical evidence. Within the parameters of the current research, i.e., a 22-minute narrated online lecture featuring instruction of a conceptual nature, these results indicate that the traditionally negative conception of the redundancy principle in regards to text-only narrated presentations delivered online is also not supported. This finding implies that learners' cognitive processes are more robust than is often described within cognitive learning theories and that learners are able to discern meaning and initiate processes related to the storage and recall of information without regard to whether the information is relayed in a verbally redundant or temporally contiguous manner. It also implies that the nature of the instructional topic, e.g., conceptual rather than technical or scientific, may affect the amount of cognitive load experienced by learners.

Mayer and Johnson (2008) said that boundary conditions for redundancy would be most significant for low-knowledge learners, but that supposition was not borne out by the current study. Despite this result, the general applicability of the redundancy principle remains closely connected to issues of accessibility. Learners with other challenges, e.g., vision impairment,

hearing impairment, cognitive impairments, or low language proficiency, have traditionally been overlooked in research on multimedia learning, but together they comprise a substantial, growing portion of the student body. Multiple researchers have identified the possibility that a combination of verbally redundant and temporally contiguous presentation styles could be efficacious for learners who use English as a second language; in particular, Toh, Munassar, and Yahaya (2010) showed the value of combining verbal redundancy and temporal contiguity when Yemeni learners were exposed to English language instruction.

Surprisingly, some researchers have decried attempts to elevate the needs of second-language students or learners with disabilities. For example, the influential scholar Sweller has strongly advocated against the use of fully redundant text and narration as a waste of precious cognitive resources. Despite admitting that “information that is redundant for one person may be essential for another” (Sweller, 2005, p. 165), he remains adamant that “information should be presented in a single form only, i.e., with all other versions and all unnecessary explanation eliminated” (p. 167). The current study’s empirical data does not support such a conclusion.

The efficacy of instruction via online and assistive technologies can also be affected by the degree of user control allowed over the speed and timing of the presentation; this is especially pertinent for non-native language speakers and learners with physical or neurological limitations. Early examples of computer-mediated experiments presumed higher cognitive loads when learners were exposed to system-paced presentations. However, the results of several experiments described by Mayer (2014, p. 392) indicate that even when the pace of presentations is beyond the control of the participant, the expected increase in cognitive load arising from verbally redundant text and narration can be managed through the use of temporal contiguity,

i.e., progressive disclosure of the text in synchronization with the narration. System-controlled timing was employed throughout the current study's text-only presentations, but it appears to have had little impact on cognitive load. It is possible that the purely conceptual content of the instruction was not sufficiently challenging to reveal the benefits of a temporally contiguous and verbally redundant presentation style; more significant differences between treatments may have emerged if the material covered were more technical or if the presence of graphics had increased the overall cognitive load.

As previously described, the landmark results of early research into cognitive learning principles have proven difficult to replicate, inspiring questions as to whether we should strictly adhere to CTML's principles when multiple empirical studies can offer only a mixed record of support. As early as 1975, Penney reviewed published studies showing that learners could best remember lists in short-term memory when the information was presented auditorily rather than visually, i.e., the modality effect, but contemporaneously to her research, dissenting studies found visual presentation to produce superior outcomes (Kroll, Parkinson, & Parks, 1972; Marcer, 1967; Scarborough, 1972). Since that time, numerous experiments designed to investigate individual multimedia effects have produced contrary or inconclusive results (Jeung, Chandler, & Sweller, 1997; Kalyuga, Chandler, & Sweller, 1999; Leahy & Sweller, 2011; Savoji, Hassanabadi, & Fasihipour, 2011; Tabbers et al., 2004) that cast doubt on the immutability of CTML's oft-cited principles. Research conducted by Tabbers, Martens and van Merriënboer (2004) found that use of visual learning material was superior to audio in terms of student transfer and retention: "Replacing visual text with spoken text even had a negative effect

on learning, contrary to what both cognitive load theory and Mayer's theory of multimedia learning would predict" (p. 80).

Delimitations, Limitations, and Assumptions

Delimitations

As the objective of this study was to examine the effect of verbal redundancy and temporal contiguity on achievement within the context of narrated text-only presentations, the current research consciously avoided the use of graphics such as diagrams or animations in favor of a tight focus on the interaction of narration and unaccompanied text. Similarly, I did not experiment with variations in graphic design elements, e.g., highlight colors or font selection, or the use of sound effects. While previous studies have involved a broad range of participant ages, this research was restricted to first-year university students enrolled in an online general education course; by doing so, it allowed me to draw a convenience sample from a group of subjects who had purposely selected a distance learning experience. The instrument developed for the current study was composed entirely of selected-response multiple-choice questions rather than open-ended or task-based assessments to allow for automatic scoring of subject responses that could eliminate potential scoring variability due to researcher bias. This choice coincided with the decision to use an experimental quantitative approach in conjunction with a postpositivist realist perspective, so that the current study could consider its subjects' performance within a sociopolitical context.

Limitations

Internal Validity Issues. This research is subject to the constraints of quantitative experimental research design and statistical analysis, including the limitations of each

treatment's sample size and the satisfaction of assumptions for ANCOVA analysis, i.e., independence of observations and multivariate normality as well as homogeneous variances and correlations. Acceptance of the research instrument, a 15-question selected response online quiz, also assumes that the forms of evidence offered in support of validation are adequate to ensure accurate assessment and extensible findings. The veracity of results from the current study's pretest/posttest experimental design may also have been limited by testing threat, where participants were able to learn about the subject via the pretest prior to being exposed to a specific treatment. Distractions in students' immediate surroundings while online as well as the speed and reliability of each student's Internet access were additional concerns.

External Validity Issues. Participants were drawn from a convenience sample of students participating in a popular multisection online class at a midwestern university ($m = 21$, range 17-39). The ecological validity of the proposed experiment was aided by a waiver of informed consent; this helps assure that participants are thoroughly engaged in the learning process, as the instrument is presented within the context of a class assignment within an actual online course. The suitability of this sample group combined with the authentic environment of the test and an adequate sample size should allow the extensibility of this study's findings to college classroom environments. Additionally, this experiment was carried out during the onset of the coronavirus pandemic, an additional stress factor that may have affected students' participation.

Additional Limitations. While the topic of this instructional module was not unfamiliar to me, I do not have formal training in the field of information literacy; more personal familiarity with this subject might have generated a more challenging assessment rather than the low-

difficulty questions used in this research. The instructional content of the presentations was conceptual in nature, rather than technical, diagrammatic, or procedural; this may have contributed to the ease with which subjects were able to learn the material without experiencing substantial cognitive load. It is also unclear whether ramifications of the coronavirus pandemic may have affected subject responses, as the current study's experiments took place just as COVID-19 was driving Americans into self-isolation.

Performing an ANCOVA analysis based on traditional test scoring meant that all assessment items were weighted equally. Alternative approaches, e.g., item response theory or factor analysis, could have made it possible to weight the difficulty of each question; unfortunately, the current study's sample size falls below the threshold for a credible item response analysis. The multiple-choice questions developed for the instrument addressed only the lowest levels of Bloom's taxonomy of cognitive thought processes, i.e., knowledge and comprehension, so the impact of this study's independent variables on higher order cognitive processes went unexamined. On a related note, the absence of questions related to subjects' experience of the treatment meant that preferential issues regarding learner affinity towards various presentation styles were not examined. Perhaps most importantly, the low number of subjects per treatment ($n \geq 26$) limits the power of this study's significance testing, thereby increasing the possibility of a Type II error. Additionally, the sample did not include a significant number of non-native English speakers ($n = 9$) and the number of learners with cognitive impairments is unknown; these limitations mean that it is not possible to draw conclusions regarding such groups. These limitations are important to consider before Mayer's proposition of boundary conditions for the redundancy principle is reconsidered.

