The Effects of Online Course Pedagogies on Learning and Cognitive Workload

Kallan Kay Christensen
San Jose State University

THE EFFECTS OF ONLINE COURSE PEDAGOGIES ON LEARNING AND
COGNITIVE WORKLOAD

A Thesis

Presented to

The Faculty of the Department of Psychology

San José State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

by

Kallan K. Christensen

May 2016

The Designated Thesis Committee Approves the Thesis Titled

THE EFFECTS OF ONLINE COURSE PEDAGOGIES ON LEARNING AND
COGNITIVE WORKLOAD

by

Kallan K. Christensen

APPROVED FOR THE DEPARTMENT OF PSYCHOLOGY

SAN JOSÉ STATE UNIVERSITY

May 2016

Dr. Ron Rogers   Department of Psychology

Dr. Sean Laraway   Department of Psychology

Dr. David Schuster   Department of Psychology

ABSTRACT

THE EFFECTS OF ONLINE COURSE PEDAGOGIES ON LEARNING AND
COGNITIVE WORKLOAD

by Kallan K. Christensen

The use of recorded lecture video is a common online practice, although the
implementation of this pedagogy varies and may have differential impacts on learning
depending on the extent to which it increases cognitive workload. This study compared
two different online statistics lecture formats: interactive and non-interactive. The
interactive lecture video stopped approximately every minute and the participant had to
answer a question regarding the material before moving on to the next section of the
video. The non-interactive lecture video was the exact same video edited to remove all
interactive breaks in the video. Cognitive workload was measured using the NASA Task
Load Index and learning was assessed using performance on a quiz. Seventy participants
from the San Jose State research pool were included in the analysis. The researcher
predicted that cognitive workload would be a mediating variable in the relationship
between online pedagogy and learning, such that an interactive pedagogy would be
associated with lower cognitive workload and result in larger learning gains compared to
a non-interactive pedagogy. The researcher concluded that cognitive workload was not a
mediating variable for online pedagogy and learning. Online pedagogy did not predict
learning. However, the type of pedagogy had an effect on cognitive workload, such that
the non-interactive group had lower cognitive workload scores. Developers of online
pedagogy should recognize that cognitive workload differs across pedagogies and the
impact on learning should remain the final assessment of any given approach.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Introduction

Many universities in the United States have rapidly increasing numbers of students enrolling in classes each year. In California, this has caused many schools to have an impacted status (California State University, 2014). Other states have requirements that students have to fulfill before choosing a major or have caps on the number of people who can be accepted into a major. There is not enough room in classrooms to accommodate the new demand for higher education. Universities offer online courses as one solution to this problem. Online courses are able to accommodate more students, as seats are not needed in a classroom (Bowen, Chingos, Lack, & Nygren, 2014).

In addition, more people are going back to school to pursue higher education for job purposes because there has been a shift in educational needs for upper level job positions. For many jobs, graduate level degrees are required in order for an employee to get promoted. This creates a need for remote education. Students who have jobs and want to go back to school are not as capable of getting to a physical campus to complete courses (Jenkins & Downs, 2003). Online education allows for flexible schedules (Leibold & Voelpel, 2013). This gives students the flexibility to complete courses since classes are not always offered at convenient times for commuter or working students. Online education allows students to avoid long commutes to come to a physical campus everyday (Jenkins & Downs, 2003). Online courses can fulfill a clear need and are also practical and convenient. Research needs to be conducted to investigate techniques and methods to ensure that online education is as effective as possible.

Teaching Online Courses

There have been several approaches to teaching online courses. The approaches vary in the level of teacher interaction, the use of video, and the types of video lecture given (Arbaugh, 2014). For the purposes of the present study, only the literature reviewing online lectures with video was reviewed. The level of teacher interaction refers to how much interaction is inherent in the method and describes the relationship between students and teachers. Some instructors send daily emails to students while others only answer emails if a student deliberately seeks the instructor's help. Arbaugh (2014) found that teaching presence was a strong predictor of a student's acceptance of the course medium. The researchers defined student acceptance as the caliber of the class and how involved the student was likely to be in the course. Involvement in the online course refers to how much a student participates in discussions, class assignments, and other activities related to the course. Students who reported a high acceptance of the online course also gave higher ratings of perceived learning. This finding suggests that while designing online courses, attention to the student-instructor interaction should follow.

The type of lecture video made by the instructor to present to students who take online courses can also differ. There are three common types of video lecture (Bowen et al., 2014). The first is lecture capture. Lecture capture is simply a video recording of a lecture that took place in classroom on a college campus and is put on the web for students to view. The second is picture-in-picture. Picture in picture refers to videos that show course material in a big window and have a smaller window that shows the

instructor as he or she talks. The last type of lecture is voice over. Voice over shows the

instructor writing and has the instructor's voice as audio for the recording of the writing.

These distinctions in media type are important to keep in mind while reviewing the online

course literature, as results from one form of media may not generalize to all forms of

online course media.

Chen and Wu (2015) found learning performance with lecture capture online

lectures and picture-in picture online lectures are more effective than voice-over type

online lectures. However, the researchers speculated that individual learning preference

might have influenced the result. Chen and Wu (2015) used a Visualizer-Verbalizer

hypothesis to help explain the finding that voice-over type lectures were the least

effective learning method. The Visualizer-Verbalizer hypothesis states that some

students learn better from seeing information while other students may learn better from

hearing information. However, when the two are combined, better results are seen. Chen

and Wu (2015) applied working memory models to explain the findings. Working

memory helps people learn and remember information. However, working memory is

limited. Working memory is composed of three parts: central executive, phonological

loop, and the visuospatial scratchpad (Baddeley & Hitch, 1974). The central executive

processes information stored in the other two components. When seeing information and

hearing information, both the phonological loop and then visuospatial scratchpad are used

rather than one or the other. The Visualizer- Verbalizer hypothesis states that providing

both the phonological loop and visuospatial sketchpad with information is better than

providing just one with information. There is a greater chance the central executive will process the information and the information will then be retained.

Chen and Wu (2015) also implied that the cognitive load theory of multimedia learning could also help to explain why voice-over type courses were the least effective methods. Lecture capture and picture-in-picture lecture types, using auditory and visual aspects, are better for working memory because they increase the chance that information will be identified by the student as important. If a stimulus is determined as important, the student can employ the necessary tools for attention and encoding to store the information in long-term memory. Voice-over presentations do not employ these components. However, it should be noted that voice over presentations could increase sustained attention and increase cognitive workload because cognitive resources in both the visual and auditory loops are needed to understand the material presented in this form (Chen & Wu, 2015). Sustained attention can help a participant to pay attention to a stimulus that presents information for a longer period of time, therefore increasing the likelihood of the information getting into long-term memory. However, sustained attention is effortful and effort increases as people spend more time on a task. If cognitive workload starts to become too high, increasing cognitive workload may not benefit learning and memory.

Book learning and online learning have also been compared. It was found that lecture capture might be better than books for complex subjects (Chen & Wu, 2015). While learning complex subjects, instructors break apart complex material into smaller parts so the concepts are not too difficult for the students to learn. Complex subjects

learned from books may require too many working memory resources resulting in high cognitive workload, which is associated with poor learning performance. The smaller parts in an online course reduce cognitive workload thus, employing fewer working memory resources. Online courses offer flexibility to the students and teachers so optimal learning conditions can be achieved. Students can move at their own pace in an online course and may benefit because the student can work at a speed that is not cognitively taxing for them.

The perceived environment in which student learning occurs may also influence cognitive workload. Emerson and MacKay (2011) were interested in investigating how students respond to learning in different environments. Emerson and MacKay used pre and post lesson questionnaires as well as the NASA-Task Load Index (Hart & Staveland, 1988) measure of subjective cognitive workload to compare cognitive workload to test scores. Emerson and MacKay (2011) also suggested that there might be a correlation between higher cognitive workload and lower test performance. Two online courses that resembled different environments, classroom or online, were compared; paper-lesson mode and web-based mode. The paper-lesson mode was meant to resemble learning in a physical classroom. The paper-lesson mode had a video of an instructor writing information on a piece of paper while talking. The web-based mode course had a videotape of an instructor giving a lecture in front of a classroom. The researchers concluded that the cognitive workload scores for the paper lesson mode and the web-based mode were negatively correlated with learning performance. Emerson and MacKay (2011) suggested that cognitive workload might increase if a student views an

online course as being harder than classroom based courses, or reversely, cognitive workload may decrease because the student views online courses as generally being easier than classroom based courses.

Types of Online Courses

Although there are many types of online courses, the Massively Open Online Course (MOOC) has taken prominence in the research literature. MOOCs are a type of online course that offer interactive aspects, such as discussions with participation from the teacher and student, and are self-paced. Each MOOC can be modified by the instructor to best fit the material. For this reason, the present study focused on common types of MOOCs to compare learning outcomes and identify characteristics that may help students effectively learn material. A brief overview on the history of MOOCs and the need for MOOCs will be discussed next to highlight the importance of MOOCs and why the previous literature has found mixed results regarding the effectiveness and usefulness of MOOCs.

MOOCs offer both free education for those who are curious about a subject as well as college accredited courses that students may be able to take for course credit. There are many different types of MOOCs, but research on the types of MOOCs has produced mixed results. In the past, learning outcomes have not been great for any of the types of MOOCs (Alraimi, Zo, & Ciganek, 2015). However, most of the factors that lead to the mixed results are confounded due to the intentions students have for taking the course and the number of students who actually complete the course (Milligan &

Buckenmeyer, 2008). The reasons for why some MOOCs work and others do not have yet to be identified.

