Jl. of Computers in Mathematics and Science Teaching (2019) **38**(1), 49-76

Dynamic Versus Static Presentation Formats, Do They Impact Performance Differently?

ELIZABETH MCALPIN, DAVID SHILANE,
AND SELIN KALAYCIOGLU
New York University, USA
em341@nyu.edu
sk168@nyu.edu
dshilane@nyu.edu

According to the Cognitive Theory of Multimedia Learning (Mayer, 2009) learning with multimedia requires the integration of verbal and visual information (e.g., printed text and static images, narrated animation). Studies exploring the effects of dynamic versus static presentations yield mixed results. Few studies explore comparing dynamic versus static presentations to teach mathematics. Most experimental studies occur in a controlled lab setting. The present study aimed to assess the impact of multimedia learning formats beyond the experimental lab to an authentic setting (i.e., classrooms) to teach mathematics. In Spring 2017, we compared weekly outcomes for students who accessed online interactive video modules (DYNAMIC) demonstrating calculus level 1 topics to students who accessed the same content presented in printed text and static image format (STATIC). Students in two sections of undergraduate course Calculus 1 (011) and (001) alternated each week the online pre-work material presented in either interactive video module or text-image document. They were assessed with problem-solving and conceptual questions at the beginning and end of each week. The results demonstrated no statistically significant differences on presentation formats for the assessments provided at the beginning and end of each week. Participants also answered the same survey question at the end of each online topical lesson related to the perceived effectiveness of the instructional material to help students master the topic. Survey results indicate students perceived the weekly instructional content, presented in a static or dynamic format, to be equally helpful to prepare them to master the material.

Keywords: *multimedia, multimedia learning, flipped class, higher education, mathematics, calculus*

INTRODUCTION

There is a lot to consider when designing multimedia instruction for an entire course. When consulting the research literature on multimedia learning to help make informed instructional design decisions to meet the needs for a specific context one must synthesize a plethora of research that compares only a few variables at a time. We find few controlled experimental studies repeated but rather many variations. These numerous studies may vary by the type of knowledge (e.g., declarative, procedural, conceptual), the subject domain (e.g., physics, math, language), the type of assessment (e.g., retention-based, deep understanding), the type of uni-modality (e.g., text, image, animation), the type of multimodality (printed text and image, narrated animation, narrated video with static images, etc.), the type of visual format (e.g., illustrations, diagrams, and photos), the method of study, the type of student (e.g., high versus low prior knowledge), and more. There is also a gap in the research literature on multimedia learning that examines generalizable outcomes from controlled experimental designs in authentic contexts such as a live course. Narrowing down the research to meet specific needs in terms of guidance, the designer is met with either literature that is too broad, too narrow, inconsistent, or incomplete.

The multimedia principles become challenging to apply successfully in practice to authentic environments especially when more research is needed to explore the possibilities (Butcher, 2014). For example, when deciding whether to design instructional presentation materials on calculus level I topics in a dynamic (e.g. narrated video or animation) or static (e.g., image and text) format the research literature suggests an advantage of instructional animations over static pictures for retention based learning outcomes although the effect size reported small-moderate strength, 0.24 (Hoffler and Leutner, 2007). The strength of the effect on learning varies as one dives into the weeds to more specific factors and contexts. For example, research demonstrates the impact of multimedia learning materials (concrete and abstract visual representations) on mathematical problem solving skills applied to near-transfer tasks (Moreno, Ozogul, & Reisslein, 2011) is stronger than the multimedia effect (representational animations) on problem solving skill applied to isomorphic problems (Hoffler and Leutner, 2007).

Other specific factors may also play a role in determining the strength of the effect on learning such as length of time spent on the materials, immediate versus delayed assessment, the type of image with vary levels of detail, and prior knowledge levels in students, to name a few (Butcher, 2014). When applied, the effective strength of any multimedia design becomes exponentially impacted from the number of these design choices to be considered and how one will evaluate the effectiveness of those choices as one ventures down a complex decision tree of factors. The reason is in part due to controlled studies focusing on just a few factors at a time. Yet, to our knowledge, there are no studies that have examined the impact and role of well-designed multimedia mathematical materials in an authentic context.

Literature Review

A. *Multimedia Learning Formats*

According to Schnotz (2005), "comprehension is highly dependent on what kind of information is presented and how it is presented" (p. 49). The type of information may be declarative, procedural, or conceptual. Multimedia presentation format types may be presented as printed text and static images, narrated text and static images, printed text and animation (dynamic images), or animation and narration, to name a few. Multimedia learning "occurs when students build mental representations from words and pictures that are presented to them (e.g., printed text and illustrations or narration and animation)" (Mayer, 2003, p. 125). The goal of multimedia instructional material is to foster deep or meaningful learning (Mayer, 2003). It does not require the use of high technology (Schnotz, 2005) as it is possible through books and classrooms equipped with a simple blackboard. Multimedia learning is only effective if the instructional material is well-designed by employing other supporting multimedia design principles such as split-attention, modality, signaling, and redundancy, to name a few (Mayer, 2003). Multimedia learning theory and the design principle are widely supported by research under lab conditions comparing only one format (e.g., words alone) to two combined formats (e.g., words and pictures) (Mayer, 2003). According to Butcher (2014), "the multimedia principle provides a rationale for research investigating the optimal design of multimedia learning materials, as well as the ways in which learning strategies and learner processing can combine with multimedia materials to result in specific outcomes" (p. 174).

The design of multimedia learning materials leans toward the use of narrated video or animations because they support the integration of words and pictures, through the use of the modality principle (Low & Sweller, 2014). The modality principle supports working memory capacity and reduces cognitive load resources to facilitate learning by presenting instructional information in simultaneous mixed modes (visual and auditory) rather than one mode (visual or auditory) alone. To digest written text and static images requires both formats to share one mode (visual) whereas narrated animations or narrated static pictures use two channels visual and auditory modes. Multiple studies conducted with low prior knowledge students demonstrated spoken over written or printed texts as more effective for learning (Ginns, 2005). The assumption many designers make is that the modality principle should be used to design and present most instructional content because it advocates the dual mode over single mode thus reducing cognitive load and increasing working memory capacity. As a result, media choices such as video or animations take precedence over printed text and static images for presenting most types of instructional content.

Using animations or video, with or without narration, to present instructional material comes with its own set of cognitive processing challenges, namely, the temporal and spatial. Too fast or slow and information will be dropped on behalf of the learner struggling to manage cognitive attention to the material. Since there is a temporal component to the pacing of animations research supporting the design of animations suggest ways to allow the learner to control the pace to match or support their cognitive processing abilities (Schwan & Riempp, 2004). Successful learning from printed text and static image formats despite well integrated designs, requires the learner to process slowly and carefully to create a proper mental representation. Depending on the type and complexity of content learning from text and image may also impose cognitive demands on the learner rendering the experience unsuccessful. Thus, successful learning required from printed text and static image formats compared to learning from narrated animation or video formats may or may not differ based on the affordances of the design, prior knowledge, type of knowledge or skill, type of assessment, complexity of instructional material, and ease of cognitive processing through the selection, organization, transformation, and integration of information (Kombartzky et al, 2010).