Assumptions

Even to the constructivist researcher, the utility of cognitivist theories cannot be denied. In particular, the assumption of cognitive load, i.e., that human information processing has a finite limit that can be overburdened, was embraced by this research. The legitimacy of multiple related theoretical principles and effects, e.g., modality effect, dual coding, split-attention effect, and the redundancy effect, was also assumed, with recognition that the definition of these principles remains an ongoing concern. In reflection, however, the results of the current study should be interpreted as casting doubt on the strength of these principles and effects within online narrated text-only presentations administered to first-year university students.

On a more practical level, this study assumes that its research subjects participated honestly and were motivated only by a desire for high personal achievement, an assumption that was aided by this study's waiver of informed consent prior to the classroom-based experiment. An assumption is also made that extensible results were obtained from the sample size of $n \geq 26$ per treatment, given that the prerequisites for an analysis of covariance were satisfied. The validity and reliability of the instrument developed for this study is also assumed, subject to the validity evidence provided in Appendix A.

Implications

The postpositivist realist perspective of the current study foregrounds the need to consider sociopolitical contexts, including the existence of learners who experience challenges in accessing online instruction. Relevant teaching theories such as Universal Design for Learning (Rose et al., 2006) urge us to adopt the most broadly accessible approach in every situation, as opposed to the most familiar, the most convenient, or the most exclusive. Most aspects of CTML

are already compatible with Universal Design for Learning: the multimedia effect, contiguity, coherence, pretraining, signaling, and personalization are all helpful to learners with disabilities; only CTML's modality and redundancy effects preclude the use of narration with redundant on-screen text. Despite the inability of the current research and some prior studies to substantiate verbal redundancy's boundary conditions for typical participants, the use of verbally redundant multimedia presentations offers the advantage of freeing hearing-impaired learners as well as those studying in noisy environments from the need for captions. (In such cases, however, closed captions should still be made available for use with assistive technologies such as Braille terminals.) This provides just one example of how practitioners should carefully weigh their available design choices, regardless of whether specific CTML principles can be proven beneficial to an idealized "standard" learner.

As the findings of this research do not support the proposition of a boundary effect for the redundancy principle within these constricted conditions, it could be possible to advocate for the abandonment of presentation designs that are both temporally contiguous and verbally redundant. However, it is important to consider that the current study's lack of significant findings may be due to methodological issues rather than a true explication of the issues. Instead, this study calls for further research into this important issue, with the recommendation that content creators who are willing to perform the additional programming needed to create presentations that are both verbally redundant and temporally contiguous should continue to do so until the potential benefits for non-native English speakers and learners with cognitive challenges are better understood.

Recommendations for Further Research

Additional research is needed to quantify the validity and applicability of Mayer's multimedia learning principles to both general and nontraditional audiences. Specifically, more quantitative research must be conducted on a variety of learners within actual online courses so that our accepted approaches to multimedia design may be contrasted with recent scholarship expanding or challenging those practices. Further discussion regarding the implications of multimedia theory on current practices should provide meaningful insights into our fundamental assumptions regarding instructional design and its impact, if any, on learners' underlying cognitive processes.

The academic community would be well served by further studies on boundary conditions for the redundancy principle within online narrated text-only instruction that incorporate the following attributes:

1. Research based on substantially larger sample sizes, which would both increase the statistical power of the observations and allow for elimination of the pretest.
2. Research involving special populations, including non-native English speakers as well as learners facing a range of common cognitive challenges such as dyslexia, attention deficit disorder, and low vision.
3. Research that contrasts conceptual topics with more technical, scientific, or procedural content.
4. Research using more challenging assessments that are capable of addressing participants' higher order cognitive processes.

5. Research in which results are analyzed using item response theory techniques to enable estimation of question difficulty, guessing, and individual learner ability.
6. Research involving additional age groups.
7. Research that captures learner preference from among various presentation styles.

Conclusions

Studies that began well before the advent of the Internet have long cast shadows onto our current practices. By relying on theoretical rationales from the past century, we have codified these potentially outmoded concepts into rigid principles founded more on tradition than empirical insights. As the use of online instruction expands, not only in the number of learners it serves but also in the kinds of learners it can include, we should feel compelled to revisit the fundamental paradigms underlying our pedagogies and praxis.

Through a carefully constructed experiment, this study sought evidence of boundary conditions under which the redundancy principle would no longer apply. While no evidence of a reverse redundancy effect was found, there was also no support shown for the negative implications of redundancy as it has traditionally been defined.

The most compelling conclusion derived from the current research is to acknowledge that previously held beliefs in the *negative* impact of redundant text and narration, i.e., the redundancy principle, may not be applicable to text-only online presentations. While temporal contiguity and verbal redundancy were not proven to be beneficial to learner outcomes, neither were they shown to be harmful. The experiment conducted for the current study highlights the robust nature of human comprehension and indicates that established cognitivist absolutes

regarding the behavior of these complex learning processes may not be empirically supported for narrated text-only online instruction.

REFERENCES

- Anderson, R. C. (1977). The notion of schemata and the educational enterprise. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.) *Schooling and the Acquisition of Knowledge*. Erlbaum.
- Atkinson, R. C., & Shiffrin, R. M. (1968). Chapter: Human memory: A proposed system and its control processes. In Spence, K. W., & Spence, J. T. *The psychology of learning and motivation* (Volume 2). Academic Press. pp. 89–195.
- Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8(4), 293-332.
- Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, 62, 233-246.
- de Jong, T. (2010). Cognitive load theory, educational research, and instructional design: some food for thought. *Instructional Science*, 38(2), 105–134.
- Dewey, J. (1900) *The school and society*. University of Chicago Press.
- Dewey, J. (1938). *Experience and Education*. Simon & Schuster, Inc.
- Drasgow, F., & Lissak, R. I. (1983). Modified parallel analysis: A procedure for examining the latent dimensionality of dichotomously scored item responses. *Journal of Applied Psychology*, 68(3), 363–373.
- Dwyer, F. M., & Lamberski, R. (1977). *The human heart: Parts of the heart, circulation of blood and cycle of blood pressure*. Learning Services.
- Fives, H., & DiDonato-Barnes, N. (2013). Classroom test construction: the power of a table of specifications. *Practical Assessment, Research and Evaluation* 18, 2–7.
- Fleming, M., & Levie, W. H. (1978). *Instructional message design: Principles from the behavioral sciences*. Educational Technology Publications.
- Frish, A., Camerini, L., & Schulz, P. J. (2013). The impact of presentation style on the retention of online health information: A randomized-controlled experiment. *Health Communication*, 28, 286–293.
- Garner, J. K., & Alley, M. (2011). PowerPoint in the psychology classroom: Lessons from multimedia learning research. *Psychology Learning and Teaching*, 10, 95–106.

- Garner J. K., Alley, M., Gaudelli, A. F., & Zappe, S. E. (2009). Common use of PowerPoint versus the assertion–evidence structure: A cognitive psychology perspective. *Technical Communication*, 56(4), 331-345.
- He, P., Liu, X., Zheng, C., & Jia, M. (2016). Using Rasch measurement to validate an instrument for measuring the quality of classroom teaching in secondary chemistry lessons. *Chem. Educ. Res. Pract.*, 17, 381-393.
- Jeung, H., Chandler, P., & Sweller, J. (1997). The role of visual indicators in dual sensory mode instruction. *Educational Psychology*, 17, 329–343.
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Harvard University Press.
- Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, 13, 351–371.
- Kress, G. (2005). Gains and losses: New forms of texts, knowledge, and learning. *Computers and Composition*, 22(1), 5-22.
- Kroll, N. E. A., Parkinson, S. R., & Parks, T. E. (1972). Sensory and active storage of compound visual and auditory stimuli. *Journal of Experimental Psychology*, 95 (32-38).
- Leahy, W., & Sweller, J. (2011). Cognitive Load theory, modality of presentation and the transient information effect. *Applied Cognitive Psychology*, 25 (943–951).
- Marcer, D. (1967). The effect of presentation method on short-term recall of CCC trigrams. *Psychonomic Science*, 8 (335-336).
- Mayer, R. E. (2002). Multimedia learning. *Psychology of Learning and Motivation*, 41, 85-139.
- Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 201-212). Cambridge University Press.
- Mayer, R. E. (2014). Multimedia instruction. In J. Michael Spector, M. David Merrill, Jan Elen, M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (pp. 385-399). Springer.
- Mayer, R. E. (2017). Using multimedia for e-learning. *Journal of Computer Assisted Learning*, 33, 403-423.
- Mayer, R. E., Bove, W., Bryman, A., Mars, R., and Tapangco, L. (1996). When less is more: Meaningful learning from visual and verbal summaries of Science textbook lessons. *Journal of Educational Psychology*, 88(1), 64-73.