Currently, MOOCs are divided into two categories (Alraimi et al., 2015). MOOCS are divided into categories to distinguish between professor-centric models and models that are set up for collaboration among others to learn and solve problems. One type is professor-centric and is referred to as an xMOOC. The xMOOCs are similar to traditional courses and contain content chosen by the professor. The xMOOC usually consists of a combination of mostly video lectures and computer-marked assignments, in combination with a few discussions and peer assignments. The other type of MOOC is the connectivity or cMOOC that has structure but also has more student involvement. The cMOOCs are more likely to incorporate social media and communication within the class. Although there are two distinct types of MOOCs, curiosity and job advancement are the main reasons why people decide to enroll in MOOCs (Alraimi et al., 2015).

MOOCs have not been extensively researched regarding the learner or instructional design. Despite the gap in the research, it should be noted that MOOCs provide a complete package for learners (Margaryan, Bianco, & Littlejohn, 2015). The MOOC formats often follow basic principles of instruction: problem-centered, activation, demonstration, application, integration, collective knowledge, collaboration, differentiation, authentic resources, and feedback. Of the principles of learning, the one that may be the most difficult for online courses to achieve is the feedback aspect. Automated learning systems cannot answer student questions in the same way traditional classroom instructors can (Margaryan et al., 2015). The need for human feedback is also

crucial in online courses. Students are more likely to respond to this type of feedback than automatically generated feedback (Margaryan et al., 2015).

Learning

As previously discussed, there is a clear need for online education, however, the most effective format for online learning has yet to be established. Many different learning models can be applied to online course formats and offer insight into how students most effectively learn. Chi and Wylie (2014) proposed a model for learning called the ICAP: interaction, connectivity, active, and passive. In the ICAP model there is a hierarchy for behaviors that predict learning. The four categories of behaviors, listed from most to least engaging, are interactive, constructive, active, and passive. The ICAP model assumes that the more engaged students are, the better learning outcomes they will have. Constructive behavior refers to students who generate inferences and ideas that go beyond the material provided by instructors. Interactive refers to students who participate in interpersonal activities. Active behavior refers to students who incorporate a physical action while learning. In contrast to active behaviors, passive behavior refers to students who take actions that minimally deepen their understanding of the material and who listen to lectures.

Others have also proposed that active learning results in better outcomes than passive learning. Abeyeskera and Dawson (2015) analyzed the flipped classroom approach and why research should be geared towards active learning styles. In addition, the authors stated that active learning might also help to manage cognitive load. Active learning can help students to pace their learning, which would help to manage cognitive

load because students would be asked to stop and process information before moving on to another section. Although research has not been conducted on the topic, the authors proposed that throughout an online course, advance learners can fast forward through material they already know and students who are struggling can re-watch the video. The student actively manipulating the pace of the video could have positive effects for reducing workload and increasing learning.

The learning process is complicated because there are many factors that can influence the outcomes. There are also multiple methods used to measure learning. Cognitive Workload and Engagement are two constructs relevant to the learning process (Johnson et al., 2011). Cognitive workload can also be used as a measurement that indicates a better chance that learning will occur. Appropriate cognitive workload in a task can increase the likelihood material is learned (Johnson et al., 2011).

Cognitive Workload

Cognitive workload theory describes the relationship between the amount of mental effort people use and the limits to working memory (Sweller, 1988). Baddeley and Hitch (1974) proposed a model for working memory. Working memory is a commonly used model to describe short-term memory and how short-term memory works. Working memory holds and processes new information and is an important process in learning and understanding. However, there are storage limits to working memory. Cognitive workload helps to describe how the limits in working memory are reached, and how people can avoid reaching these limits when learning and processing new information.

Originally, cognitive workload theory was applied to problem solving, but has since been applied to learning in general. New information, steps to complete a task or to learn something new, must stay in working memory until the information is moved to long-term memory. If information is not moved to long-term memory, the information is forgotten and therefore, is not learned. Working memory keeps the information active while the brain processes the information into long-term memory. If there is too much information to process, cognitive overload may occur and learning will not be possible (Sweller, 1988).

Cognitive Load Theory breaks down cognitive load into four different categories: intrinsic, extrinsic, extraneous, and germane. Intrinsic load is inherent to the particular topic being learned. Thus, if the topic is more complex, intrinsic load will be high. Intrinsic load can also be referred to as task load. Extrinsic load refers to the manner in which information is given to learners. Extrinsic load is also referred to as instructional load. The teacher or instructional designers control extrinsic load. It is important that intrinsic and extrinsic load stay within a learners working memory limits.

Extraneous load is cognitive load that is imposed from the information or learning activities. Extraneous load hinders learning. In contrast, germane load is load imposed from the information or learning activities that foster learning. For learning to be maximized, the sum of intrinsic, extrinsic, extraneous, and germane load cannot exceed the memory resources that are available. Cognitive load is additive, therefore, there has to be a balance between each load type. It is also important to remember that in a classroom setting, as working memory is processing information, new information is

simultaneously given to learners. Thus, the new information acts as extraneous load. The increase in extraneous load can be detrimental to learning.

Wong, Leahy, Marcus, and Sweller (2012) suggested that presenting potentially transient information in short segments could help to capture a learners attention and reduce extraneous workload. Wong et al. (2012) used mixes of animations and speech in videos to help children learn origami. It was found that when the information was presented in short segments, advantages of presenting information in both animation and speech forms were seen. Thus, the natural advantages of each learning technique were able to manifest themselves. However, when learners viewed information in longer segments, the animations did not seem to help learners as much as in the short segments.

Segmenting information may be a way to reduce extraneous load. Lowering extraneous load may be necessary to prevent cognitive overload by freeing up capacity for other load dimensions and cognitive processes relevant to learning. Segmenting information may also help with increasing germane load; increasing germane load is beneficial to learning. Segmenting may help to engage learners, providing an artificial way to induce germane load; higher levels of interacting would result in higher amounts of germane load.

Paas, van Gerven, and Wouters (2007) found that instructional efficiency was increased in interactive teaching methods when compared to non-interactive teaching methods. There was an increase in cognitive load for the non-interactive learners. However, it was noted that replication studies are needed and the material being learned will have an effect on instructional designs.

Current Study

Online education is an alternative form of education. Educational pedagogies for online courses have been examined, however, there are questions about the most effective delivery of online content. Cognitive workload is a measure that can help predict learning outcomes. Therefore, the current study aimed to examine how cognitive workload is affected by online course pedagogy. Based on previous findings, an online course that is interactive and that contains a component that makes the student participate in active behavior should result in better learning outcomes than a course that does not contain these components. The component that promoted active behavior in the present study was a question that the participants were required to answer before moving on to the next segment of the course. The motion of the participant moving the mouse makes the learning situation active rather than passive. The learner is using different cognitive resources to complete the motor action, and it may also give the participant more time to process information that had just been presented to him or her.

The researcher focused on the MOOC format produced by Udacity. Udacity created an online course format in which students watch videos that include video and audio components. Udacity is different from other online course formats because it shows the hand of the teacher writing out notes and giving examples as students progress throughout the course. In addition, the course forces students to answer questions throughout the lesson segments. The current study compared two different online lecture formats: interactive and non-interactive. The purpose of the current study was to explore how segmentation, by means of the student answering a question, affects learning. The

course material that was used for the present study was obtained from the *Intro to Descriptive Statistics* course at Udacity.com (Rogers, Laraway, & Kormanik, n.d.). The lecture videos covered measures of central tendency: mean, median, and mode.

In line with previous studies, the current study used a subjective measure for cognitive workload, the NASA Task Load Index (NASA-TLX), to assess if segmentation, an aspect of online course pedagogies, can lead to greater learning outcomes. However, it should be noted that the results from the subjective measures do not always match up with the objective results for measures of cognitive workload. That is, self-report measures do not always provide the same results as measures that do not involve self-reports, such as electroencephalography. Matthews, Reinerman-Jones, Barber, and Abich (2015) compared several subjective and objective workload measures of cognitive workload to investigate how general workload could be defined. Physiological methods seem to explain the results from objective learning outcome measures with fewer inconsistencies than subjective measures. Despite this finding, Mathews et al. (2015) also reported that the NASA-TLX had the strongest effect sizes in their comparison of multiple psycho-physiological workload metrics for sensitivity to task type and dual tasking. The NASA-TLX is also one of the most widely used and extensively validated scales. Furthermore, Mathews et al. (2015) concluded that the different metrics of cognitive workload were sensitive to cognitive workload but it remains unclear if all measures of cognitive workload measure the same cognitive processes. Applied uses of workload metrics need to be interpreted with some caution.

Colleges often assess students learning based on test scores. Many students in online courses take computer graded multiple-choice tests as an assessment of their leaning. To ensure the proposed pedagogies of online learning result in high-test scores, the current study also used a multiple-choice exam to measure learning outcomes. The researcher compared the results from the subjective cognitive workload measures with the participants' test scores to examine the overall effectiveness of the online course pedagogy.

Two conditions were tested in the present study: interactive and non-interactive. It was first hypothesized that participants in the interactive condition would have a significantly higher mean score on the final multiple-choice test than participants in the non-interactive condition. It was also hypothesized that the interactive group would show lower cognitive workload scores than the non-interactive group. Finally, it was hypothesized that cognitive workload would be a mediating variable in the relationship between online course pedagogy and learning outcome.

Method

Institutional review board approval was obtained before the data collection

process began.