The research comparing dynamic versus static formats across different subject matter topics, knowledge types, and assessment types, shows mixed results but overall leans positively toward the dynamic. Some research favors alternatives to animation such as printed text and static images to be more successful for learning (Betrancourt, Morrison, & Tversky, 2002; Betrancourt & Tversky, 2000). In a meta-analysis by Hoffler and Leutner (2007) they reported an advantage of instructional animations over static pictures, albeit, with only a moderate effect size. The meta-analysis conducted by Hoffler and Leutner (2007) testing different types of knowledge also suggest that when the animation is highly realistic (e.g., video-based) and/or when procedural-motor knowledge is to be acquired the learning effects are even greater over static visualizations. In terms of specific instructional domains, such as mathematics and physics, teaching problem-solving skills demonstrated a moderate advantage using animations over static images. The studies in the meta-analysis included annotated text along with the visual representations but they do not appear to include cases with full text, narrated animations or video, or learning outcomes aimed to promote deeper forms of learning. It also does not include cases with video instruction containing static images, animations, or a sequence of both. The results are mixed suggesting it depends on the subject domain, type of content, prior knowledge level of the student, and the design.

To fill in some of the gaps from the meta-analysis by Hoffler and Leutner (2007), Arguel and Jamet (2009) observed an advantage of presenting a mixed video format, namely, narrated video with static images for learning procedural first aid knowledge over narrated static images and narrated video. Open-ended, paper-based questions were used to evaluate the conditions. The design of the narrated video with static images allowed the learner to view the static images in the video throughout as opposed to editing them into the video. This obviated the transient learning effects that can occur with video or animation. It is interesting to note that the Arguel and Jamet (2009) example focused on a very specific to a type of learning, namely procedural-motor skills suggesting the type of media may be more or less effective based on the type of knowledge and skill. König et al. (2012) conducted an experimental study comparing narrated video to static images with printed text on the topic of physical exercise. Again, the type of skill in the design focused on procedural-motor skills. The retention based learning outcomes were higher for participants using the static presentation. The design of the narrated video did not include the use of pervasive static images suggesting possible transient learning effects of the information in the video condition.

Examining multimedia research specifically in the field of mathematics, Nathan et al. (1992) conducted a study using narrated animation to facilitate learning algebra word problems. Learners solved problems more accurately in the animation condition than the non-animation condition. Moreno and Mayer (1999) compared the use of printed text (control group) to animations plus the same printed text (experimental group) to teach elementary students addition and subtraction of signed whole numbers. Participants in the animation and printed text condition performed better on difficult items than participants in the control group. These results suggested a positive effect for both easy and difficult items for high achieving students assigned to the animation condition compared to their peers in the control condition. No differences were detected to low-achieving students (Moreno & Mayer, 1999).

Atkinson (2002) conducted research on ways to present worked examples. In Atkinson's studies comparing spoken text with animations and printed text with animations to printed text with animation, performance on problems solving learning outcomes enhanced significantly for the group using the narrated animations. Scheiter et al. (2009) compared the use of animations with worked-examples to text-based worked-examples for teaching students to solve algebra word-problems. The animations used in the study used a combination of concrete and abstract representations. The experimental condition using the text-based worked examples augmented with animations performed better on problem-solving activities than the control condition using text-based worked examples alone. Gebre Yohannes et al. (2016) conducted a controlled experiment examining the differences between traditional teaching methodology (lecture) and a multimedia based-approach (computer based presentation using MATLAB) on Calculus topics. The experimental MATLAB group received worked examples and isomorphic practice problems. Students in the experimental condition using MATLAB outperformed the control group. The effects of this study may be due to what is commonly called a "practice" effect. Students performed better in the experimental condition because they had more practice compared to the lecture-based condition.

Changing presentation formats from traditional forms such as text and image to newer formats generates debates towards their innovative power. New modalities have the appearance of redefining the experience yet the cognitive and emotional effects do not appear to improve substantially, particularly in certain domains. Pincus et al. (2016), studied the effects of embedded multimedia journalism comparted to print-only and traditional multimedia reporting using the publication of *Snow Fall*. In this study, participants were randomly assigned to three conditions. Each condition contained the same content (news story) but differed in terms of modality and layout. One condition received a printed text only news story with no additional presentation formats. The second condition received a traditional multimedia story whereby the primary storytelling technique was printed text but included add-on elements such as graphics, images, or video. The traditional multimedia format did not embed the add-on elements into the printed narrative rather it provided them off to the side for the reader to integrate with the text. The third condition received an embedded multimedia story whereby the printed narrative embedded the graphics, images, and video into the flow of the layout to provide a more integrated experience rather than off to the side. After engaging with the new story, participants completed a post-test to assess information recall and emotional reactions.

The results from the Pincus et al. (2016) study indicated that participants in the printed text only format of the news story learned slightly more, on average, about the topic than either the traditional multimedia format or the embedded multimedia format. In terms of statistically significant results using ANCOVA, the printed text reached marginal significance over the traditional media format. In sum the printed text ranked highest in terms of knowledge gain, followed by the embedded multimedia format and the traditional multimedia format. The authors of the study concluded that the multimedia format using a combination of modalities within a story may have added cognitive complexity leading to overloading the processing system. The study is not without limitations such as no ability to control for which media components participants attended to most in the multimedia formats. Other research techniques such as eye-tracking may help future studies parse out these layout features. Pincus et al. (2016) admit that the topic used in the study may have been best suited for text only whereas if they used a different topic it may have been better suited for audio and/or video formats, and yet other topics may benefit from a mix of media approaches to tell the story. More research is needed to parse out the differences based on topic.

While the research literature supports “the general prescription that effective learning materials should combine visual and verbal materials targeting to-be learned concepts” (Butcher, 2014, p 175) the literature on specific verbal and visual formats, for specific kinds of information tends to lean towards dynamic presentation formats such as narrated video and animation. But, to date we find examples of research studies conducted in controlled experimental conditions comparing a few variables that do not get tested or replicated in authentic environments. While the literature appears to have stopped around 2016 the relevance has not because implementing principles culled from research into authentic environments does not happen readily. Does any combination of visual and verbal material work the same or differently in any live class for any subject at any grade level? Is there any guidance for an educator to help decide? Does the presentation format, dynamic or static, matter when used in a live course?