- Mayer, R. E., & Fiorella, L. (2014). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 279-315). New York: Cambridge University Press.
- Mayer, R. E., & Johnson, C. I. (2008). Revising the redundancy principle in multimedia learning. *Journal of Educational Psychology, 100*(2), 380–386.
- Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology, 90*(2), 312-320.
- McLaren, B. M., DeLeeuw, K. E., & Mayer, R. E. (2011b). A politeness effect in learning with web-based intelligent tutors. *International Journal of Human-Computer Studies, 69*, 70–79.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*(2), 81-97.
- Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology, 91*(2), 358-368.
- Moreno, R., & Mayer, R. E. (2002). Verbal redundancy in multimedia learning: When reading helps listening. *Journal of Educational Psychology, 94*(1), 156-163.
- Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology, 87*, 319-334.
- Orlikowski, W.J., & Yates, J. (1994). Genre repertoire: Examining the structuring of communicative practices in organizations. *Administrative Science Quarterly, 39*, 541-574.
- Paivio, A. (1986). *Mental Representations: A Dual-Coding Approach*. Oxford University Press.
- Penney, C. G. (1975). Modality effects in short-term verbal memory. *Psychological Bulletin, 82*(1), 68-84.
- Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory & Cognition, 17*(4), 398-422.
- Reese, S. D. (1984). Visual-verbal redundancy effects on television news learning. *Journal of Broadcasting, 28*(1), 79-87.
- Reeves, T. D., & Marbach-Ad, G. (2016). Contemporary test validity in theory and practice: A primer for discipline-based education researchers. *CBE-Life Sciences Education, 15*, 1-9.

- Reiser, R. A., & Gagné, R. M. (1983). *Selecting media for instruction*. Educational Technology.
- Rose, D., Harbour, W., Johnston, C. S., & Abarbanell, L. (2006). Universal design for learning in postsecondary education: Reflections on principles and their application. *Journal of Postsecondary Education and Disability*, 19(2), 135-151.
- Savoji, A. P., Hassanabadi, H., & Fasihipour, Z. (2011). The modality effect in learner-paced multimedia learning. *Procedia - Social and Behavioral Sciences*, 30, 1488 – 1493.
- Scarborough, D. L. (1972). Stimulus modality effects on forgetting in short-term memory. *Journal of Experimental Psychology*, 95 (285-289).
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2001). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.
- Smith, J. K. (1983, March). Quantitative versus qualitative research: An attempt to clarify the issue. *Educational Researcher*, 6-13.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science* 12, 257-285.
- Sweller, J. (2005). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 159-168). Cambridge University Press.
- Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load as a factor in the structuring of technical material. *Journal of Experimental Psychology*, 119, 176–192.
- Tabbers, H. K., Martens, R. L., & van Merriënboer, J. J. G. (2004). Multimedia instructions and cognitive load theory: Effects of modality and cueing. *British Journal of Educational Psychology*, 74, 71–81.
- Tarmizi, R. A., & Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology*, 80(4), 424–436.
- Tindall-Ford, S., Chandler, P., & Sweller, J. (1997). When two sensory modes are better than one. *Journal of Experimental Psychology: Applied*, 3(4), 257–287.
- Toh, S. C., Munassar, W. A. S., & Yahaya, W. A. J. W. (2010). Redundancy effect in multimedia learning: A closer look. In proceedings of ASCILITE - Australian Society for Computers in Learning in Tertiary Education Annual Conference 2010 (pp. 988-998).
- Tufte, E. (2003). *The Cognitive Style of PowerPoint*. Graphics Press.
- van den Broek, G. S. E., Segers, E., & Verhoeven, L. (2014). Effects of text modality in multimedia presentations on written and oral performance. *Journal of Computer Assisted Learning*, 30(5), 438-449.

- Wang, N., Johnson, W. L., Mayer, R. E., Rizzo, P., Shaw, E., & Collins, H. (2008). The politeness effect: Pedagogical agents and learning outcomes. *International Journal of Human- Computer Studies*, 66, 98–112.
- Woolgar, S., Pawluch, P. (1985). Ontological gerrymandering: The anatomy of social problems explanations. *Social Problems*, 32(3), 214–227.

APPENDIX A

DETAILS OF PILOT TESTING AND INSTRUMENT VALIDATION

During the spring and summer of 2019, I conducted then analyzed an initial pilot test for my dissertation experiment using an existing instrument developed by Dr. Frank Dwyer. The results of the pilot led me to determine that the Dwyer materials were not a good fit with the nature of this experiment, which involved the delivery of text-only narrated presentations as generally used for conceptual instruction. The topics discussed in the Dwyer materials seemed too spatial to be properly conveyed via text-only materials. As a result, this instrument was rejected.

Search for an alternate established instrument began. Before resigning myself to the need to create a new instrument, I first conducted an analysis of the existing course content covered within the online presentation this experiment would replace. Given the opportunity presented by this research venue (a fully online class), my search was for materials that could be delivered in an impactful one-shot online presentation, followed by an assessment that could be automatically graded to avoid subjectivity in grading, e.g., using only selected response questions.

Although my investigation turned up a wealth of instructional materials on the topic of information literacy, the number of available assessments was very low. Of the assessments located, few appeared to have been used by more than one instructor. These included the Tools for Real-Time Assessment of Information Literacy Skills (TRAILS) from Kent State University Libraries, and the Information Literacy Competency Standards for Higher Education (ILCSHE), as developed by the Association of College & Research Libraries (ACRL). Although the TRAILS materials have been used since 2004, I found that the scenario-based questionnaires used to assess the participants would require far more instructional time than my ETT 234 opportunity would allow.

After being unable to locate a suitable existing instrument, I began the development of a new instrument. This effort was shaped by the opportunity for which it was created: as the coordinator for a general education class, I obtained permission from the ETRA department chair to replace the existing course's presentation on Information Literacy and its accompanying assessment with a new presentation of my own design. Both the presentation and the assessment were prepared in time for inclusion in two Fall 2019 sections of this course. My intent was to use this opportunity to validate the research instrument (the assessment) as well as pilot test one of the four treatments to be used in the experiment (a presentation that featured both verbal redundancy and temporal contiguity). Unfortunately, the instrument's pilot test produced highly skewed results, with a large number of participants achieving perfect or near-perfect scores. Working with my committee, the decision was made to revise the instrument, then conduct a second pilot test. The remainder of this section provides evidence to support the validity of the revised research instrument's test content and internal structure.

Procedure 1: Table of Specifications. Reeves and Marbach-Ad (2016) provided a primer on contemporary practices in test validation for education researchers in which the development of a table of specifications (Fives & DiDonato-Barnes, 2013) was presented as one form of validity evidence. This "test blueprint" provides a structure for developing test questions that align with both content and cognitive dimensions (Reeves & Marbach-Ad, 2016, p. 3). A table of specifications (ToS) for the information literacy assessment was developed, based on seven learning objectives derived from ACRL's ILCSHE. Next, I authored questions that would align with the ToS. Finally, I created text-only PowerPoint slides to present the instructional content

that aligned with the ToS. Table 9 cross-references this content against two levels of cognitive thought processes taken from Bloom's taxonomy.