Participants

Using convenience sampling, 98 participants from the General Psychology course

research pool at San José State University were recruited for the study. The participants

in the study had to meet the following requirements: (1) must be 18 or older; (2) could

not have taken a college level statistics course. A math ability questionnaire was given to

participants at the end of the study to make sure they met the second requirement

(Appendix A). The participants filled out the questionnaire at the end of the study to

avoid priming effects. If the survey was given to the participants at the beginning of the

study, participants could have been aware of a potential math-related inability to

complete the experiment. Although all 98 participants completed the procedure, 28

participants did not meet the second requirement and were excluded from data analysis.

Thus, 70 participants were included in the analysis.

Approximately 46% of the participants in the experiment were males and 54%

were females. Participants had a mean age of 18.55 and ranged between 18 and 22.

Participants either indicated that their primary language was English or that they

understood English very well. A question was given to participants that did not mark

English as their primary language to ensure they understood English. Participants chose

a number on a scale of one to seven, one indicating "not at all" and seven indicating

"very well," their understanding of English. All participants chose seven, "very well,"

for their understanding of English. This was important because taking a course in a language that a participant did not understand could have influenced workload. The general student population at San José State University identify themselves as Asian (33%), Hispanic (23%), White (20%), Foreign National (12%), African American (3%), or Other (9%) (San José State University, 2015). All participants must have previously demonstrated high school level math proficiency and a readiness for college-level math. San Jose State University determines this through the completion of either two years of high school algebra or passing scores on or exemption for the Entry Level Math (ELM) test, which ensures that students have basics knowledge of mathematic concepts. Although either two years of high school algebra or passing the ELM is required to attend San Jose State University, there were questions regarding this information on the math abilities questionnaire to ensure participants had the math abilities needed to complete the experiment. All participants demonstrated that they had the mathematical skills needed for the experiment. Informed consent was obtained from each participant prior to the start of the study (Appendix B). Participants in this study received partial course credit for their participation. In addition, participants had the option to enter into a raffle for one of four $25 gift cards.

Materials

NASA-Task Load Index (NASA-TLX). Cognitive workload was measured using the NASA-TLX (Hart & Staveland, 1988). Hart and Staveland (1988) conducted an extensive validation study and concluded the six dimensions of the NASA TLX were able to account for a significant percentage of variance: r^2 values ranged from .78 to .90

with an overall rating of $r^2 = .86$. There are three ways to administer the NASA-TLX survey: verbally, digitally, and paper and pen. Using a test/retest reliability rating technique, the researchers also concluded that the correlation between each of the three administrations was .83. It was noted that the verbal method was the least favorable method to use. Participants in the present study completed a digital version of the NASA-TLX to measure subjective cognitive workload.

The NASA-TLX consists of six dimensions: mental demand, physical demand, temporal demand, performance, effort, and frustration (Hart & Staveland, 1988). Descriptions of each of the six dimensions were given to participants to ensure interpretation of each of the dimensions was the same. The description of mental demand was, "How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?" The description of physical demand was, "How much physical activity was required (pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?" The description of temporal demand was, "How much time pressure did you feel due to the rate or pace at which the task or task elements occurred? Was the pace slow and leisurely or rapid and frantic?" The description of performance was, "How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?" The description of effort was, "How hard did you have to work (mentally and physically) to accomplish your level of performance?" The

description of frustration was. " How insecure, discouraged, irritated, stressed, and annoyed versus secure, gratified, content, relaxed, and complacent did you feel during the task?"

Participants first filled out a ratings sheet. The ratings sheet required the participants to rate each of the six dimensions on a scale of zero to one hundred, where zero indicated "low" and one hundred indicated "high." However, the performance dimension was reversed scored where zero indicated "good" and one hundred indicated "poor." In addition to the written description of the labels, the performance dimension was clearly labeled differently in efforts to have the participants notice this question was revered scored. Participants moved a marker across the scale and the participants could see the exact number they were choosing. After filling out the ratings sheet, the participants completed the weighted rating section of the NASA-TLX. In this section, participants were given fifteen pairwise comparisons of the six dimensions. The participant clicked on which of the two dimensions was a more important contributor to their workload. The dimensions that were picked as contributing more to workload were later summed and then used as weights when calculating the overall workload score. Higher scores indicated higher levels of subjective cognitive workload.

Lab. Participants completed the experiment on a desktop computer in a shared lab space at San José State University. The lab space consisted of 30 computers with a keyboard and mouse. Each computer was in an individual cube with sound proofing barriers. Participants were tested in groups of one to nine. If more than one participant was scheduled, there was at least one computer station separating one participant from

the next. All surveys and videos were viewed using the Google Chrome web browser. The videos were linked and streamed through screencast.com. Participants were required to wear a pair of headphones while viewing the videos. Each computer was preset to have the headphone volume at 50%. If participants needed to, they could adjust the volume to a comfortable volume. Each computer station also had a piece of scratch paper, a pen, and a calculator that could be utilized by the participants in a manner they felt appropriate. Some participants wrote notes on the paper, while others did not. The paper and calculator were also provided to participants as a resource to help them solve statistics questions in both the practice and quiz portions of the experiment.

There was also a large table with nine chairs in the lab. The table faced a blank wall in which the experimenter projected an orientation slide presentation. The orientation slide presentation gave an overview of the experiment. The participants were familiarized with the experiment setup, video functions, and the NASA-TLX survey. The experimenter sat at a desk in the room so she was available to answer any questions the participants had. The desk did not face any of the participants' computers.

Orientation Video. All participants viewed the same orientation video, which was designed to serve two purposes. First, the orientation video provided a general overview of what the participants would be learning in the lecture video that followed (i.e., basic statistical concepts like the measures of central tendency). Second, the orientation video served to introduce the participant to the video interface and response system. That is, the video was segmented and required the participants to respond to questions using check boxes or typed-in responses before navigating to the next segment.

19

The questions ensured that the participants could hear, see, and navigate easily through the videos. The participants were also oriented to how the information would be presented to them. The video explained to the participants that they would see various figures and drawings while also hearing the instructor talk. The video was approximately 45 seconds in duration.

      Instructional Videos. The participants viewed one of two lectures: non-interactive or interactive. The non-interactive video played without any interruptions. The interactive video required participants to answer questions as the video played. The participant had to think about an answer and enter the answer into the video as it played. The course was derived from the *Introduction to Descriptive Statistics* course at Udacity.com (Rogers, Laraway, & Kormanik, n.d.). The course was an entry-level college statistics course designed for a general education audience. Mean, median, and mode and how each relates to normal and skewed distributions were extensively covered in the video. In the original Udacity course, there are pauses in the course where students are required to answer a question before progressing to the next video in the course. The video shows a hand overlaid onto any pictures or figures that are drawn during the videos (Appendix C). All of the pictures and figures look hand-drawn, although the drawing of them in the video is sped up to match the audio instruction. A female instructor working from a script in a studio narrated the video. The videos do not show any recordings of actual classroom lectures. Instead, they have been scripted before hand and incorporate many real-life examples.

      For the present study, the original Udacity videos were edited using a video

editing software, Camtasia. Camtasia provided a platform in which questions could be inserted into the videos. The videos were also edited to remove any information pertaining specifically to the Udacity class (i.e. access to spreadsheets and instructions for class assignments). White boxes that made information invisible to participants were also placed over information in the original video that the participants in the experiment would not have the background knowledge to understand. Many symbols introduced in previous lectures appeared in the original videos, but were removed for the purposes of the experiment. The Camtasia platform also allowed participants to enter an assigned user name and email address. The assigned user name allowed the experimenter to track participant information. This kept the participants' identities confidential and allowed the ability to check if participants had watched all the video content.

The non-interactive video was edited so there were no breaks in the video. The participants who viewed this video did not have to click on the screen or answer questions. The video was approximately 21 minutes in duration. At the end of the video, the participant was required to complete a set of fifteen practice questions. If the participant answered the question wrong, a brief description as to why the correct answer was correct was given in text format.

The interactive course was the exact same lecture recording as the non-interactive course. However, the video stopped approximately every minute and the participant had to answer a practice question before moving on to the next section of the video (Appendix D). This lecture resembled a traditional Udacity course, but as previously stated, the Udacity platform was not used. The practice questions that interrupted the

video were similar to the questions that the non-interactive group received at the end of the video.

Procedure

This study used a between subjects design. This study required one 90-minute session. The first step was to assign participants the condition that would be presented to them during the study using quasi-random assignment. This was done prior to the participant arriving to the study. Odd numbers corresponded with the non-interactive procedure and even numbers corresponded with the interactive procedure.

When the participant arrived at the lab, he or she was given consent information and asked to electronically sign consent forms. The consent process took approximately five minutes. Then, participants listened to an orientation presentation given by the experimenter. The experimenter read through a script while showing the participants slides with various bullet points and screenshots on them. Participants were oriented to different functions of the videos and the overall procedure of the experiment.

In the orientation presentation, the participants were also told that they would have the opportunity to win one of four $25 gift cards if they score above a 70% on the final test over the statistics concepts they learned in the experiment. The gift cards were offered as extra incentive so the participants were motivated and would do their best to learn the concepts in the videos. In a normal online class, student motivation is driven, at least in part, by an incentive to do well and pass the class. The gift cards took the place of grades as an incentive. However, every participant was entered into the raffle

regardless of performance. The participants were told at the end of the study that everyone had an equal chance to win a gift card.