Based on this body of literature researchers of this study support the recommendations provided by Butcher (2014) that suggest continued research in multimedia learning of mathematics that moves beyond single experimental conditions to authentic classroom conditions using delayed tests. Much of the aforementioned research focused on geometry instruction, a suitable subdomain to explore combinations of images and words in various modalities. Additional research is needed in other sub domains of mathematics besides algebra and geometry. Lastly, expanding the grade levels of participants to include higher education populations is yet another area for research development and expansion.

B. *Flipped Classroom*

Instructors seeking to change their teaching approach from the traditional lecture format where students just take notes in the classroom to a more cognitively active format leverage a flipped classroom approach often facilitated by media and technology. In this approach lectures are often digitized into videos and accessed by students prior to attending an active learning classroom. By definition, the flipped class approach is any method of delivering part or all instructional content prior to meeting in-class so that the in-class time focuses on student-centered learning activities (Lage et al., 2000). But not all flipped classroom designs are the same and vary based on active learning strategies for the classroom, technologies used to facilitate an active model, as well as the design of pre-class work material. Pre-class work may provide only video lectures, assigned reading, or other multimedia formats. Depending on how well the pre-work materials are designed, they may induce cognitive overload on students' working memory and interfere with cognitive processing abilities thereby reducing effective learning (Clark & Mayer, 2016). Additionally, the design of pre-work materials may or may not be accompanied by student learning activities to support learning objectives and with or without immediate feedback. All of these factors contribute to the success and failure of the flipped model.

In a comprehensive review of the flipped classroom approach, Giannakos et al. (2018) summarized research that, overall, demonstrated increased learning performance, attitudes, and engagement using the flipped approach. The authors also provided challenges to the model as well as future research opportunities namely the lack of specific pedagogical strategies used in the flipped model and high use of the model in computer science and information technology domains. It is also not clear from the prior research provided in the report which aspects of the flipped model provide the most value to students in terms of improved learning, enjoyment, and engagement. But in a study by Gross et al. (2015) evaluating a flipped model in an undergraduate course in biochemistry, researchers attributed increased student outcomes and behaviors to the pre-class preparation material. In this study researcher examined the way students in both formats interacted with the pre-class work materials. A limitation to this study was lack of control for the in-class active learning component in the flipped class section. While the researchers demonstrate how the students in the flipped class section performed better, attempted pre-class homework more often, and engaged earlier with the online pre-class work than the traditional format differences may also have been attributed to the active learning component over and above the way students interacted with the pre-class work materials.

The limited research on flipped class for mathematics and Calculus specifically, demonstrated mixed results in terms of increased performance (McGivney-Burelle and Xue, 2013; Guerrero et al., 2015; Adams & Dove, 2017). McGivney-Burelle and Xue (2013) conducted an experimental study on an undergraduate Calculus course. Results provided only mean comparisons between the two groups with the flipped class, on average, scoring about the same for exam 1, about five points higher on exam 2, and about four points higher in WeBWork assignments. Guerrero et al. (2015) found no statistically significant difference between a flipped and traditional format of a Finite Mathematics course. Success implementing a new approach depends widely on how well the pre-class work and in-class activities are designed and managed. It is unclear from this study how well the new approach was implemented with all the variability factors at play. Making adjustments to the approach and refining it can radically change the outcome.

In a study conducted by Adams and Dove (2017) on the impact of the flipped model on Calculus students' performance and perceptions, despite low sample sizes (N=19) for each condition, results demonstrated substantial improvement over the traditional lecture model while student beliefs about learning mathematics in either condition did not change. In this model, students were asked, but not required or graded, to watch video demonstrations of weekly topics before attending class. In class, students often worked in small-groups to solve problems while the instructor navigated the space to provide guidance. And occasionally students were required to record themselves working through a Calculus problem and submit for feedback from the instructor. While the results on pre-post achievement tests, final exams, and grades demonstrated substantial improvement for the flipped class over the traditional class the course continued to make improvements to the model in order to overcome challenges such requiring students to watch the videos and clearer expectations for group work and creating student videos.

Based on the current body of research on the flipped class model specifically targeting Mathematics and specifically Calculus it is unclear how the design of the elements and features of the model contribute more or less to student success. Research on the design specifically of pre-work material controlling for other elements of the model appear to be missing from the literature.

Background

Calculus I, the largest enrolled undergraduate course at New York University (NYU), provides about 14 sections each year with over 100 students in each section. In the past, the course delivered in-class lecture demonstrations followed by homework to be completed outside of class. As of Fall 2015, in an effort to introduce more time for student-centered activities and decrease instructor-centered demonstration 2 sections out of 14 experimented with a "flipped" model. The flipped teaching method presents an alternative design to the traditional attend lecture then completed homework sequence. In the flipped approach, part of the in-class lecture content and homework is accessed before class, thus, freeing in-class time to provide more student-centered problem solving activities. Even though the flipped model statistically improved overall grades compared to the traditional format we wanted to understand more about the specifics of the model.

In the design of the multimedia learning materials for the Calculus I course, designers adhered to lab-researched multimedia principles (Mayer 2001; Clark & Mayer 2011; Clark & Mayer 2016) supporting well-designed presentation of instructional materials (e.g., multimedia, modality, contiguity, segmenting, coherence, signaling, learner control, and personalization) in order to best support cognitive processing of Calculus concepts, procedures, and theorems. The educational designers embedded practice questions with immediate feedback in the materials to further support learning and retention of this material. After completing the materials students completed additional practice-assessment questions with immediate feedback. These design decisions were based on research by Skuballa et al. (2018) affirming that immediately after presenting the instructional material to be learned, to answer questions, solve problems, and receive immediate feedback related to the instructional material engages and strengthens long-term memory. Additional, albeit delayed, retention testing occurred at the end of the week.

Specifically, in the design of the interactive video modules (*Figure 1*), educational designers and faculty scripted and produced close to 50 modules. Each module contained a mixed sequence of narrated diagrams, illustrations, motion graphics, and animations to teach, through examples, mathematical concepts and principles to support problem-solving skills. Students could pause, stop, start, and rewind the modules to manage processing. Knowledge retention problem-solving questions were embedded at various time points in the video module design to keep students engaged. The design of each interactive video module varied in terms of length, number of diagrams, illustrations, animations, and embedded practice problems. No interactive video module exceeded 15 minutes in length.

To design the printed text and image presentation (*Figure 2*), educational designers used the transcript and screen shots of imagery from the video modules to construct the static format adhering to similar best-practice design principles of the same instructional content. The printed text and static image format also embedded the same practice questions into the design although student responses were captured separately in the learning management system in order to collect data and provide feedback to students. After completing the static presentation students completed the same additional practice-assessment questions with immediate feedback as the alternative multimedia format.