Table 9

Information Literacy Table of Specifications

<i>Content</i>	<i>Knowledge</i>	<i>Compre- hension</i>	<i>Application</i>	<i>Analysis</i>	<i>Synthesis</i>	<i>Evaluation</i>	<i>Total</i>
<i>A) Searching for Sources</i>	2(1, 3, 5)	1(8)					2(4)
<i>B) Evaluating Sources</i>	3(6, 9)						1(2)
<i>C) Using Sources</i>	4(10)						1(1)
<i>D) Respecting Intellectual Property</i>	5(11); 6(12)						2(2)
<i>E) Citing Sources</i>		7(13, 14, 15)					1(3)
<i>Total</i>	5(8)	2(4)					7(12)

The *learning objectives* to which these questions relate are identified as follows:

1. Recognize the primary characteristics of sources and search methods. (Based on the ILCSHE standard: "...determines the nature and extent of the information needed.")
2. Recall the steps of an effective research search. (Based on the ILCSHE standard: "...accesses needed information effectively and efficiently.")
3. Identify the attributes of scholarly sources. (Based on the ACRL ILCSHE standard: "... evaluates information and its sources critically and incorporates selected information into his or her knowledge base and value system.")

4. Recognize the effective use of sources. (Based on the ACRL ILCSHE standard: "...uses information effectively to accomplish a specific purpose.")
5. Identify when copyright law applies. (Based on the ILCSHE standard: "... understands many of the economic, legal, and social issues surrounding the use of information and accesses and uses information ethically and legally.")
6. Recall the relationship between plagiarism and citation. (Based on the ILCSHE standard: "... understands many of the economic, legal, and social issues surrounding the use of information and accesses and uses information ethically and legally.")
7. Recognize correct APA formatting style. (Based on the ILCSHE standard: "...uses information effectively to accomplish a specific purpose.")

Table 9, the finalized ToS, reflects the feedback described in the next section. Only the highest cognitive level for each learning objective is shown.

Procedure 2: External Systematic Review by Subject Matter Experts. A team of subject matter experts (SMEs) was assembled by posting requests for help to two Facebook groups dedicated to instructional designers; an invitation to participate in this survey was posted to Facebook on January 14, 2020 (Appendix B). These SMEs provided feedback on the questions in the instrument as well as the corresponding portions of the instructional presentation: Thomas Smith, Professor; Tina M. Kister, Information Developer; Pamela Bryan Williams, Director, Learning Design and Development; Jean Marrapodi, Chief Learning Architect; Leann Poston, Instructor; Gailey Tugadi, Instructional Designer; Kim Muephy, Associate Director, Instructional Design and Capabilities; and Rema Merick, Instructional Designer.

A Qualtrics survey form was created to capture the reactions of the SMEs. The Qualtrics form began with this introduction:

Thank you for providing your information. Next, you'll be shown one of the 16 questions in this test (the correct answer is always shown first), accompanied by the instructional presentation's relevant PowerPoint slide (some PPT slides are related to more than one question). Please note that the PPT slides are purposefully text-only, as this is part of an experiment to test the effectiveness of various presentation styles for the delivery of text-only content. After viewing the two items, please rate the correlation of the instructional content to the question, then rate the degree of correspondence to the desired educational outcome, and finally give any feedback that you would like to provide.

Each SME responded by leaving comments on the Qualtrics survey form (Appendix C). This data was collected and analyzed, leading to the determination that two items (Q15 and Q16) were flawed and should be replaced. The SMEs' remaining suggestions were implemented, leading to updates in the assessment, the instructional presentation, and the learning objectives. The full list of pertinent suggestions from the reviewers is shown in Appendix D.

Procedure 3: Analysis of Pilot Test Results. Following the results of the external systematic review, the updated questions, instructional slides, and matching narration were converted into a SCORM-compatible interactive assessment. This was accomplished with iSpring, a software plug-in for PowerPoint that converts slides into HTML5 presentations for viewing within an iSpring-generated "player" (viewing environment). The iSpring player allowed for a degree of control over the viewing experience, including the ability to prevent students from skipping ahead or returning to previous slides once the assessment began. The sequence of questions was set in advance, but the sequence of responses was randomized. The pilot test treatment was designed to exhibit the characteristics of both verbal redundancy (the full text of the narration was visible on each slide) and temporal contiguity (each bullet point appeared in synchronization with the recorded narration). My department chair approved the

inclusion of this information literacy module as an extra credit assignment within another general education course. This assignment (the final pilot test of the instrument) was made available to all four sections of the course on Monday, March 2, 2020.

Participants: This pilot testing occurred within a general education course offered by the university. As is typical for a general education course, its enrollees hailed from a wide variety of majors, including many students from the Nursing, Kinesiology, Communications, and Business programs. As this pilot test featured the same treatment for all participants, IRB approval was not sought. The students were unaware that these modules were anything other than a standard part of the course.

Sample Size: A total of 53 students participated in the pilot test, which launched on March 2 and accepted participants until March 8. Two cases were considered invalid and deleted due to total test-taking times of less than 2.5 minutes, leaving a total of $n = 51$.

Evidence of test content validity from Procedure 1 (table of specifications). The table of specifications created at the outset of this instruments' development (Table 9) provided assurance of two important considerations: that the assessment questions contained within the instruments would speak to major aspects of the topic and that these topics would be addressed using multiple cognitive processes. By linking key content areas directly to a standard from a reputable source, this table of specifications serves as evidence of validity for this instrument.

Evidence of test content validity from Procedure 2 (SME review). Working directly with a panel of experts allowed for independent external review of both the instructional materials and the assessment instruments developed for the current research. Multiple panel members suggested alterations to the assessment questions and instructional materials; these modifications

were carried out prior to the pilot test. The reactions of the SMEs and the remediation of the issues they identified provide additional evidence of validity for these instruments.

Evidence of internal structure validity from Procedure 3 (pilot test analysis). Following my interactions with the SMEs, the revised assessment questions and the corresponding content were published to Blackboard in the form of an iSpring-generated SCORM package. On March 2, 2020, the information literacy module was made available to students as an extra credit assignment. Students were told that they would receive extra credit only if they passed the assessment with a 70% score or higher. They were also told that they could watch the presentation, then take the assessment multiple times until they were able to pass. Only the first attempt of each participant was used in the following analysis. There were no reported problems with the technology and no complaints about the format of the presentations or the assessment questions.

Internal consistency reliability was evaluated via Cronbach's alpha using SPSS, producing an acceptable value of .706 for the initial 16 items. Rasch modeling was also used to examine the pilot test data within R. Calculating the overall goodness-of-fit produced a problematic value of 0.16. Calculating item fit indicated a problem with Item 12 ($p = .048$). In response to this concern, these tests were run again with Item 12 omitted. Analysis of the reduced dataset showed an improved Cronbach's alpha of .713 for 15 items. In R, the Rasch overall goodness-of-fit increased to 0.24. Item fit analysis showed above-threshold values for all 15 items. In examining statistics on local dependence, the pair values in the upper diagonal never exceeded the threshold of 2.0, indicating that there was no cause for concern. The lower half of the matrix contained many locally dependent items at the .05 significance level, but these alone

are not enough reason to say that this data fails the assumption of local independence. Plotting the remaining 15-Item Characteristic Curves sans Q12 revealed the instrument's tendency towards low-difficulty questions (Figure 5).