After the orientation presentation, the participant was assigned to a computer in the lab. A piece of scratch paper, pen, calculator, and extra procedural instructions were at each of the computer stations for the participant to use as he or she wished. The participant completed the rest of the procedure on the computer. An online survey system (Qualtrics) was used to instruct participants through the various parts of the experiment. The experimenter was available to help participants with technical difficulties. The experimenter was not able to answer questions regarding the content of the statistics videos.

The participants watched an orientation video to introduce him or her to the stimulus. The orientation video briefly described the features of the online course format and what was expected of the participant. Immediately following the orientation video, the participant filled out the NASA-TLX survey in regards to cognitive workload experienced while watching the orientation video. This allowed the participant to be exposed to the NASA-TLX before the experimental portion of the procedure.

The participant then completed either the non-interactive procedure or the interactive procedure depending on which group he or she was assigned. The non-interactive procedure started with the participant viewing a lecture video that was approximately 21-minutes in length. The participant then completed another NASA-TLX survey. Next, the participant was instructed to complete practice math problems related to the concepts the participant learned about in the lecture video. The participant

then received feedback on the questions. The participant could see which questions were answered correctly and incorrectly. A brief written explanation was also given so that the participant could understand why the correct answer was correct.   After completion of the practice problems, the participant took a one-minute break. After the break, the participant completed a short demographics questionnaire (Appendix E). The participant then had to walk to the experimenter to turn in the original scratch piece of paper and receive a new piece of scratch paper. The participant was then instructed to take a final exam over the material learned in the lecture video. The final exam consisted of 15 questions and the participant had 15 minutes to complete the test (Appendix F).

The interactive procedure began with the participant viewing the lecture that was interrupted with questions about every minute. The participant had to answer the question before the video resumed. The participant received feedback on if the question was answered correctly or incorrectly. A green check mark appeared next to correctly answered questions and a red "x" appeared next to questions answered incorrectly. The answer to the questions was then explained in the following video segment. The questions that appeared in this lecture type were similar to those in the non-interactive lecture. After the participant finished watching the video, the participant was prompted to take the NASA-TLX. After the NASA-TLX was completed, the participant took a one-minute break. After the break, the participant completed the short demographics questionnaire. The participant then had to walk to the experimenter to turn in the original scratch piece of paper and receive a new piece of scratch paper. The participant then

took the final exam. The final exam was the exact same as the exam given to the participants in the non-interactive condition.

Lastly, the participants completed a questionnaire about their previous mathematics and statistics exposure. After the participant completed all of the components of the experiment, he or she was debriefed. The debrief script appeared on the computer screen after the participant finished the final exam. The experimenter also asked the participant if he or she had any questions regarding the study's procedure. The participant then had the option to enter their name into the raffle drawing for one of four $25 gift cards.

Calculations

Before data analysis could begin, both the quiz scores and NASA-TLX scores had to be calculated. The following describes the procedure for calculating the measures.

Quiz Score. The experimenter graded each question. For the fill-in-the-blank type questions, the participant did not have to give the exact answer for the response to be considered correct. There were no directions given to the participant on how to round numbers. Therefore, the experimenter counted any answer within a one number range as correct. If the answer was not in the range, it was counted as incorrect. One test question, question 13, was deleted when calculating the test score because of the number of participants who incorrectly answered the question (Appendix F). In the interactive group, 27 of the 36 participants answered the question incorrectly. In the non-interactive group, 25 of the 34 participants answered the question incorrectly. Since almost 75% of the participants in both groups got the question wrong, the researcher inferred that this would not be a diagnostic question. An independent samples t-test was also performed to make sure there were no differences between the groups on this question. Correct answers were coded as "1" and incorrect answers were coded as "2." The interactive group (M = 1.75, SD = .44) did not differ from the non-interactive group (M = 1.74, SD = .45) on the test question that was thrown out, t (68) = .14, p = .89. It was obvious that the question was poorly written or tricky for participants to answer. Thus, the highest score possible on the quiz was a 14.

NASA-TLX Score. NASA-TLX scores were calculated based on the NASA-TLX instructions. The weights for each of the dimension were calculated based on the 15 pairwise comparisons. Then, the ratings from each of the six dimension scales were multiplied by the calculated weights. Thus, each dimension had its own score. The scores were also summed creating the overall NASA-TLX score. This process was completed for both the first NASA-TLX the participant completed and the second NASA-TLX the participant completed. The first NASA-TLX score measured the amount of perceived workload from the orientation video. This served as a baseline task because it measured how much cognitive workload the participant experienced just from watching a video. The workload scores did not differ between the interactive group (M = 15.45, SD = 18.88) and the non-interactive group (M = 22.53, SD = 22.29) on the baseline NASA-TLX survey total, t(68) = -1.44, *p* = .15 (Figure 1).
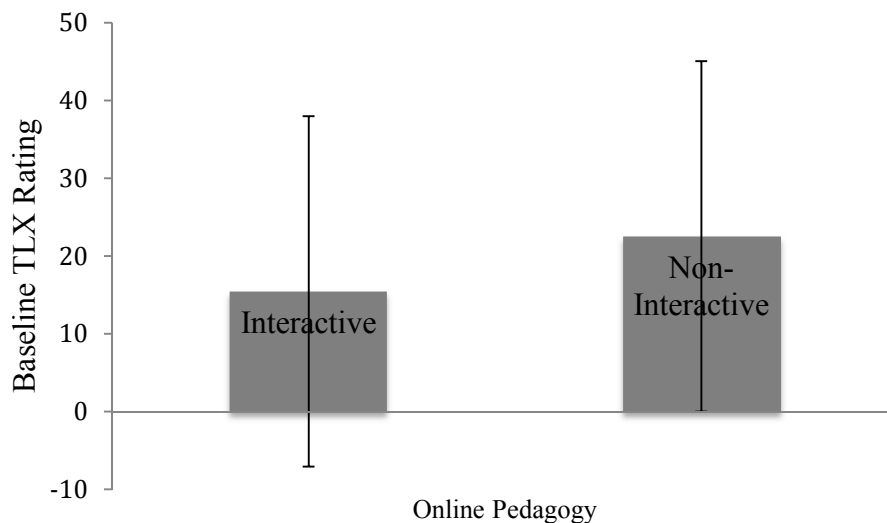
Figure 1. Baseline perceived workload scores (*N* = 70). Workload was measured using the NASA-TLX. In the interactive condition M = 15.45, *SD* = 22.53. In the non-interactive condition M = 22.53, *SD* = 22.29. Error bars represent ±1 *SD*.

The second NASA-TLX was the amount of cognitive workload experienced by the participant from the instructional video and orientation video. The score calculated from the second NASA-TLX was referred to as "overall workload." In addition, the first NASA-TLX score was subtracted from the second NASA-TLX score to create a measure of the workload from just the treatment video. This score was referred to as "treatment workload." Thus, two separate TLX scores were used in the subsequent analyses.

Planned Analyses

Learning. Using an independent samples t-test, it was found that there was not a significant difference between the types of online pedagogies, such that viewing either the interactive ($M = 10.83$, $SD = 2.21$, n = 36) or non-interactive ($M = 11.08$, $SD = 2.21$, n = 34) video resulted in similar test scores, $t(68) = -.483$, $p = .63$. Both the interactive and non-interactive videos had high scores of 14 and low scores of 6.

Subjective Cognitive Workload. Cognitive Workload was measured using the NASA-TLX. Participants completed two NASA-TLX surveys. The first survey was given to participants to get a measure of perceived workload of watching a video. The second NASA-TLX survey measured a participant's perceived workload after watching an instructional statistics video. A paired samples t-test revealed that there was a significant difference between the first NASA-TLX ($M = 18.89$, $SD = 20.77$) and second NASA-TLX survey ($M = 36.63$, $SD = 22.52$), $t(69) = -9.42$, $p < .001$, such that workload increased between when the first NASA-TLX was completed and when the second NASA-TLX was completed.

Mediation Analysis for Overall Workload. The hypothesized model tested the relationship between online pedagogy, overall cognitive workload, and learning (Table 1). The mediation model proposed by Baron and Kenny (1986) was used for data analysis. Step one established that there was not a significant relationship between online pedagogy and learning ($\beta = .058$, $p = .63$) and step two established there was not a significant relationship between online course pedagogy and cognitive workload ($\beta = -.008$, $p = .95$). Step three examined the relationship between cognitive workload and learning, which was significant, ($\beta = -.214$, $p = .045$), such that higher workload scores resulted in lower learning performance (Figure 2). In a true mediation analysis, a fourth step (if online pedagogy and workload were significant predictors of learning) would have been examined. However, the mediation analysis failed to meet the requirement that steps one through three show significance. Therefore, step four was not completed. Step four was not completed for any of the subsequent analyses either for the same reason. The relationship that online pedagogy is mediated by cognitive workload on learning could not be established. The subsequent mediation analyses look at other relationships between online pedagogy, cognitive workload, and learning; however, no mediating variables were found.