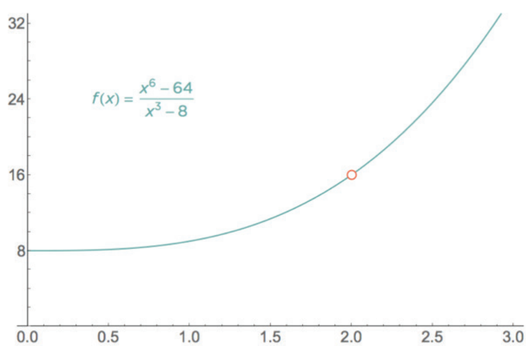


Figure 1. Interactive Video Presentation (DYNAMIC).

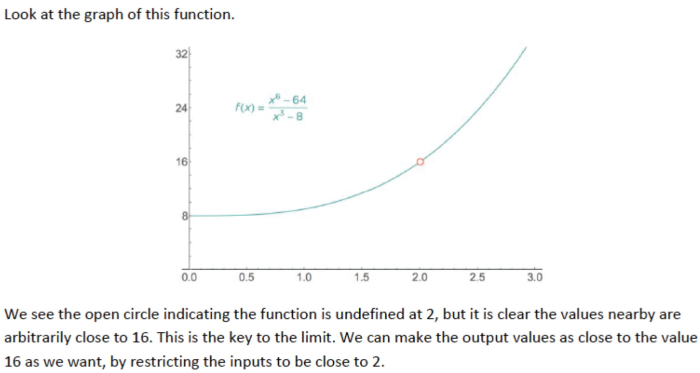


Figure 2. Text & Image Presentation (STATIC).

Essentially everything was the same for both conditions except the format, dynamic or static (*see additional comparative images in appendix*). Students in both sections accessed either weekly presentation format and completed the same activities in the same learning management system before attending the first, in-class lecture demonstration. At the end of the week, during recitation, students were assessed with a paper quiz.

Thus, the goal of this research was to examine whether the presentation of pre-work instructional material in either interactive video format (DYNAMIC) or printed-text and image (STATIC) would affect students' learning outcomes differently on pre-class practice-assessments (immediate) and end-of week quizzes (delayed). Sections were randomly assigned to one of two treatment groups (DYNAMIC or STATIC) that alternated presentation formats for pre-class work each week.

We propose the following **hypothesis**; there is no difference on weekly performance outcomes dependent on the presentation format (DYNAMIC versus STATIC) of pre-class work material.

METHODS

Study design

A quasi-experimental design was used for this project. At NYU, undergraduate students participated in this project during regular scheduled instructional class time for the Calculus I course. Two out of eight sections of the Calculus course participated in the study. The same instructor facilitated both sections. Students self-enrolled into either section of the course using normal enrollment processes. To compensate for the lack of randomization, the two sections used control measures (e.g., SAT/ACT scores, GPA, Diagnostic) to establish a baseline. The Formal Internal Review Board (IRB) approved access and analysis to de-identified data collected at the end of the semester for this project.

A typical semester consists of sixteen weeks. The two sections alternated each week between learning from online lesson modules presented as interactive video or printed text and image formats *(Table 1)*. Students completed formative assessments presented at the end of each online lesson module. Both online lesson module and formative assessments were required to be completed before the first live in-class lecture demonstration of the week. Students were required to attend two 75-minute live in-class lecture demonstrations during the week. At the end of the week students attend one 75-minute in-class recitation period managed by teaching assistants where a paper-based quiz was administered covering topics related to the week. Week six, thirteen, and sixteen were dedicated to exams and did not include online lesson modules nor weekly assessments. Week eight and nine were combined into one lesson to accommodate the holiday break. Week one included formative assessments but did not include an end of the week quiz. Week fifteen included an end of the week quiz but did not include any formative assessments.

Table 1

Study Design

Week	**Section 001**	**Section 011**	**Assessment**
1	Dynamic	Static	Immediate Formative Only
2	Static	Dynamic	Immediate Formative & Delayed Quiz
3	Dynamic	Static	Immediate Formative & Delayed Quiz
4	Static	Dynamic	Immediate Formative & Delayed Quiz
5	Static	Dynamic	Immediate Formative & Delayed Quiz
6	Midterm 1	Midterm 1	Midterm 1
7	Dynamic	Static	Immediate Formative & Delayed Quiz
8/9	Static	Dynamic	Immediate Formative & Delayed Quiz
10	Dynamic	Static	Immediate Formative & Delayed Quiz
11	Static	Dynamic	Immediate Formative & Delayed Quiz
12	Dynamic	Static	Immediate Formative & Delayed Quiz
13	Midterm 2	Midterm 2	Midterm 2
14	Static	Dynamic	Immediate Formative & Delayed Quiz
15	Dynamic	Static	Delayed Quiz Only
16	Final Exam	Final Exam	Final Exam

Participants

Undergraduate students enrolled in Calculus I must meet one prerequisite (e.g., SAT score of 650 or higher). *Table 2* describes the distribution of the sample size based on the total number of students in both sections. A combined 217 students totaled at the start of the semester. Section 011 had a female to male ratio of 1.62 while Section 001 had a female to male ratio of 1.36. Overall the class had a female to male ratio of 2.36. Mathematic majors and non-majors could not be determined since many students had not yet declared majors. Students with declared majors come from a wide range of over 50 undergraduate disciplines.

Table 2

Participants

	Section 011	Section 001
Total N = 217	n=93	n=124
Male	35	36
Female	57	85
N/A	1	3

Study Procedures and Materials

Both sections of the course completed the same diagnostic, weekly online pre-work, weekly homework, weekly quizzes, mid-term and final exams. Both sections attended two 75-minute face-to-face classes throughout the semester. Both sections met on the same days during the week but at different times. Section 011 met in the afternoon while section 001 met in the morning. Additionally, students attended one 75-minute recitation period managed by teaching assistants who fielded questions from students as well as administered weekly paper quizzes.

The intervention focused on the format of the weekly online pre-work. Sections alternated each week the presentation format and assessed immediate and delayed performance measures for any differences. Performance measures included formative assessment questions presented after the online instructional presentation, DYNAMIC or STATIC, and delayed end-of-week quizzes.

The online lesson modules were required to be completed prior to the first live lecture demonstration class of the week. The interactive video format includes embedded practice questions, whereas for the printed text and image format, practice questions were referenced in the materials but presented at the end of the instructional content in a separate module to collect data and provide feedback. After completing each lesson module, students completed *additional* practice-assessment problems (formative assessments). Students were required to complete the online instructional material as well as the additional practice-assessment questions to receive participation credit for the whole lesson. Each formative assessment question had a point value of one although these points were not counted toward the overall grade except in terms of participation. Students received participation points for answering the questions. Students completed formative assessments in presented in eleven weeks over a sixteen-week semester period.