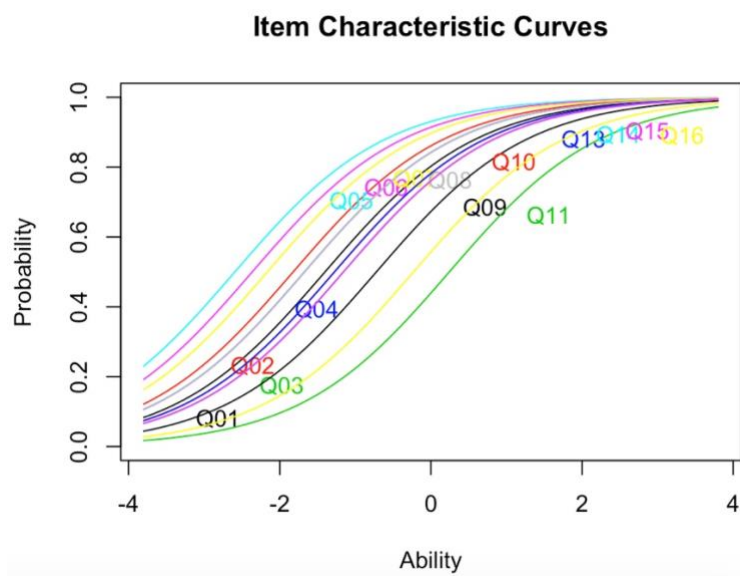


Figure 5

Item Characteristic Curves from the 15-item Dataset

He, Liu, Zheng, and Jia (2016) note, “Items and item responses are examined in Rasch modeling for their degree of fit between the person responses and the measurement model. The mean square residual (MNSQ) and the standardized mean square residual (ZSTD) are typically used as the fit indices to examine how well each item is coherent with the Rasch model” (p. 15).

In the current research, examining infit mean square residual fit statistics for items that are “overfitting” (less than .7 infit MSQ) or “underfitting” (more than 1.3 infit MSQ) relative to model expectations revealed underfit on Q01, which was just beyond the threshold at 1.312 (Table 10). Outfit mean square residual fit statistics also indicated underfit on Q01 (1.320) as

well as Q14 (1.564). However, these issues were not considered significant enough to force invalidation of these questions. Drasgow and Lissak (1983) calculated a test p -value of 0.673, indicating that the 15-item dataset satisfies the assumption of unidimensionality.

Table 10

Item Fit Statistics for the 15-item Dataset

Itemfit Statistics:								
	Chisq	df	p-value	OutfitM	InfitM	Outfit t	Infit t	Discrim
Q01	62.048	46	0.057	1.320	1.312	1.058	1.564	0.025
Q02	35.713	46	0.863	0.760	0.897	-0.540	-0.368	0.506
Q03	30.101	46	0.966	0.640	0.760	-1.482	-1.513	0.621
Q04	40.082	46	0.717	0.853	0.860	-0.314	-0.597	0.490
Q05	42.736	46	0.610	0.909	1.038	0.051	0.224	0.298
Q06	49.590	46	0.332	1.055	1.081	0.280	0.354	0.220
Q07	32.405	46	0.935	0.689	0.846	-0.572	-0.475	0.498
Q08	33.985	46	0.905	0.723	0.834	-0.740	-0.730	0.557
Q09	62.735	46	0.051	1.335	1.076	1.504	0.585	0.229
Q10	35.501	46	0.869	0.755	0.894	-0.854	-0.561	0.524
Q11	44.204	46	0.548	0.941	0.985	-0.210	-0.086	0.288
Q13	45.064	46	0.511	0.959	0.851	-0.054	-0.828	0.533
Q14	73.524	46	0.006	1.564	1.197	1.918	1.166	0.110
Q15	54.582	46	0.181	1.161	1.101	0.669	0.644	0.244
Q16	42.026	46	0.639	0.894	0.924	-0.511	-0.618	0.409

Visual examination of a person-item map shows good distribution of the assessment items, but only across a restricted range (Figure 6). The results described in this section are offered as evidence of this instrument's validity.

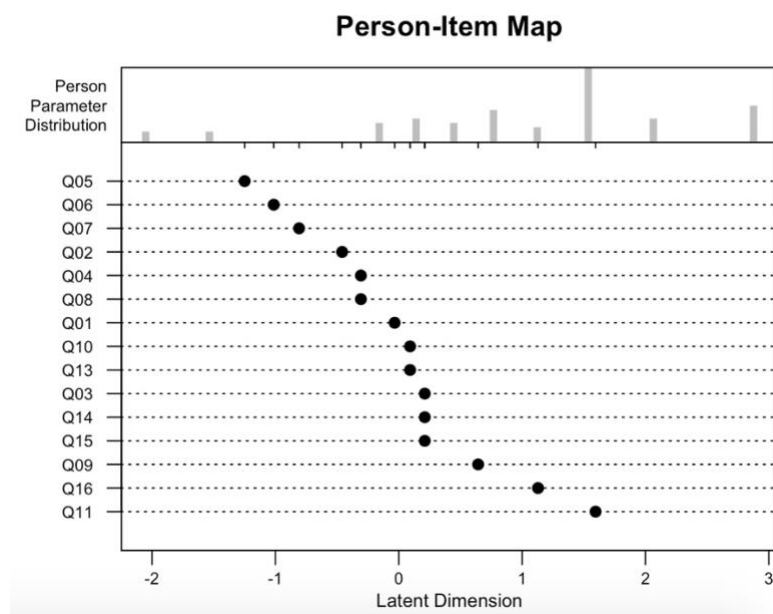


Figure 6

Person-Item Map for the 15-Item Dataset

Discussion of validity procedures. While this validity testing did not exhaust the full range of evidence-gathering procedures outlined by Reeves and Marbach-Ad (2016), it has presented three methods of evaluation. Through the construction of a table of specifications, an expert panel review, and an analysis of pilot testing results, the current research presents multiple forms of evidence that support the validity of this new instrument. Despite these precautions, important limitations persist in that the usability of this study's results is subject to the satisfaction of assumptions for ANCOVA analysis and that adequate evidence has been presented to support the validity of its newly created research instrument.

APPENDIX B

DIRECTIONS TO SUBJECT MATTER EXPERTS



Hal Hinderliter shared a link.



January 14

Hi everyone! Hey, would you please take 15 minutes (or thereabouts) to evaluate my new instrument? I've built a survey form that will show you 16 questions and the supporting instructional content for each assessment item, then allow you to give feedback. This is for my Dissertation research at NIU. Thanks in advance for any help you might be able to provide!! Here's the link to the survey:

NIU.AZ1.QUALTRICS.COM



Test Design Evaluation - Hal Hinderliter

Questions for evaluating an online instructional presentation and...



2

3 Comments



Like



Comment



Rema Merrick I'll take it. 😊

Love · Reply · 7w



1



Hal Hinderliter Rema, thanks so much! I will really appreciate your help 😊

Like · Reply · 7w



1



Rema Merrick Hal Hinderliter - I apologize for the delay in completing your survey. I finally finished it. Thanks for sharing it with the group and allowing us to be part of your dissertation research. Please keep up updated.

Like · Reply · 5w



1



Write a reply...



APPENDIX C

QUALTRICS FORM USED TO COLLECT SME FEEDBACK

niu.az1.qualtrics.com/jfe/form/SV_ddpBQNQzx1O5yQJ?fbcl

Test Design Evaluation - Hal Hinderliter

Thank you for taking part in this Test Design Evaluation! Before reviewing the 16-question test and related instructional content, please provide a few bits of information about yourself...

Your name:

Your email address:

Your job title:

Please provide a short description of your terminal degree in Instructional Technology (or relevant experience):

Thank you for providing your information. Next, you'll be shown one of the 16 questions in this test (the correct answer is always shown first), accompanied by the instructional presentation's relevant PowerPoint slide (some PPT slides are related to more than one question). *Please note that the PPT slides are purposefully text-only, as this is part of an experiment to test the effectiveness of various presentation styles for the delivery of text-only content.*

After viewing the two items, please rate the correlation of the instructional content to the question, then rate the degree of correspondence to the desired educational outcome, and finally give any feedback that you would like to provide.