Table 1

Summary of Mediation Regression Analysis for Overall Workload as a Mediator (N = 70)

Variable	B	*SE* B	β	R^2
Step 1: online pedagogy predicts learning				
Online pedagogy	.255	.528	.058	.003
Dependent variable: Learning				
Step 2: online pedagogy predicts cognitive workload				
Online pedagogy	.351	5.425	-.008	.001
Dependent variable: cognitive workload				
Step 3: cognitive workload predicts learning				
Cognitive workload	.023	.011	.241*	.058
Dependent variable: Learning				

* $p < .05$
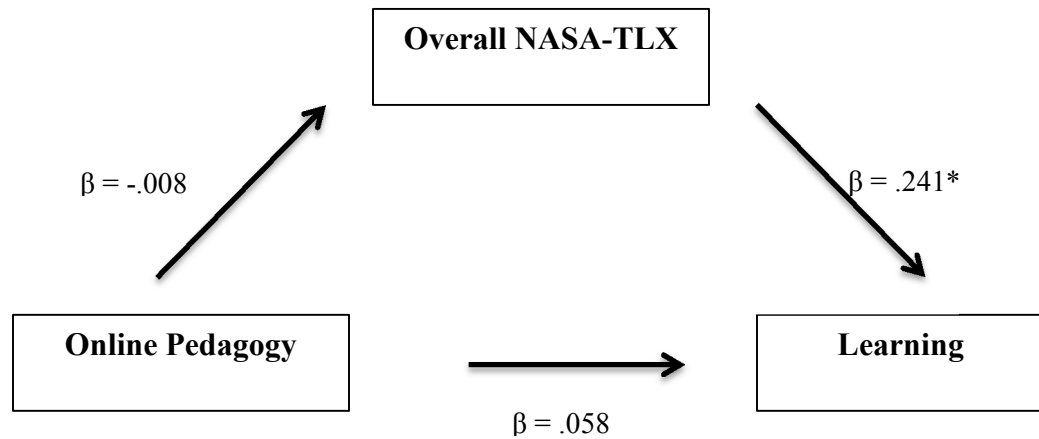
Figure 2. Mediation analysis overall workload (*N = 70).* * $p < .05$

Meditation Analysis for Treatment Workload. The hypothesized model tested the relationship between online pedagogy, treatment cognitive workload, and learning (Table 2). Step one established that there was not a significant relationship between online pedagogy and learning ($\beta = .058$, $p = .63$). Step two established there was a significant relationship between online course pedagogy and cognitive workload ($\beta = -.237$, $p = .048$), such that the interactive video resulted in higher perceived workload scores than the non-interactive video. Step three examined the relationship between treatment cognitive workload and learning, which was not significant ($\beta = .018$, $p = .88$) (Figure 3).

Table 2

Summary of Mediation Regression Analysis for Treatment Workload as a Mediator (N = 70)

Variable	B	SE B	β	R^2
Step 1: online pedagogy predicts learning				
Online pedagogy	.255	.528	.058	.003
Dependent variable: Learning				
Step 2: online pedagogy predicts cognitive workload				
Online pedagogy	7.429	3.688	.237*	.056
Dependent variable: cognitive workload				
Step 3: cognitive workload predicts learning				
Cognitive workload	0.003	.017	.018	.001
Dependent variable: Learning				

$* p < .05$
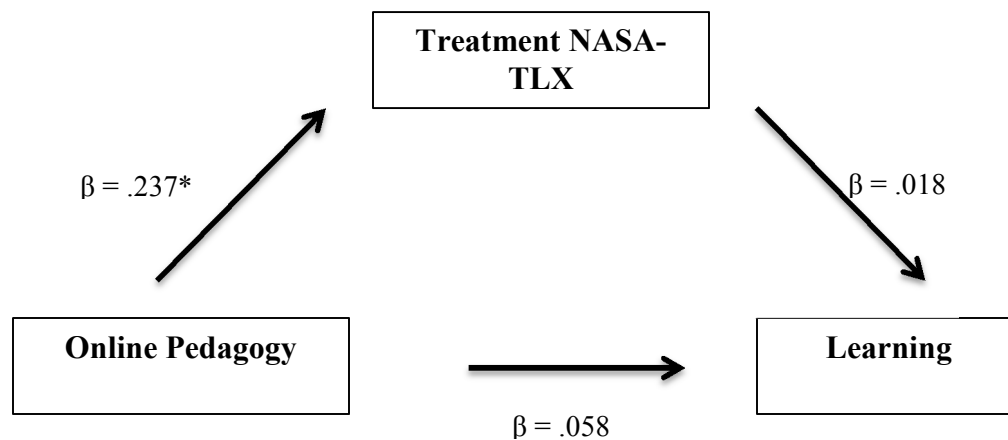
Figure 3. Mediation analysis treatment workload (*N* = 70). * *p* < .05

Subsequent Analyses

Based on the results of the mediation analyses, cognitive workload was not a mediator in the relationship between online course pedagogy and learning. Thus, more exploratory analyses were conducted to test the relationships between online course pedagogy and cognitive workload. Quiz score was not included in the analyses, as the previous analyses showed null results. Only treatment workload scores were used because the researcher was interested in looking at online the effect of the type of online course on workload rather than the workload from the orientation and instructional video. Using online the treatment scores allowed the researcher to evaluate any potential differences that occurred only from the segmentation.

The interactive course had a higher treatment workload score (M = 21.35, *SD* = 17.47) than the non-interactive workload score (M = 13.92, *SD* = 12.90). First, an independent samples t-test was conducted between the online course pedagogy and the treatment workload score. There was a significant relationship between online course pedagogy and treatment workload score, t(68) = 2.01, *p* = .048, d = .48, such that the overall workload experienced between the two groups was affected by the online course pedagogy. Those in the interactive video condition experienced more workload than the non-interactive condition.

Then, a series of independent samples t-tests were conducted with online pedagogy as the independent variable and each of the six treatment NASA-TLX dimensions as the dependent variable. There was a significant relationship between online course pedagogy and the effort dimension of the NASA-TLX, t(68) = 2.80, *p* = .007, d = .67. The interactive video (M = 110.75, *SD* = 113.62) was significantly more effortful for participants than the non-interactive video (M = 36.41, *SD* = 107.53). However, there were no significant differences in the other workload dimensions.

There was no significant relationship between interactive mental demand (M = 106.39, *SD* = 104.03) and non-interactive mental demand (M = 115.59, *SD* = 126.66), t(68) = -33, *p* = .74. There was no significant relationship between interactive physical demand (M = 2.56, *SD* = 29.42) and non-interactive physical demand (M = 3.79, *SD* = 60.40), t(68) = -.11 *p* = .91. There was no significant relationship between interactive temporal demand (M = 36.33, *SD* = 89.13) and non-interactive temporal demand (M = 57.18, *SD* = 95.74), t(68) = -.94, *p* = .35. There was no significant relationship between

interactive performance (M = 17.33, *SD* = 156.73) and non-interactive performance (M = -31.23, *SD* = 170.56), t(68) = 1.25, *p* = .22. There was no significant relationship between interactive frustration (M = 46.94, *SD* = 91.64) and non-interactive frustration (M = 27.15 *SD* = 74.99), t(68) = .99, *p* = .33.

Meditation Analysis for Overall Effort. After finding that the effort NASA-TLX scores were significant, the effort dimension of the NASA-TLX was included in a separate mediation analysis. The effort dimension of the NASA-TLX was the only dimension that was significantly related to cognitive workload (Tables 3 and 4). This model tested the relationship between online pedagogy, overall effort, and learning (Table 5). Step one established that there was not a significant relationship between online pedagogy and learning, (β = .058, *p* = .63). Step two established there was a not significant relationship between online course pedagogy and overall effort, (β = -.046, *p* = .70). Step three examined the relationship between overall effort and learning, which was significant, (β = -.265, *p* = .03), such that higher levels of perceived effort resulted in lower learning performance (Figure 4).

Table 3

Pearson Correlations, Treatment Workload (N = 70)

Variable	1.	2.	3.	4.	5.	6.	7.	8.
1. Online Pedagogy	--							
2. Quiz Score	.06	--						
3. Treatment NASA-TLX: mental demand	.04	.01	--					
4. Treatment NASA-TLX: physical demand	.01	.15	-.07	--				
5. Treatment NASA-TLX: temporal demand	.11	.01	.21	-.14	--			
6. Treatment NASA-TLX: performance	-.15	.03	-.23	-.23	-.01	--		
7. Treatment NASA-TLX: Effort	-.32 **	-.09	.03	.23	.01	-.12	--	
8. Treatment NASA-TLX: Frustration	-.12	.01	-.07	-.15	.21	.02	-.19	--

* p < .05
** p < .01

Table 4

Pearson Correlations, Overall Workload (N = 70)

Variable	1.	2.	3.	4.	5.	6.	7.	8.
1. Online Pedagogy	--							
2. Quiz Score	.06	--						
3. Overall NASA-TLX: mental demand	.19	-.20	--					
4. Overall NASA-TLX: physical demand	.07	-.08	.23	--				
5. Overall NASA-TLX: temporal demand	.13	-.14	.43**	.21	--			
6. Overall NASA-TLX: performance	.17	-.04	.05	-.04	.05	--		
7. Overall NASA-TLX: Effort	.05	-.27*	.44**	.23	.28*	.04	--	
8. Overall NASA-TLX: Frustration	.14	-.01	.00	-.05	.20	.13	.03	--

$* p < .05$

$** p < .01$

36

Table 5

Summary of Mediation Regression Analysis for Overall Effort as a Mediator (N = 70)