Students also completed the same survey question at the end of each online topical lesson. The question asked: On a scale of 1-5 (1 being the least effective and 5 being the most effective), how effective was the material to help you master the topic? The question aimed to collect student perceptions and attitudes to determine if there were any perceptual differences depending on the presentation format of pre-work materials.

After completing the online pre-class work students attend two 75-minute in-class lecture demonstrations during the week. At the end of the week students attend one 75-minute in-class recitation period managed by teaching assistants where a paper-based quiz is administered on the weekly instructional content. Students completed eleven total quizzes over a sixteen-week semester period. There was no end-of week quiz for week one.

MEASURES

We analyzed various measures of prior knowledge, student behaviors and performance in the course (*Table 3*). We provide definitions of the most important variables used in this study. *SAT/ACT Math score* is defined as either Scholastic Aptitude Test (SAT) out of 800 points or American College Test (ACT) standardized tests out of 36 points (converted to 800-point scale). *Current GPA* is the current Grade Point Average (out of 4 points). *Pretest Diagnostic* is defined as a score out of 42 from an algebra test administered at the beginning of the course was used a proxy for students' prior math ability (converted to a 0 to 100 percentage). *Online Lesson Module* is the average completion rate of online lesson modules (dynamic or static formats) completed prior to attending class (out of 48). *Formative Assess scores* is the average score across all immediate formative assessments (11, each with varying total points) in the course. *Quiz scores* is the average score across all the quizzes (11, each for 10 points) in the course. *Course* *grade* is a weighted sum of attendance and lesson modules (5%), WebAssign problem sets (5%), quizzes (10%), homework (10%), midterms (40%), and final (30%).

Table 3
Measures

Prior Knowledge	SAT/ACT Math		**Intervention Groups**	Section
	Current GPA		**Performance (Scores)**	Formative Assessment
	Diagnostic Pretest			Homework
Behaviors (Completion Rates)	Attendance			Quiz
	Online Lesson Modules			Midterm 1
	Formative Assessments			Midterm 2
	Homework			Final Exam
	Quizzes		**Final Outcome**	Course Grade

RESULTS

To compare differences between the two sections we performed *t-test* on behaviors and performance measures. *Analysis of covariance* (ANCOVA) was performed to compare sections on outcomes while controlling for prior differences. *Regression* was performed to compare total weekly pre-class practice-assessments and quiz scores for all students who used printed text and static images (STATIC) and all students using the interactive video modules (DYNAMIC). The primary outcome measures were pre-class practice-assessments (immediate) and weekly student performance measures namely average quiz score (delayed).

Prior Knowledge Abilities by Section

Descriptive statistics including mean and standard deviations (SD) show similarities and differences between the two groups in terms of prior knowledge abilities (*Table 4A*). Despite small average differences on prior knowledge levels, we note no statistically significant differences between the two groups in terms of prior knowledge indicators, SAT/ACT Math, GPA, or Pretest Diagnostic. There was no statistically significant difference on any prior ability (SAT/ACT Math, GPA, Pretest Diagnostic) between males and females in each section (*Appendix: Table 4B*). Since there were no differences between males and females on prior knowledge, we proceed with the analysis comparing sections.

Table 4A

Summary of Prior Knowledge Abilities (All students by Section)

		Section (001) (*n= 124*)	Section (011) (*n= 93*)	*t-test Differences*
Indicators	*Measure*	*Mean (SD)*	*Mean (SD)*	*t-value, p-value*
Student prior ability	SAT/ACT Math	659.35(73.42)	673.27(71.77)	-1.341, 0.182
	GPA	3.31(.39)	3.27(.53)	0.549, 0.584
	Pretest Diag. (%)	73.73(14.10)	71.12 (17.08)	1.123, 0.263

Behaviors by Section

As shown in *Table 4C*, both sections applied, on average, about the same effort to attend class and complete weekly quizzes. Attendance rates were collected via an audience response system, namely TopHat™, for both sections. Completion rates on the online lesson modules were collected through the university's learning management system. We note a statistically significant difference between the two sections in terms of weekly homework completion rates, section (001) (M=90.44, SD=15.96) and section (011) (M=81.34, SD=25.93), t (143) =2.99, p=0.003. Completion rates for end of week homework problems and quizzes were collected manually and entered in the university learning management system.

Table 4C

Summary of Behaviors (All students by Section)

		Section (001) (*n= 130*)	Section (011) (*n= 93*)	*t-test Differences*
Indicators	*Measure*	*Mean (SD)*	*Mean (SD)*	*t-value, p-value*
Student Behaviors	Attendance rate	4.44 (.871)	4.24 (.835)	1.687, 0.093
	Modules rate%	93.70 (17.23)	91.55(18.29)	0.864, 0.389
	Formative Assess rate %	95.35 (11.94)	93.33 (13.14)	1.180, 0.241
	Homework rate%	90.44 (15.96)	81.34(25.93)	2.99, **0.003***
	Quiz rate%	99.66 (1.87)	98.67 (4.95)	1.838, 0.069

Summary of Overall Performance by Section

Descriptive statistics including mean and standard deviations (SD) show similarities and differences between the two conditions based on performance measures (*Table 4D*). *Not controlling for any prior performance differences* or *testing for homogeneity of variance* among all undergraduate students at NYU participating in the Calculus I course sections (001) and (011), (N=223), there was no finding of a statistically significant difference between the two sections in terms of weekly quiz averages, midterm exams, final exam, or course grades. We note a statistically significant difference between the two sections in terms of weekly homework performance scores, section (001) (M=91.68, SD=13.41) and section (011) (M= 78.23, SD=24.55), t (130) =4.721, p=0.00.

Table 4D
Summary of Overall Performance (All students by Section)

		Section (001) (*n= 130*)	Section (011) (*n= 93*)	*t-test Differences*
Indicators	*Measure*	*Mean (SD)*	*Mean (SD)*	*t-value, p-value*
Student Performance	Formative Assess scores	64.44 (14.68)	64.76 (16.01)	-0.150, 0.879
	Homework scores	91.68 (13.41)	78.23 (24.55)	4.721, **0.00***
	Quiz scores	68.90 (19.68)	70.24 (21.55)	-0.465, 0.642
	Midterm1 scores	70.23 (18.01)	72.80 (19.71)	-0.983, 0.327
	Midterm2 scores	61.96 (19.51)	63.89 (23.42)	-0.629, 0.530
	Final scores	77.31(18.81)	74.62 (24.51)	0.874, 0.383
Final outcome	Course grade	74.44 (14.83)	72.62 (19.49)	0.739, 0.461

Course grade is a weighted sum of attendance and lesson modules (5%), WebAssign problem sets (5%), quizzes (10%), homework (10%), midterms (40%), and final (30%).