If you would like to view the entire instructional presentation as a PDF (optional), please download it from <http://www.halhinderliter.com/img/Infoliteracy.pdf>. Let's get started!

→

Please review the proposed question and the relevant instructional content from the PPT presentation, then respond to the subsequent questions...

QUESTION 2:

Which attribute is exclusively associated with scholarly sources?

☒ Peer reviewed before publication.
☐ Includes source citations.
☐ Author and publication date clearly noted.
☐ Professionally edited.

Here is the relevant content from the preceding instructional presentation:

Scholarly Sources

The most credible sources come from **peer-reviewed academic publications**, where experts review each others' work prior to publication.

- Typically, only scholarly journals and books use peer reviews prior to publication.

These types of publications can be found in several ways:

- Google Scholar is a search engine just for academic sources.
- Most university libraries have online portals for their students.
- If you have a specific journal or book publisher in mind, go directly to the publisher's website to search for specific topics and authors.

Search results from Google Scholar (and many university portals) show the number of times that document has been cited by other researchers.

Please rate the degree of correlation between this question and the relevant instructional content.

The stated **instructional outcome** for this question is: *Know how to search for scholarly sources*. Please rate the degree to which this combination of instruction and question corresponds to this outcome:

In your opinion, should this question and/or the relevant instructional content be changed?

Please describe your recommended changes (if any) to the test question and/or the instructional content for the question shown above.

→

Powered by Qualtrics

APPENDIX D

RESPONSES FROM SUBJECT MATTER EXPERTS

Substantial feedback from reviewers. Yellow highlighting indicates feedback about instructional content, not assessment questions. Question order differs from Appendix E.

Q1 Feedback:

Your question states that we are looking for credible sources, but you haven't shared how your learners are determining what constitutes a credible resource. The instructional content needs to have this somewhere and, at this point, we can only assume this was covered. The actual question was what can you determine about the source. The instructional content correlated well to date, time, volume, title.... if the preceding statement of "you are looking for credible sources" is removed. Understanding APA style doesn't equate to a credible source.

Remove credible statement. The question and the instructional content deal with correct APA format not credibility of the source.

You cannot tell just from a citation that the media is a creditable source. It is suggested that there is an explanation and physical examples of creditable and non-creditable sources look like and the type of structure or information we would see. Then also with the content provide some sort of knowledge check scenarios based on the content to "test" to see if the learner knows what to look for and be able to classify materials.

Credible source goes further than the citation. Citations and information included in the source can be varied and this instruction doesn't get into that level of detail. An appropriate validation for this question would be what type of citation or is APA citation correct

I would remove the first sentence from the question. The instructional outcome asks the user to recognize a credible and non-credible source but the relevant content and the question refers to the structure of the APA reference.

Q2 Feedback:

Regarding, "Google Scholar is a search engine just for academic sources," the term "Academic sources" seems to be a combination of two distinct terms that you used. Do you mean "scholarly sources" or "academic publications?" Also, be aware that Google Scholar also contains non-peer-reviewed content; e.g., white papers published in ERIC.

For this particular slide, you have two additional problems related to language:

1) Is the focus on SEARCHING for scholarly sources? If so, then revise to focus on searching. If it's on IDENTIFYING scholarly sources, then revise accordingly.

2) The slide content is vague and qualified with words like ""most"" and typically,"" while the question is definitive and uses the word ""exclusively."" These need to match.

Some of your readers may struggle with your statement "Most credible resources...." and "only peer-reviewed journals....." The "most" and "only" may cause some to question the understanding of the exclusivity.

The suggestions for places to search are very basic but how does a user know what they would be looking for when they do end up searching.

If the goal is to know how to find, knowing that they are peer reviewed is only part of the story. Where would you go to look for scholarly articles, journals, book. That would be google scholar, university portal, etc.

Q3 Feedback:

For library searches, most (rather than all) results are presented in scholarly language.

Searches from library search engines can sometimes include articles from popular publications.

For example, NIU's library has a search engine for the Chicago Tribune.

Q4 Feedback:

Your content started after the search and didn't talk through workflow to get to the point of evaluation.

Your content addresses understanding the intent and purpose of the article. Your question is about refining the initial search.

Instructional content should be changed.

Q5 Feedback:

"The attributes of a scholarly source include the following:" Change to "The attributes of a scholarly source include all of the following:"

Q6 Feedback:

To drive home learner connection to the content, relate how they refine the search, looking at sources found, not scanning and finding other search terms.

Q7 Feedback:

Your outcome is about the process of identifying differences. Your instruction is about what to do with them. Perhaps the discussion should be about expecting differences and counting the sources as equal rather than comparing all the new against the first find.

Something needs to be adjusted. There is nothing about prior research, it could have been same time when you found the opposing views. If it is about the author addressing the differing views then the question should be reworded.

Q8 Feedback:

Same comments as provided previously for this PowerPoint slide. Additionally, "Have been professionally edited" probably isn't the most appropriate phrase. Professional editing only implies that a professional editor has proofread the article for grammar, spelling, etc. Maybe you mean "professionally vetted?"

Q9 Feedback:

If this is the only question on copyright, you're focusing on a tangent. Important to research is accurate crediting of the source, and quoting correctly or paraphrasing. If I find a source from Disney, it's unlikely the author would be mentioned for this to be an issue. If you are helping them learn their own work would be protected, that's different. Better to help them hone in on proper use of copy written material.

Q10 Feedback:

I'm not sure how recognizing the intended audience matters. If I write a piece for ELearning practitioners vs SMEs, I would be more specific than general. I'm not sure how this influences credibility of the piece.

Q11 Feedback:

The intro phrase "You may have noticed that" is unnecessary. This is important and should be stated factually beginning with Violating or better still, switching around so the subject is Using someone's work without acknowledging it is plagiarism and a violation of copyright.

Q12 Feedback:

Why are you asking, “closest to?” It is APA. Which citation is formatted in APA style?

Q13 Feedback:

I would give an example of in text (author, year) in your instruction so someone who is unfamiliar with this has a visual to connect with it since your goal is use and formatting.

Use was discussed, but formatting was not.

Q14 Feedback:

Your questions have all been about APA and your examples have included MLA and others. Is the goal for them to know all or is/isn't APA? Should your outcome change to include format in APA?

Q15 (Removed) Feedback:

The instructional content doesn't provide any information about how the author names should be formatted; i.e., last name, or both first and last name.

"Your slide is titled Reference Page: Page Title. The paragraph about matching citations and references doesn't belong here since it's off topic. This belongs on a page about in-text citations.

The question doesn't need the info about author last name and page in the answers. If you do include it, you need to add year. It's not asking about that though, but looking to see if they know it must have a matching reference. Better to ask about including as matching reference as a footnote, reference page, end note or something about superscripts. "

Q16 (Removed) Feedback:

None of the response options for the question include the hanging indent that is referred to in the instructional content. This reiterates my comment before whether the goal is to know APA or all styles. Why ask is closest to? Which of these is formatted in APA style would be better or Select the APA reference.

Only APA was covered though the instructional outcome did not restrict to APA

Regarding the question, is there a reason the question states "Which option is closest to APA style?" I suggest rewording the question to "Which option represents an APA style reference citation?" The question seems a lot more confident and clear.

APPENDIX E

FINAL ASSESSMENT QUESTIONS

For all multiple-choice items, the correct answer is shown first in this Appendix.

Question 1: *A good research question will inquire about a subject (the phenomenon being studied) as well as:*

- A: some specific attribute of that subject.
- B: the antithesis of that subject.
- C: many other possible subjects.
- D: All of the above.

Question 2: *Select the option for which both statements are true:*

- A: Google's searches will provide more results than a library search.
Library searches can be restricted to peer-reviewed sources.
- B: Google's searches will provide more results than a library search.
Library searches are based on full sentences, not search terms.
- C: Google's searches will always be relevant to your search topic.
Library searches will display only available full-text sources.
- D: Google's search results are not affected by paid advertisers.
Library searches can be restricted to available full-text sources.