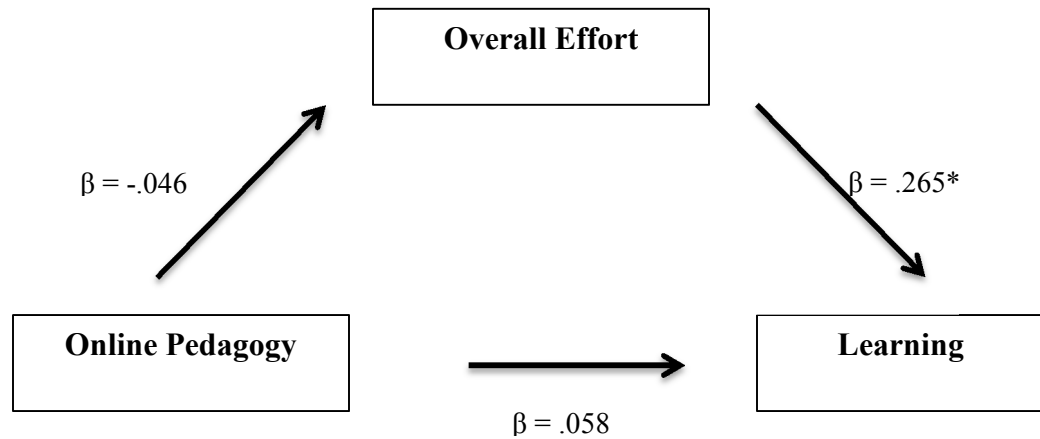| Variable	B	*SE* B	β	R^2
Step 1: online pedagogy predicts learning				
Online pedagogy	.255	.528	.058	.003
Dependent variable: Learning				
Step 2: online pedagogy predicts effort				
Online pedagogy	10.123	26.582	-.046	.002
Dependent variable: effort				
Step 3: effort predicts learning				
Effort	-0.005	.002	.265*	.070
Dependent variable: Learning				

* $p < .05$
** $p < .01$

Figure 4. Mediation analysis overall effort ($N = 70$). * $p < .05$

Meditation Analysis for Treatment Effort. The final model tested the

relationship between online pedagogy, treatment effort, and learning (Table 6). Step one

established that there was not a significant relationship between online pedagogy and learning, (β = .058, p = .63). Step two established there was a significant relationship between online course pedagogy and treatment effort, (β = -.322, p = .007), such that the interactive video resulted in higher perceived effort scores than the non-interactive video. Step three examined the relationship between treatment effort and learning, which was not significant, (β =-.087, p = .47) (Figure 5).

Table 6

Summary of Mediation Regression Analysis for Treatment Effort as a Mediator (N = 70)

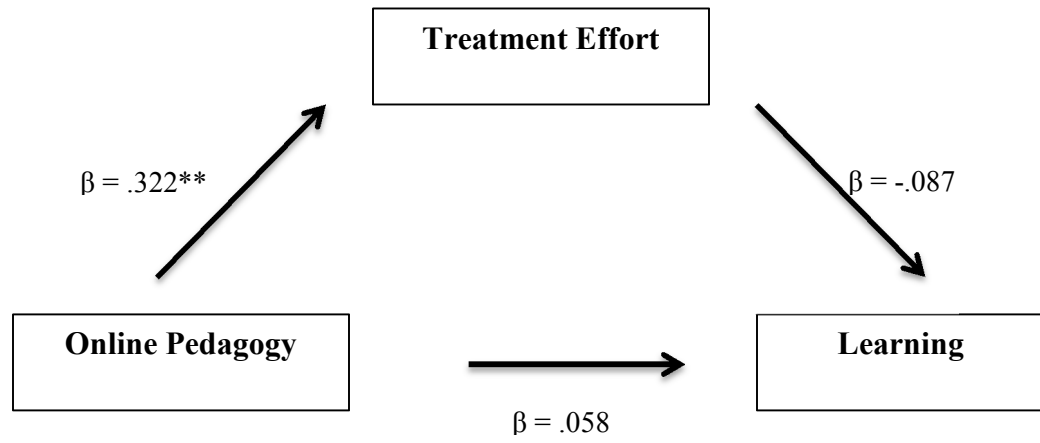| Variable	B	*SE* B	β	R^2
Step 1: online pedagogy predicts learning				
Online pedagogy	.255	.528	.058	.003
Dependent variable: Learning				
Step 2: online pedagogy predicts effort				
Online pedagogy	74.338	26.475	.322**	.104
Dependent variable: effort				
Step 3: effort predicts learning				
Effort	-0.002	.002	-.087	.008
Dependent variable: Learning				

* $p < .05$
** $p < .01$

Figure 5. Mediation analysis treatment effort ($N = 70$). $*$ $p < .05$

Discussion

Analyses were conducted to examine the relationships between online course pedagogy, cognitive workload, and learning. Cognitive workload was not a significant mediator in the relationship between online course pedagogy and learning. However, online course pedagogy was significantly related to cognitive workload, specifically, on the dimension of effort. The results did not support the hypotheses. The interactive course pedagogy had higher cognitive workload scores than the non-interactive pedagogy, which was opposite of what was predicted. However, although the interactive course had higher cognitive workload scores, the end quiz results were very similar for both groups.

Engagement

One possibility for why the results differed from the hypotheses is that participants in the interactive group had to put more effort into the course. However, since both groups had similar quiz scores, it can be inferred that the limits of working memory were not reached in the interactive course. On the one hand, putting more effort into a task could result in a student being more engaged. Engagement is another dependent variable that could be measured in future studies. Although cognitive workload scores were higher, cognitive workload may have not impaired learning because the students were just more engaged in the course. It is only when working memory resources are overworked that cognitive overload becomes a detriment to learning.

On the other hand, the students may have perceived the task as being more effortful because they had to answer questions during the course. Instead of passively watching the video, the interactive video required the participants to think about the topics being presented to them, thus engaging them in the course material more than the non-interactive group. College students may have perceived the task as being effortful because it was different than what college students are normally exposed to. Most classroom lecture styles present information to students in a slide presentation. Students do not have to actively think if they do not choose to do so. Students may also not be as engaged.

Quiz

Another reason why the results were not significant may be due to the tool used to measure the dependent variable. The quiz may not have been a sensitive measure for learning in this experiment. Several aspects of the quiz may have caused it not to be a sensitive measure in this experiment. First, the researcher tried to make questions of varying level of difficulty, but the questions may have still been too easy for the participants. In actuality, the questions may have not been as varied as originally thought. Some questions required the participants to simply define concepts while other questions required more abstract thought, but the questions were all either multiple choice or fill-in-the-blank. For future studies, short-answer questions or an essay question may work better to distinguish between the learning in each group, as there may be more variability in participants' answers.

The quiz may have also not been composed of enough questions to see a difference in scores. There were only 14 questions on the quiz, so the distribution of possible scores on the quiz was not inherently large. More questions may have made it easier to distinguish between the groups. A better idea of which concepts were mastered from the course could be obtained with more questions. However, it may be difficult to administer a longer test in an experimental situation. Participants may lose motivation and start to guess or randomly answer the questions. A larger reward may need to be offered in this case. However, a longer quiz may better mimic a real online class. This predicament would need to be carefully thought out in a future experiment.

Finally, the quiz may have been given to the participants too close to the ending of the video. Participants may have not had enough time to forget any information; thus, the information was still in short-term memory. Administering the quiz to students after a day or a week could have been valuable because it could measure how much information was retained long-term. Online courses differ on when students can take quizzes or tests, which may be another aspect of online course pedagogies that needs to be researched.

Effort

Effort was the dimension of the NASA-TLX cognitive workload survey that was significant. Overall effort was significant in the relationship between effort and learning. Treatment effort was significant in the relationship between online course pedagogy and effort. The type of online course and effort were related such that the interactive course was more effortful than the non-interactive course. This makes sense because

participants had to answer questions as they watched the video. Participants may have perceived the course as being more effortful because they were not familiar with the format. If the participants had completed a longer lecture or multiple lectures, the perceived effort may have decreased as the participants became more familiar with the course format. However, it may be good that perceived effort was greater for the interactive video. Effort was defined as "how hard did you have to work (mentally and physically) to accomplish your level of performance?" Participants in the non–interactive group may have had lower effort scores after the short course (remember that the course was only about 20 minutes long), but later had higher effort scores. As the non-interactive group watched the videos for longer, it may have become harder for them to concentrate because of the monotonous task of watching a video without completing other tasks.

Implications for Teaching

Based on the results of this study, teachers should be cautious when determining the online pedagogy for their online courses. While the videos resulted in the same learning outcomes, the work required for the instructor to create the two learning experiences is quite different. The interactive lectures require considerably more editing and pre-production compared to the non-interactive lectures. This represents a potentially significant cost with no apparent learning benefit, at least in the current study. Long-term studies may find different results that justify the added production costs. The type of students that a teacher has in an online class may also make a difference. Some students may prefer a more interactive video while others prefer a non-interactive video.

The present study did not take into account the attitudes of students towards online classes. If students perceive the online course as being more effortful, they may give the class lower ratings because the students had to work harder than they needed too to achieve the same learning outcome. There may have also been differences in previous exposure to online courses and/or openness to online courses. Students who have not had previous exposure to or are not open to online courses may have perceived the online course as being more effortful.

Limitations

There are several other limitations to this study. First, participants in this study were not able to ask questions while learning the material. Students may have had lower cognitive workload scores if they were able to ask clarifying questions during the online lecture video. In addition, interaction with other classmates or a teacher through online communication may have reduced cognitive workload because the participant would be able to clarify concepts and not have to ruminate about information that was unclear.

Another limitation to the study was the use of a subjective measure of cognitive workload. Subjective workload measures do not always match up with objective workload measures. Although there is much debate, it would be interesting to use an objective measure to see if the results were similar. Participants may have also not carefully read the directions for the NASA-TLX. If the participants did not have the same understanding of the terms, the results may be skewed or the survey may have been measuring different things because of the different interpretations of the terms. There was no check to see if the participants all had the same understanding of the dimensions

of the NASA-TLX. Also, the performance measure of the NASA-TLX was reversed scored. Some participants could have overlooked this, although it was mentioned in the directions. These limitations to the cognitive workload measure may have influenced the results.