Summary of Performance by Presentation Format

A descriptive analysis was performed to assess the differences between the two presentation formats (*Table 4E*). The selected performance measures included weekly quizzes and formative assessments. *Not controlling for any prior performance differences* or *testing for homogeneity of variance* among all undergraduate students at NYU participating in the Calculus I course sections (001) and (011), (N=223), there was no finding of a statistically significant difference between students using either presentation format on immediate formative assessment scores or delayed quiz scores.

Table 4E

Summary of Performance by Presentation Format

		STATIC	DYNAMIC	*t-test Differences*
Indicators	*Measure*	*Mean (N, SD)*	*Mean (N, SD)*	*t-value, p-value*
Student Performance	Formative Assess scores	64.32 (1133, 25.43)	65.98 (1044, 24.33)	-1.556, 0.120
	Quiz scores	74.69 (1069, 25.39)	76.56 (1032, 24.22)	-1.724, 0.085

Start of Week: Immediate Formative Assessment ANCOVA

In eleven out of sixteen weeks, the students were required to complete formative assessment problems immediately after reviewing either presentation format covering relevant weekly topics. The mean for each week was analyzed for differences with and without prior knowledge scores (diagnostic, GPA, and SAT/ACT) as individual and grouped covariates. Only the week 04 results showed a statistically significant effect favoring the Dynamic presentation, F (1, 168) = 6.417, $p<0.05$. All other weeks demonstrated insignificant results controlling for prior knowledge scores as covariates.

End of Week: Delayed Quiz ANCOVA

In eleven out of sixteen weeks, the students were required to complete an end-of week quiz containing problem solving question items related to weekly topics. Differences in the mean for each section were analyzed using ANCOVA models (*Table 5A*) incorporating prior knowledge scores (diagnostic, GPA, and SAT/ACT math). Four of the eleven weeks demonstrated statistically significant ($p < 0.05$) differences, with 3 in favor of DYNAMIC presentations, 1 for STATIC presentations, and 7 with neutral results. Additionally, we found the effect size for each week demonstrating significance to be small. Neither presentation format appears to dominate the semester.

Table 5A
ANCOVA on Weekly Quiz

		STATIC	DYNAMIC	*Effect Size*
Indicators	*Measure*	*F-test*	*F-test*	r^2
Performance	Quiz Score [Week 03] Section 001		*F* (1, 166) = 6.696, p=0.011	0.040
	Quiz Score [Week 05] Section 001		*F* (1, 168) = 6.945, p=0.009	0.041
	Quiz Score [Week 11] Section 011		*F* (1, 156) = 9.642, p=0.002	0.062
	Quiz Score [Week 15] Section 011	*F* (1, 152) = 7.315, p=0.008		0.048

Regression on Aggregated Immediate Formative Assessment Model

A linear regression on aggregated immediate formative assessment scores analyzed scores with prior knowledge scores (diagnostic, SAT/ACT, and GPA) as individual and grouped control measures (*Table 5B*). The model's estimated mean difference between students using Static or Dynamic formats (0.0171 points higher for Dynamic) is not statistically significant (p=0.1092) while controlling for prior knowledge scores. The model explains about 11.4% of the variation in scores which was a statistically significant amount [p<.05] but not a meaningfully large overall effect.

Table 5B
Modeling Immediate Measures (N= 200)

Dependent Variable	*Source*	*Unstandardized B*	*Coefficients Std. Error*	*Standardized Coefficients*	*t value*	*p value*
Course Grade	Intercept	-0.2757	0.0671	NA	-4.1085	0.0000
	Dynamic	0.0171	0.0107	0.0351	1.6025	0.1092
	Diagnostic	0.0032	0.0004	0.1877	8.4247	0.0000
	GPA	0.1395	0.0147	0.2126	9.4990	0.0000
	SAT/ACT	0.0003	0.0001	0.0962	4.3375	0.0000

Linear Regression had an $R^2 = 0.114$ with 1852 degrees of freedom, which was statistically significant [*F* (4, 1852) = 59.609, p < 0.05].

Regression on Aggregated Quiz Model

A linear regression on aggregated immediate formative assessment scores analyzed scores with prior knowledge scores (diagnostic, SAT/ACT, and GPA) as individual and grouped control measures (*Table 4B*). The model's estimated mean difference between aggregate quiz scores between students using Static or Dynamic formats (0.106 points higher for Dynamic) is not statistically significant (p=0.326) while controlling for prior knowledge scores. The model explains about 11.8% of the variation in scores which was a statistically significant amount (p<.05) but not a meaningfully large overall effect.

Table 5B
Modeling Immediate Measures (N= 200)

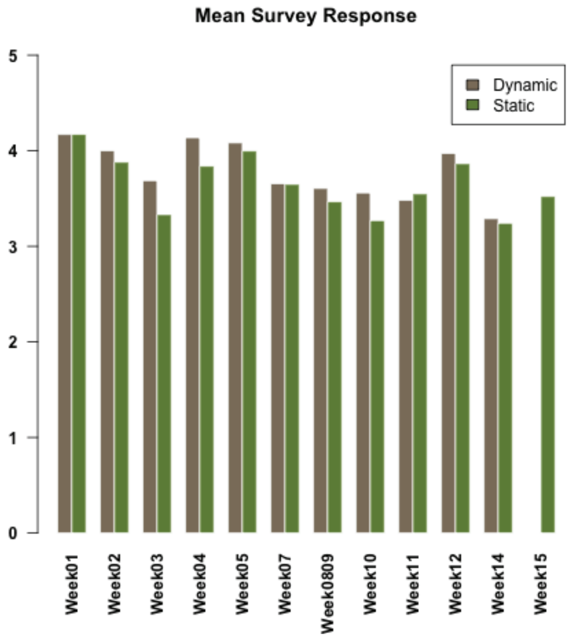
Dependent Variable	*Source*	*Unstandardized B*	*Coefficients Std. Error*	*Standardized Coefficients*	*t value*	*p value*
Course Grade	Intercept	-1.960	0.646	NA	-3.033	0.002
	Dynamic	0.106	0.107	0.022	0.983	0.326
	Diagnostic	0.009	0.004	0.052	2.310	0.021
	GPA	1.765	0.149	0.274	11.867	0.000
	SAT/ACT	0.004	0.001	0.130	5.749	0.000

Linear Regression had an $R^2 = 0.118$ with 1813 degrees of freedom, which was statistically significant [$F (4, 1813) = 60.75$, $p < 0.05$].

Survey: Student Perception on usefulness of pre-work materials

At the end of each topical section in each week students in both sections (001 and 011) were presented with the following survey question: *On a scale of 1-5 (1 being the least effective and 5 being the most effective), how effective was the material to help you master the topic?* We observed no difference in student perceptions regarding the instructional material whether presented as static or dynamic (*Chart 1*).