Question 3: *You're writing an academic report. Which option will produce the best search results?*

- A: Use the university library's online portal to find peer-reviewed full-text sources.
Based on your results from the first search, modify your search terms and search again.
- B: Use the university library's online portal to find peer-reviewed full-text sources.
Determine which source is the most highly cited, then include it in your research.
- C: Use a Google search to find websites, then view the first three results.
Determine which source is the most recent, then include it in your research.
- D: Use a Google search to find websites, then view the first three results.
Determine which source is the most highly cited, then include it in your research.

Question 4: Which attribute is exclusively associated with scholarly sources?

- A: Peer reviewed before publication.
- B: Includes source citations.
- C: Author and publication date clearly noted.
- D: Highly illustrated.

Question 5: You've made a first-round search using the phrase "cognitive processes," then skimmed two of the sources you found. Identify the best follow-up step(s).

- A: Decide if your first search produced useful results, then run a modified search to look for more sources.
- B: Decide which of your two top choices to cite in your report.
- C: Use both sources in your report, with proper citations.
- D: Reverse the order of your keywords and search again.

Question 6: Which condition indicates that you should revise and rerun your search for sources?

- A: Most of your results were off-topic.
- B: Your results included many credible sources.
- C: Your top results were scholarly sources.
- D: Your top results were all published within the last ten years.

Question 7: Identify the search result(s) most suitable for academic research into "best practices for online course design."

- A: Peer-reviewed journal articles with empirical test results for online classes.

- B: A press release describing a new model of laptop designed for student use.
- C: Articles from the Washington Post about the effectiveness of online classes.
- D: TEDx presentations on “effective online classes,” posted on YouTube.

Question 8: Which option is a scholarly source?

Reeves, T. D., & Marbach-Ad, G. (2016). Contemporary test validity in theory and practice: A primer for discipline-based education researchers. *CBE—Life Sciences Education*, 15, 1-9.

Harris, Rob, and Andrew C. Revkin. “Clinton on Climate Change.” *The New York Times*, 17 May 2007, www.nytimes.com/video/world/americas/1194817109438/clinton-on-climate-change.html. Accessed 29 July 2016.

An Inconvenient Truth. Directed by Davis Guggenheim, Paramount, 2006.

United States Department of Housing and Urban Development. (2008). *Indiana income limits* [Data file]. Retrieved from https://www.huduser.org/Datasets/IL/IL08/in_fy2008.pdf

Question 9: When evaluating the usefulness of a source document, what should you consider?

- A: The author’s intended audience
- B: The author’s alma mater
- C: The number of books published by the author
- D: If the source provides a Creative Commons license

Question 10: You have located four different sources with valuable information, but some of the sources disagree about key details. How should you handle this in your research report?

- A: Create a dialog between sources by discussing their differences.

- B: Check the dates they were written and use only the newest source.
- C: Only use the source that is the most highly cited.
- D: Choose the source that appears earlier in the search results.

Question 11: *U.S. copyright law makes sure that authors maintain control over the work that they create. Select one exception to this law:*

- A: “Work for hire,” when the author is on a company’s payroll.
- B: The time between creation and publication is unprotected.
- C: Copyright protection does not apply to digital photographs.
- D: Most new intellectual works are in the public domain.

Question 12: *Why is it necessary to provide an in-text citation within your research report?*

- A: To indicate words or specific concepts that were developed by other authors.
- B: As a way to claim credit for information that you have located.
- C: To link to the source (website or publication) of a document.
- D: To link to the source for a document and show date of latest revision.

Question 13: *You want to cite an academic journal article in your research report. Which in-text citation follows APA style?*

- A: In this case, “adding on-screen text does not create split attention” (Mayer & Fiorella, 2014, p. 299).
- B: In this case, “adding on-screen text does not create split attention” (Mayer & Fiorella, *Principles for extraneous processing*).
- C: In this case, “adding on-screen text does not create split attention” (Richard Mayer, 2014, p. 299).

D: In this case, “adding on-screen text does not create split attention” (Richard Mayer, “Principles for Extraneous Processing,” 2014).

Question 14: Which Reference list entry has been formatted in APA style?

Susen, S. (2011). In the public sphere. *Sociology Journal* 5(1), 156-163.

Susen, Simon. “In the Public Sphere.” *Sociology Journal* 5.1 (2011, Spring).

Susen, Simon. (2011). In the Public Sphere. *Sociology Journal* 5 (1): pp. 156-163.

Susen, S. (Spring 2011.) *In the Public Sphere*. *Sociology Journal* 5 (1): 156-163.

Question 15: Looking at the citation shown, what can you determine about this source?

Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of instructional design. *Human Factors*, 40, 1–17.

A: This is a journal article titled “Levels of instructional design.”

B: This was published in a journal titled “Levels of instructional design.”

C: This was published on a website titled *Human Factors*.

D: This citation does not use proper APA style.

APPENDIX F

TREATMENT ADMINISTRATION METHOD IN BLACKBOARD

Blackboard is a learning management system (LMS) that can also be used as the apparatus to conduct data collection for qualitative pretest/posttest experiments using random assignment of participant treatment conditions to n equal groups. The experiment described below takes the form of a multiple-choice instructed response assessment.

Method

Prepare the treatments. This step has options: instructional content can be constructed through Blackboard's available tools or the optional creation of instructional treatments (e.g., audio-narrated PowerPoint presentations) using software such as iSpring, Captivate, or others to export a SCORM package that can be uploaded to Blackboard.

Within a Blackboard module, each treatment must be hosted as a separate item.

Content

webcourses.niu.edu/ultra/courses/_296597_1/cl/outline

Create Item
A content item is any type of file, text, image, or link that appears to users in a Content Area, Learning Module, Lesson Plan, or content folder.
[More Help](#)

* Indicates a required field.

CONTENT INFORMATION

* Name
Treatment A

Color of Name
Black

Text
For the toolbar, press ALT+F10 (PC) or ALT+FN+F10 (Mac).

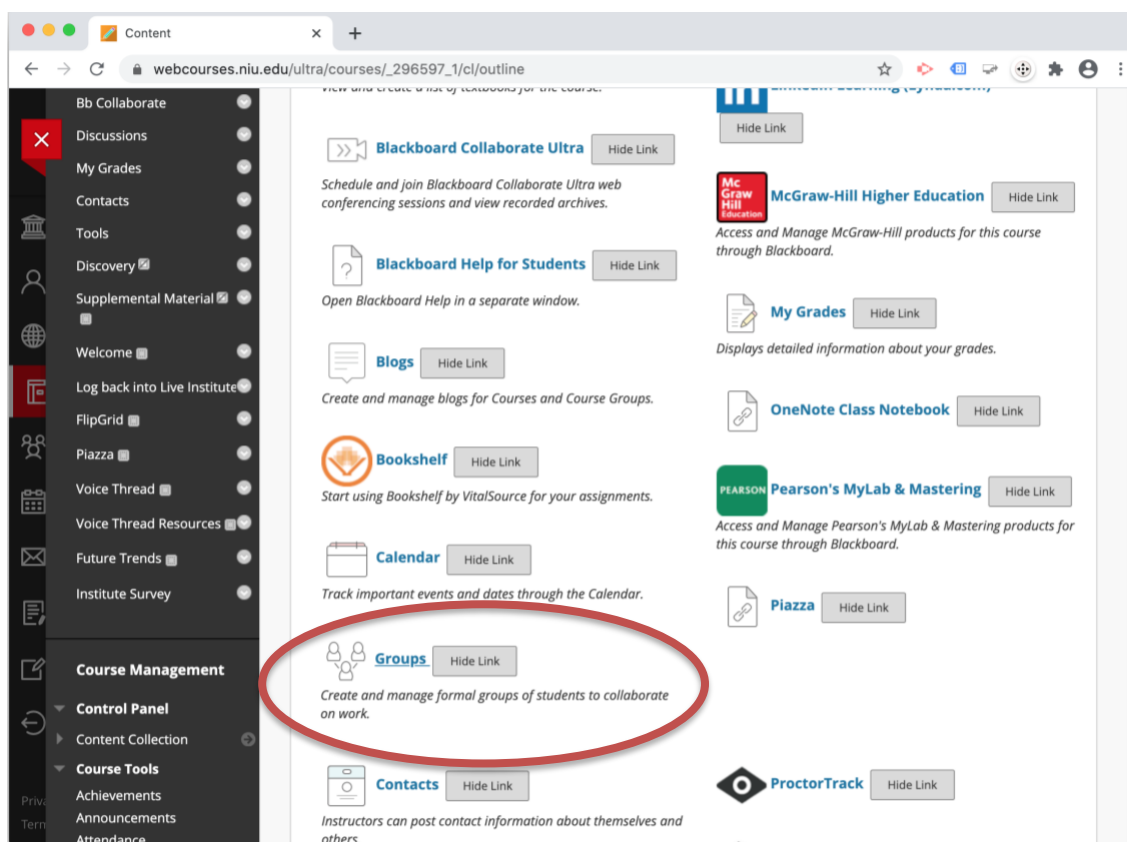
Students will be randomly selected for one treatment group