Participants were tested in groups of one to nine. The participants who had other people in the room may have had more distractions or been influenced because others were moving around or needing assistance. The participant may have felt behind or too far ahead of the other participants. The amount of noise heard by the participant was minimized, but it may have had an effect on the results. Also, participants do not usually complete online courses in a laboratory at a university; therefore, the results might differ if the participants completed the experiment at home.

The final limitation was that there were a few participants who experienced problems with video buffering. When the Internet connection was not strong, the added wait time and assistance from the experimenter may have affected cognitive workload scores. All of these factors could have made the participant more stressed and may have cause the participant to have higher perceived workload. However, since these extraneous factors were not recorded, there is no way of determining whether or not they contributed to perceived workload.

Future Directions

Previous research has focused on changing the media type in online courses. The present study focused on changing the online pedagogy. However, more research needs

to be completed on other online pedagogies. Longitudinal studies should also be carried out to see how time and practice with videos might change learning outcomes.

Allowing the participants to more freely manipulate the videos would also be something to explore in the future. Participants were not made aware of fast-forward or rewind functions of the video. Furthermore, allowing participants to manipulate video functions may reduce cognitive workload. Participants could rewind and watch something that was unclear to them again, or fast-forward through sections with material that they feel comfortable with. In addition, longitudinal studies may be beneficial to fully capture how interactive and non-interactive pedagogies affect the learner throughout the course. The difference between interactive and non-interactive pedagogies may become more pronounced as the course material becomes more complex. Concepts such as mean, median, and mode are fundamental concepts within an elementary statistics courses and are more easily comprehended compared to later more abstract discussions such as hypothesis testing or the normal distribution.

Conclusion

The online pedagogy did not significantly affect learning performance. Cognitive workload was not a mediating variable of the relationship between online pedagogy and learning. However, the type of online pedagogy had an affect on cognitive workload such that the interactive video resulted in higher cognitive workload, specifically on the dimension of effort. Cognitive workload was also able to predict learning such that higher cognitive workload scores resulted in lower test scores. The dimension of effort was significantly related to learning such that more perceived effort in the course resulted

in lower test scores. Furthermore, it is speculated that the dependent variable, the quiz, may not have been an effective measure in the study. Further research needs to be conducted to further investigate the relationships between online pedagogy, cognitive workload, and learning. Research on online pedagogies also needs to be conducted to explore the method of online learning that will result in the greatest learning gains. With a rise in the need and popularity of online courses, the research will be able to guide instructors in the best teaching method for online courses so that online courses can be an effective method for learning.

References

Abeysekera, L., & Dawson, P. (2015). Motivation and cognitive load in the flipped classroom: Definition, rationale, and a call for research. *Higher Education Research and Development*, *34*, 1-14. doi:10.1080/07294360.2014.934336

Alraimi, K. M., Zo, H., & Ciganek, A. P. (2015). Understanding the MOOCs continuance: The role of openness and reputation. *Computers and Education*, *80*, 28-38. doi:10.1016/j.compedu.2014.08.006

Arbaugh, J. B. (2014). System, scholar or students? Which most influences online MBA course effectiveness?. *Journal of Computer Assisted Learning*, *30*, 349-362. doi:10.1111/jcal.12048

Baddeley, A., & Hitch, G. (1974). Working memory. In G.A. Bower (Ed.), The psychology of learning and motivation, Vol. 8 (pp. 47-89). New York: Academic Press.

Bowen, W. G., Chingos, M. M., Lack, K. A., & Nygren, T. I. (2014). Interactive learning online at public universities: Evidence from a six-campus randomized trial. *Journal of Policy Analysis and Management*, *33*, 94-111. doi:10.1002/pam.2172

California State University (2014, September 8). Impacted undergraduate majors and campuses in the California State University 2015-2016. Retrieved from http://www.calstate.edu/sas/impactioninfo.shtml

Chen, C., & Wu, C. (2015). Effects of different video lecture types on sustained attention, emotion, cognitive load, and learning performance. *Computers and Education*, *80*, 108-121. doi:10.1016/j.compedu.2014.08.015

Chi, M. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*, 219-243. doi:10.1080/00461520.2014.965823

Emerson, L., & MacKay, B. (2011). A comparison between paper-based and online learning in higher education. *British Journal of Educational Technology*, *42*, 727-735. doi:10.1111/j.1467-8535.2010.01081.x

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), Human mental workload (pp. 139–184). Amsterdam, Netherlands: North-Holland.

Jenkins, S. J., & Downs, E. (2003). Demographic, attitudes, and personality differences

reported by students enrolled in online versus traditional courses. *Psychological Reports*, *93*, 213-221. doi:10.2466/PR0.93.5.213-221

Johnson, R. R., Popovic, D. P., Olmstead, R. E., Stikic, M., Levendowski, D. J., Berka, C,. (2011) Drowsiness/alertness algorithm development and validation using synchronized EEG and cognitive performance to individualize a generalized model. *Biological Psychology, 87*, 241-250.

Leibold, M., & Voelpel, S. (2013). Review of the innovative university: Changing the DNA of higher education from the inside out. *Academy of Management Learning and Education*, *12*, 676-686. doi:10.5465/amle.2013.0301

Margaryan, A., Bianco, M., & Littlejohn, A. (2015). Instructional quality of massive open online courses (MOOCs). *Computers and Education*, *80,* 77-83. doi:10.1016/j.compedu.2014.08.005

Matthews, G., Reinerman-Jones, L. E., Barber, D. J., & Abich, J. I. (2015). The psychometrics of mental workload: Multiple measures are sensitive but divergent. *Human Factors*, *57*, 125-143. doi:10.1177/0018720814539505

Milligan, A. T., & Buckenmeyer, J. A. (2008). Assessing students for online learning. *International Journal on E-Learning*, *7*, 449-461.

Paas, F., van Gerven, P. M., & Wouters, P. (2007). Instructional efficiency of animation: Effects of interactivity through mental reconstruction of static key frames. *Applied Cognitive Psychology*, *21*, 783-793. doi:10.1002/acp.1349

Rogers, R., Laraway, S., & Kormanik K. (n.d.) Lesson 3: Central tendency. *Intro to Descriptive Statistics*. [Video File] Retrieved from https://classroom.udacity.com/ courses/ud827/lessons/1489118552/concepts/77367476023

San Jose State University (2015). Institutional effectiveness and analytics. Retrieved from http://www.iea.sjsu.edu/cognos/cgin/cognos.cgi?b_action=xts.run&m=port al/cc.xts&gohome=

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*, 257-285. doi:10.1207/s15516709cog1202_4

Wong, A., Leahy, W., Marcus, N., & Sweller, J. (2012). Cognitive load theory, the transient information effect and e-learning. *Learning and Instruction*, *22*, 449-457. doi:10.1016/j.learninstruc.2012.05.004

Appendix A: Math Abilities Questionnaire

The following are questions related to your previous math studies. Please answer these questions as accurately and honestly as you can.

1. Have you taken the equivalent to two years of high school algebra courses?
   Yes   No

2.. Have you passed the Entry Level Math exam (ELM) at San Jose State University?
   Yes   No

3. Have you taken Advance Placement (AP) or International Baccalaureate (IB) Statistics?
   Yes   No

4a. Have you taken or are you currently enrolled in a statistics course at San Jose State University?
   Yes   No

   4b. If yes, which statistics course(s) are you enrolled in or have taken?

Appendix B: Consent Form

REQUEST FOR YOUR PARTICIPATION IN RESEARCH

TITLE OF THE STUDY
Cognitive Workload in Online Courses: An Investigation of the effects of Online Course Format on Learning Outcomes

NAME OF THE RESEARCHER
Dr. Ronald Rogers, San Jose State University
Kallan Christensen, San Jose State University Experimental Psychology Graduate Student
Department of Psychology

PURPOSE
You have been asked to participate in a research study investigating the effectiveness of online courses. This research will be used to find ways to help teachers design effective online course designs.

PROCEDURES
This research study is expected to take 90 minutes of your time and will be conducted on campus at San Jose State University. The study will include taking surveys and watching videos.

POTENTIAL RISKS
Anxiety may be an emotional discomfort associated with participating in this study. You will be asked to complete a math test that may cause anxiety due to the fact that you will be in a testing situation.

POTENTIAL BENEFITS
The research will allow you to gain knowledge on basic statistics concepts. The research will have indirect benefits that contribute to the general body of research on online courses. New methods will be explored that may help students to perform and learn more effectively in online course settings.

COMPENSATION
Course credit for the amount of time spent participating in the study will be given as compensation for your time. You will only receive credit for the sections of the study you complete. Additionally, you may have the chance to win one of four $25 gift cards.

CONFIDENTIALITY
Although the results of this study may be published, no information that could identify you will be included.

PARTICIPANT RIGHTS

Your participation in this study is completely voluntary. You can refuse to participate in the entire study or any part of the study without any negative effect on your relations with San Jose State University. This consent form is not a contract. It is a written explanation of what will happen during the study if you decide to participate. You will not waive any rights if you choose not to participate, and there is no penalty for stopping your participation in the study.

QUESTIONS OR PROBLEMS

You are encouraged to ask questions at any time during this study.