Chart 1
Mean Survey Response



CONCLUSIONS

Instructional content presented as printed text with static images (STATIC) or interactive video modules (DYNAMIC) assigned online as pre-work in the beginning of the week before class do not appear to impact outcomes on either immediate formative assessment or end of week quizzes. Either presentation format appears to produce about the same results. The few weeks that did show statistically significant effects were also varied based on the section by presentation format. Given the low r^2 values the results suggest that the presentation format of the same instructional content did not impact quiz scores in terms of weekly performances or aggregated overall. The survey results on student perceptions regarding the value of the instructional content demonstrated similar results not favoring one format over another.

IMPLICATIONS FOR PRACTICE & POLICY

The take-away for educational designers and instructors when planning the design of instructional materials is to consider the type of knowledge, learning processes, and learning outcomes targeted by the materials as well as the duration of learning time. Multimedia effects, despite the format, may diminish after practice, especially if the practice is retention-based (Butler, 2014). As educational designers know, effort placed on the design and development of materials varies given the format. Knowing where and when to put the most effort given the type of knowledge and learning outcome is important. Putting a lot of effort into the multimedia design and development of pre-work instructional materials aimed to support retention-based learning may not impact the learning outcomes as much as putting effort into designing other parts of the course that may have more impact to support goals targeting deeper understanding. Educational designers and media producers can still adhere to multimedia principles for developing content with less financial means, effort, and time. The pedagogical focus should shift from high-end content development to low-end content development to allow for more focus on innovative formative and summative assessment development. The results may also suggest different cognitive processing abilities at work that may be useful for different processing purposes depending on the learner's prior knowledge and needs. Hence both formats may be useful for the learner.

LIMITATIONS AND FUTURE RESEARCH

This study was conducted in two sections of a live course. The immediate formative assessment was not conducted in a controlled testing environment but online through the university learning management system. While the immediate formative questions were scored they only counted towards completion value points. Given the low stakes of the immediate formative questions it is unlikely that students worked together to solve these problems but we cannot account for students who may have worked together. Since the weekly quiz was a delayed outcome measure by several days the effects of the pre-work presentation formats may have diminished by other course events, namely the in-class activity sessions focused on the end of week homework assignment. The difficulty level of questions presented for the immediate and delayed assessments were easier than exam level items which may have affected the results. But given the low to moderate averages in both assessment types we find this consideration unlikely.

We have knowledge that students requested the printed text and image PDF documents for all weeks immediately prior to exams for study purposes. The documents were only provided to both sections post weekly quizzes so as to not contaminate the study. This suggests that the STATIC format had additional benefits beyond the weekly assignments. They were also useful study aids. The course now provides both presentation formats for students to use. Future studies in comparing presentation formats might compare an entire section using one format compared to another section using a different format for the entire semester. Additional research should be conducted on presentation formats for low ability students versus high ability students. Likewise, future research should also focus on deeper learning problems sets rather than knowledge retention.

Acknowledgements

A special thank you to the educational media designers Paula Torres and Carmen Louie for their hard work transforming this course from a traditional to flipped format and preparing materials for this study. They worked tirelessly to create the interactive video modules and stitch the course together. Lastly, a thank you to Dr. Jan Plass, Dr. Jennifer Hill, Dr. Ben Maddox and NYU IT leadership, and NYU's Provost for their support to not only enhance teaching and learning with pedagogically effective uses of technology and media but also to support these efforts with research to validate them.

APPENDIX

Additional Comparative Format Examples

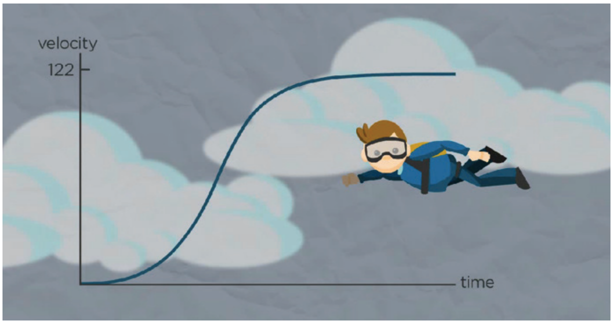


Figure 1. Interactive Video Presentation (DYNAMIC).

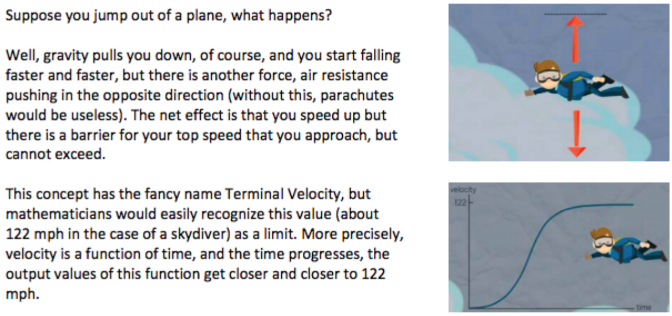


Figure 2. Text & Image Presentation (STATIC).

Table 4B
Summary of Prior Ability by Gender and Section

		Section (001) (*n= 124*)	Section (011) (*n= 93*)
Indicators	*Measure*	*Mean (N, SD)* *t-value (p≤0.05)*	*Mean (N, SD)* *t-value (p≤0.05)*
Student prior ability	SAT/ACT Math	M = 668.4 (35, 70.2) F = 655.4 (81, 74.9) *t* = 0.90, 0.373	M = 688.0 (30, 73) F = 665.1 (54, 72.7) *t* = 1.43, 0.157
	GPA	M = 3.3 (36, 0.4) F = 3.3 (85,0.4) *t* = -0.40, 0.687	M = 3.2 (35, 0.6) F = 3.3 (57,0.5) *t* = -0.84, 0.401
	Pretest Diagnostic	M = 69.8 (32, 18.1) F = 75.1 (81,12.1) *t* = 1.54, 0.132	M = 71.1 (30, 15.7) F = 71.1 (50,18.0) *t* = -0.01, 0.992

References

Adams, C., & Dove, A. (2018). Calculus students flipped out: The impact of flipped learning on Calculus students' achievement and perceptions of learning. *PRIMUS*, *28*(6), 600-615.

Ainsworth, S., & VanLabeke, N. (2004). Multiple forms of dynamic representation. *Learning and Instruction*, 14, 241-255.

Atkinson, R.K. (2002). Optimizing learning from examples using animated pedagogical agents. *Journal of Educational Psychology*, 94, 416-427.

Atkinson, R. K. (2005). Multimedia learning of mathematics. *Cambridge handbook of multimedia learning*, 393-408.

Betrancourt, M., Morrison, J., & Tversky, B. (2002). Animation: can it facilitate? *International Journal of Human-Computer Studies*, 57, 247-262.

Betrancourt, M., & Tversky, B. (2000). Effect of computer animation on users' performance: a review. *Le Travail Humain*, 63(4), 311-329.