Path: p Words: 9

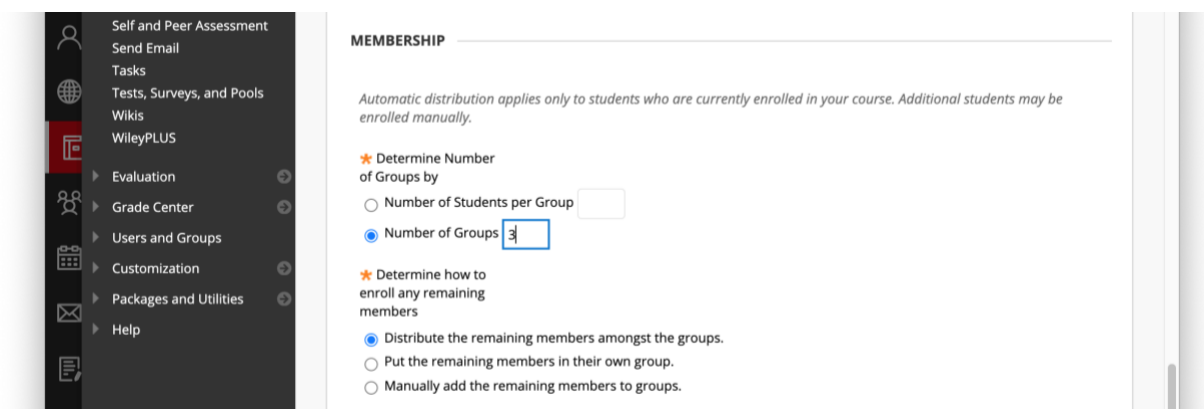
Click **Submit** to proceed.

Cancel Submit

From Blackboard's interface, choose Tools then enter Groups.

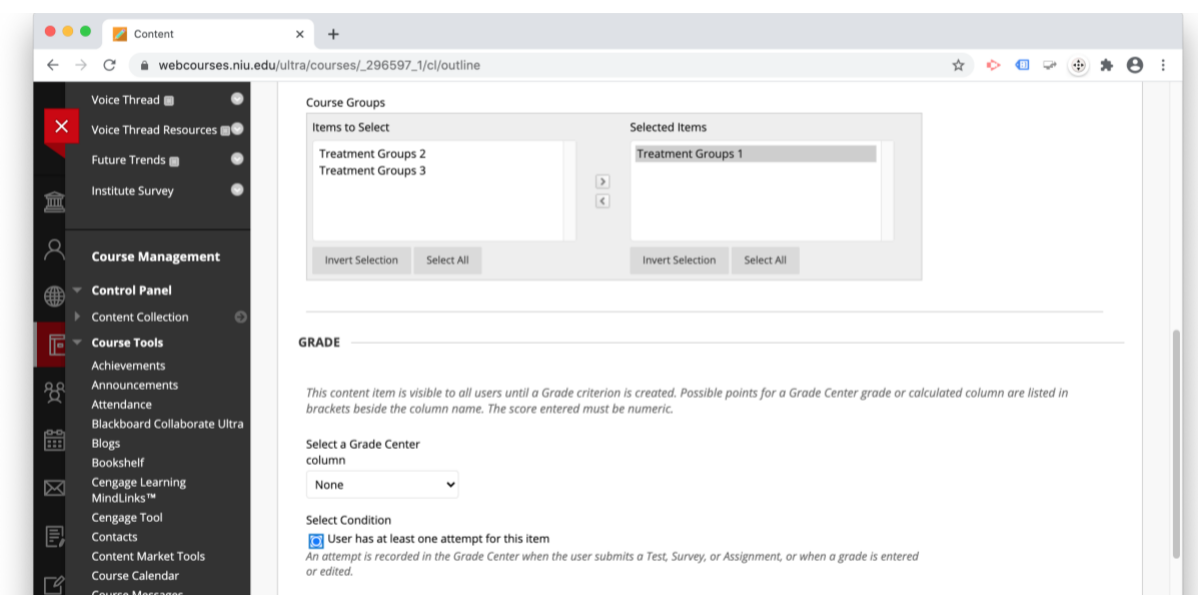


Choose Group Sets > Create > Random Enroll. Give the set a name and set student visibility to “no” until the entire experiment is ready. Uncheck all the Tools and uncheck Personalization. Set the desired number of groups, choose “Distribute the remaining members among the groups,” then click Submit.



Upload or create each treatment as separate assignments (but all within the same module) with each name extremely similar but just different enough to be able to distinguish the groups. (The name of the randomly assigned treatment name will be visible to each learner if they open the Blackboard module's Table of Contents.)

In Blackboard, edit the Adaptive Release option for each treatment. In the Membership section, select the appropriate treatment by moving the desired treatment into the right-hand window. Select the correct Grade Center column and score value, select condition of "User has at least one attempt," then click Submit.



In addition to collecting the experimental data, it is also necessary to set up the Grade Center (original view Blackboard) or Gradebook (ULTRA view) so that students' efforts will be graded. Go to the Grade Center/Gradebook. You will find that as each treatment was uploaded to Blackboard, it created its own column in the Grade Center/Gradebook. Since each column is worth the same number of points, yet you need to prevent students from getting more than that

extra number of points, you can't just let them all become part of the total; instead, create a new Total Column to show the mean of all the treatment grades.

Leave the visibility on for the new mean column and add it to the Grade Center/Gradebook total grade calculations; hide the individual experiment columns from students' view.

Implementation

When the assignments become available to students, the collection of experimental data begins. You might have all students complete the assignment synchronously within a short window (e.g., from noon until 2 pm) but you may also leave the response period open for multiple days. After the assessment period is over, download the results from Blackboard. In the ULTRA view, this is done by clicking the Download icon in the top-left corner of the Blackboard interface.

Using an externally created experiment may require adjustment of the communication protocols between the SCORM and Blackboard. Each variation of SCORM protocol provides a slight change in data collected as well as its own potential for protocol issues with Blackboard. Version SCORM 2004 v2 protocol was determined to be the most compatible with both Blackboard ULTRA and Classic view.

If this course is repeated in a later semester, the Groups will not automatically enroll users. Once all users have enrolled in the course, it will be necessary to "Edit Group Set Membership" then click the Randomize Enrollments button to push the students into the groups. (You only need to click the button once to fill all the groups simultaneously.) Also note that the default setting for this process creates equal-size groups, so the last few students will not be

assigned to a group – you’ll either have to add them manually (by editing the Group Set Membership) or change the group settings to allow unequal group sizes.