- For further information please contact Kallan Christensen.
- Complaints about the research may be presented to Ronald F. Rogers, San Jose State University Psychology Department Chair at 408-924-5652.
- For questions about participants' rights or if you feel you have been harmed in any way by your participation in this study, please contact Dr. Pamela Stacks, Associate Vice President of Graduate Studies and Research, San Jose State University, at 408-924-2479.
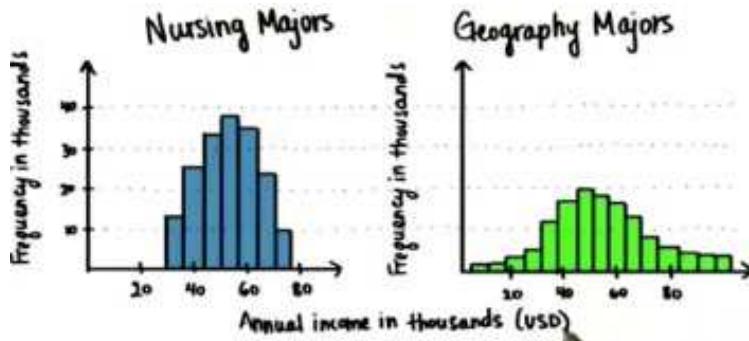
AGREEMENT TO PARTICPATE

Please select from the choices below. If you click agree, it is implied that you have read the information above about the research, your rights as a participant, and give your voluntary consent. Please print out a copy of this page and keep it for your records.

I agree to participate in the research.	I do not agree to participate in the research.

Nursing Majors

Geography Majors

Frequency in thousands

Frequency in thousands

Annual income in thousands (USD)
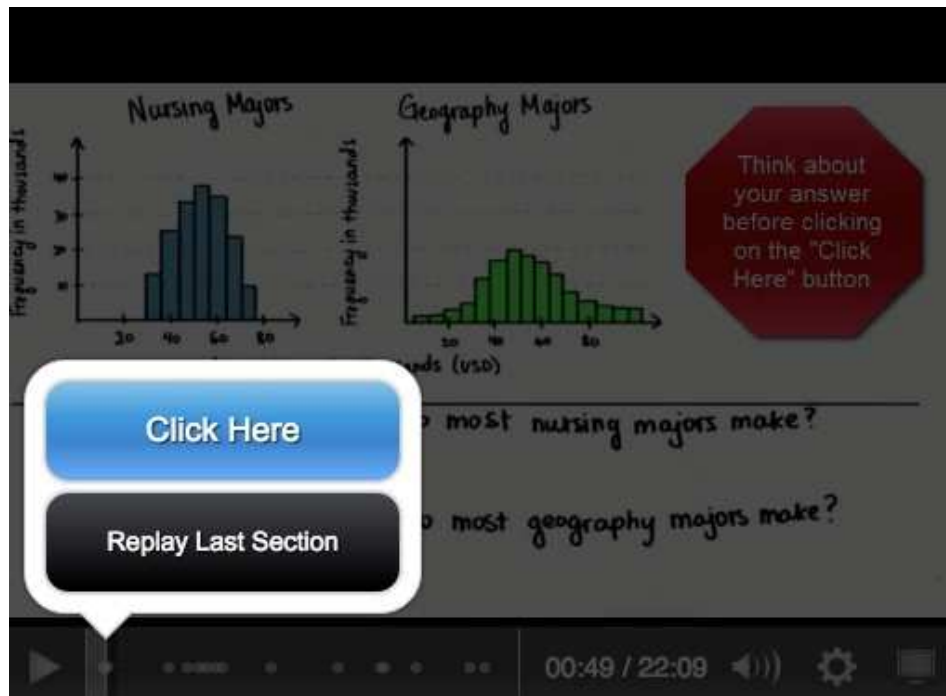
Approximately what income do most nursing majors make?

Between $50,000 and $60,000

Approximately what income do most geography majors make?

Appendix D: Sample Interactive Video Question

Appendix E: Demographics Questionnaire

Below are a series of questions related to the study. Please answer them as accurately as you can. Please make sure all questions are completed. Be assured that information provided is anonymous. Should you have any questions, please notify the experimenter.

1. Age: _____

2. Gender: □ Male □ Female □ Other

3. Is English your primary language? □ Yes □ No

4. How well do you understand English?

1	2	3	4	5	6	7
Not at all				Moderately		Very well

Note: The correct answers are in bolded text. When the exam was given to the participants the correct answers were not bolded. Also, question 13 was excluded when calculating the quiz scores.

1. Which of the following is true for the normal distribution?
The mean is greater than the median
The mean is less than the mode
The mean is greater than the mode
The mean and the mode are both greater than the median
The mean, median, and mode are all equal

2. If you wanted to know which of two websites was more popular than the other (in terms of the
frequency of unique visits last year). What measure of central tendency would be **most** appropriate for these data and why?
The mean because it is not affected by outliers
The mean because it has useful mathematical properties
The mode because the data are categorical
The mode because it has useful mathematical properties
The median because it is not affected by outliers
The median because it has useful mathematical properties

3. A researcher was interested in the average number of books read by high school students per year. Included in the sample of 25 were two students who were members of a book reading club at their high school that provided rewards for the most books read in a year. These two students were first and second place in that contest. What measure of central tendency would be most appropriate for these data and why?
The mean because it is not affected by outliers
The mean because it has useful mathematical properties
The mode because the data are categorical
The mode because it has useful mathematical properties
The median because it is not affected by outliers
The median because it has useful mathematical properties

Scenario for questions 4- 10

City A and City B both entered into a contest to see which city could attract more tourists. The guidelines for the contest stated that data on the total number of tourists would be collected over 5 years. After 5 years, a winner would be chosen. The following data set represents the number of tourists that visited each city over the past 5 years.

City A	City B
267	289
264	251
290	276
283	268
264	268

4. Find the mean for City A
   273.6

5. Find the mean for City B
   270.4

6. Find the mode for City A
   264

7. Find the mode for City B
   268

8. Find the median for City A
   267

9. Find the median for City B
   268

10. The competition rules failed to mention which measure of central tendency would determine the winner. City A and City B are making arguments to the judges about how they should measure the city with the best tourism. Is it the city with the most visitors in a single year, the middle number of visitors over the five years, or an average of total visitors over the 5 years? City A would win the competition if which measure(s) of central tendency determined the winner:

   Mean
   Median

Mode
Mode or Median
None of the above
All of the above

11. Which of the following statements are true?
The mean is in the exact middle of any distribution.
The mean is in the exact middle of skewed distributions.
The mean is affected by extreme scores (outliers).
The mean is not affected by extreme scores (outliers)

12. Which of the following statements are true?
The mode is the average of all the numbers in the dataset.
The mode can be used with categorical data, such as gender or country of origin.
The mode can only be used with numerical data.
The mode is always the best measure of central tendency to use to describe data.

13. Which of the following statements are true?
The median cuts the bottom 50% of scores from the top 50%.
The median separates the extreme scores from the rest of the dataset.
The median is the balance point of distributions of numerical data.
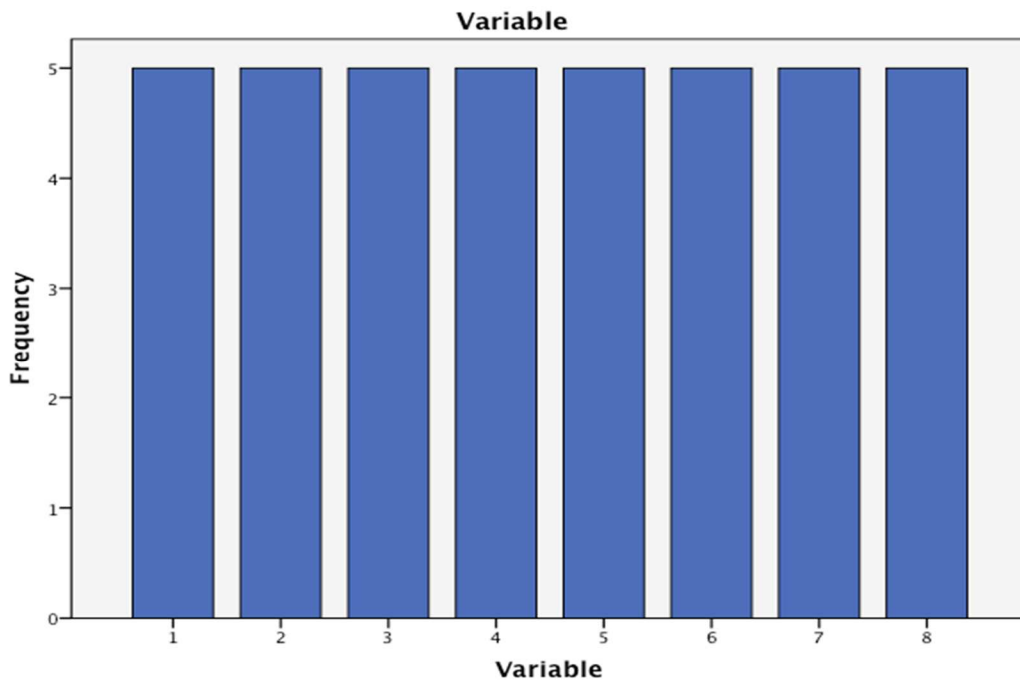The median is the most frequently occurring number.

14. What is the mode for the dataset below?
5
8
more than one mode
no mode

Variable

15. Which of the following is true about the dataset below?
mode < median < mean
median < mode < mean
mean < mode < median
mean < median < mode
mode = median = mean

Variable