Boucheix, J.-M., & Schneider, E. (2009). Static and animated presentations in learning dynamic mechanical systems. *Learning and Instruction*, 19, 112-127.

Cooper, G., & Sweller, J. (1987). The effect of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology*, 79, 347-362.

Clark, R., & Mayer, R. (2011). E-learning and the science of instruction. (3 ed.). San Francisco: Pfeiffer.

Clark, R. C., & Mayer, R. E. (2016). *E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning*. John Wiley & Sons.

GebreYohannes, H., Bhatti, A. H., & Hasan, R. (2016). Impact of multimedia in Teaching Mathematics. *International Journal of Mathematics Trends and Technology*, *39*(1), 80-83.

Giannakos, M. N., Krogstie, J., & Sampson, D. (2018). Putting Flipped Classroom into Practice: A Comprehensive Review of Empirical Research. In *Digital Technologies: Sustainable Innovations for Improving Teaching and Learning* (pp. 27-44). Springer, Cham.

Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction*, 15, 313-331.

Gross, D., Pietri, E. S., Anderson, G., Moyano-Camihort, K., & Graham, M. J. (2015). Increased preclass preparation underlies student outcome improvement in the flipped classroom. *CBE—Life Sciences Education*, *14*(4), ar36.

Guerrero, S., Beal, M., Lamb, C., Sonderegger, D., & Baumgartel, D. (2015). Flipping undergraduate finite mathematics: Findings and implications. *PRIMUS*, *25*(9-10), 814-832.

Hoffler, T. N., & Leutner, D. (2007). Instructional animation versus static pictures: a meta-analysis. *Learning and Instruction*, 17, 722-738.

Jeung, H., Chandler, P., & Sweller, J. (1997). The role of visual indicators in dual sensory mode instruction. *Journal of Educational Psychology, 17, 329-433.*

Kombartzky, U., Ploetzner, R., Schlag, S., & Metz, B. (2010). Developing and evaluating a strategy for learning from animations. *Learning and Instruction*, *20*(5), 424-433.

König, A., Stadler, M., Klepsch, M., & Seufert, T. (2012). The Effects of Visualization Forms on Usability and Learning Outcomes–Dynamic Videos versus Static Pictures. *STAGING KNOWLEDGE AND EXPERIENCE: HOW TO TAKE ADVANTAGE OF REPRESENTATIONAL TECHNOLOGIES IN EDUCATION AND TRAINING?* 103.

Lage, M. J., Platt, G. J., & Treglia, M. (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. *The Journal of Economic Education*, *31*(1), 30-43.

Low, R., & Sweller, J. (2014). The modality principle in multimedia learning. *The Cambridge handbook of multimedia learning*, *2*, 227-246.

Lowe, R. K. (1999). Extracting information from an animation during complex visual learning. *European Journal of Psychology of Education*, 14(2), 225-244.

Lowe, R. K. (2003). Animation and learning: selective processing of information in dynamic graphics. *Learning and Instruction*, 13, 157e176.

Lowe, R. K. (2004). Interrogation of a dynamic visualization during learning. *Learning and Instruction*, 14, 257-274.

Lowe, R. K. (2008). Learning from animation: where to look, when to look. In R. K. Lowe, & W. Schnotz (Eds.), Learning with animation: Research implications for design (pp. 49e68). NewYork: Cambridge University Press.

Lowe, R. K., & Schnotz, W., (Eds.). (2008). Learning with animation: Research implications for design. New York: Cambridge University Press.

Mayer, R. E., (2003). The promise of multimedia learning: Using the same instructional design methods across different media. *Learning and Instructio*n, 13, 125-139.

Mayer, R. E. (Ed.). (2005). *The Cambridge handbook of multimedia learning*. Cambridge university press.

Mayer, R. E., & Anderson, R. (1991). Animations need narrations: An experimental test of dual-coding hypothesis. *Journal of Educational Psychology, 83, 484-490.*

Mayer, R. E., & Anderson, R. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology,* 93, 444-452.

Mayer, R. E., & Moreno, R. (2002). Animation as an aid to multimedia learning. *Educational Psychology Review*, 14, 87-99.

Mayer, R. E., & Moreno, R. (1999). Cognitive principles of multimedia learning: the role of modality and contiguity. *Journal of Educational Psychology*, 91, 358-368.

Mayer, R. E., & Heiser, J., & Lonn, S. (2001). Cognitive restraints on multimedia learning: When processing more material results in less understanding. *Journal of Educational Psychology*, 93, 187-198.

McGivney-Burelle, J., & Xue, F. (2013). Flipping calculus. *Primus*, *23*(5), 477-486.

Moreno, R., Ozogul, G., & Reisslein, M. (2011). Teaching with concrete and abstract visual representations: Effects on students' problem solving, problem representations, and learning perceptions. *Journal of Educational Psychology*, *103*(1), 32.

Nathan, M. J., Kintsch, W., & Young, E. Y., (1992). A theory of algebra-word-problems comprehension and its implications for the design of learning environments. *Cognition and Instruction*, 9, 329-389.

Pincus, H., Wojcieszak, M., & Boomgarden, H. (2017). Do multimedia matter? Cognitive and affective effects of embedded multimedia journalism. *Journalism & Mass Communication Quarterly*, *94*(3), 747-771.

Scheiter, K., Gerjets, P., & Schuh, J. (2010). The acquisition of problem-solving skills in mathematics: How animations can aid understanding of structural problem features and solution procedures. *Instructional Science*, *38*(5), 487-502.

Schnotz, W., & Lowe, R. K. (2008). A unified view of learning from animated and static graphics. In R. K. Lowe, & W. Schnotz (Eds.), *Learning with animation: Research implications for design* (pp. 304-356). New York: Cambridge University Press.

Schnotz, W. (2005). An integrated model of text and picture comprehension. *The Cambridge handbook of multimedia learning*, *49*, 69.

Schwan, S., & Riempp, R. (2004). The cognitive benefits of interactive videos: learning to tie nautical knots. *Learning and Instruction*, 14, 293-305.

Skuballa, I. T., Dammert, A., & Renkl, A. (2018). Two kinds of meaningful multimedia learning: Is cognitive activity alone as good as combined behavioral and cognitive activity? *Learning and Instruction*, *54*, 35-46.

Sweller, J., & Cooper, G. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction*, 2, 59-89.

Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load as a factor in the structuring of technical material. *Journal of Experimental Psychology: General,* 119, 176-192.

Tamizi, R., & Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology*, 80, 424-436.

Watson, G., Butterfield, J., Curran, R., & Craig, C. (2010). Do dynamic work instructions provide an advantage over static instructions in a small-scale assembly task? *Learning and Instruction*, *20*(1), 84-93.