Abstract

Abstract: This paper summarizes research conducted on three computer-based video models’ effectiveness for learning based on memory and comprehension. In this quantitative study, a two-minute video presentation was created and played back in three different types of media players, for a sample of eighty-seven college freshman. The three players evaluated include a standard QuickTime video/audio player, a QuickTime player with embedded triggers that launched HTML-based study guide pages, and a Macromedia Flash-based video/audio player with a text field, with user activated links to the study guides as well as other interactive on-line resources. An assumption guiding this study was that the enhanced designs presenting different types of related information would reinforce the material and produce better comprehension and retention. However, findings indicate that
the standard video player was the most effective overall, which suggests that media designs able to control the focus of a learner’s attention to one specific stream of information, a single-stream focused approach, may be the most effective way to present media-based content.

Résumé: Cet article résume une étude vérifiant l’efficacité de l’apprentissage basé sur la mémorisation et la compréhension, conduite à partir de trois modèles basés sur la vidéo informatisée. Dans cette étude quantitative, une vidéo de deux minutes a été créée et lue sur trois types de lecteurs différents, pour un échantillon de 87 étudiants universitaires de première année. Les trois lecteurs évalués comprenaient un lecteur standard audio/vidéo Quicktime, un lecteur Quicktime avec déclencheurs intégrés qui lançait un guide d’étude en HTML, et un lecteur audio/vidéo Flash Macromedia avec un champ texte, comprenant des liens activés par l’usager vers des guides d’étude et d’autres ressources interactives en ligne. Une supposition guidant cette étude était que les designs enrichis présentant différents types d’informations interreliées renforceraient le matériel et produiraient une meilleure compréhension et une meilleure rétention. Cependant, les résultats indiquent que le lecteur vidéo standard était le plus efficace, ce qui suggère que les designs de médias concentrant l’attention de l’apprenant sur une source d’information spécifique seraient la meilleure façon de présenter du contenu médiatisé.
Introduction

As different forms of media converge via newly developed technologies, the development and use of multimedia designs that incorporate video playback as a major component have been adopted by educational multimedia creators. There is a significant body of video-based research that suggests that cognitive processing, both conscious and unconscious, is largely bypassed during the act of screening video (largely based on studies of television viewing). According to visual theorist Fred Barry and others, video and audio playback tap into the emotional part of the brain, and viewers become emotionally, but not logically, involved with the medium (Barry, 1997; Halloran, 1970). Emotional versus logical involvement may occur because as viewers watch television, their brains move into a hypnotic alpha rhythm stage, much like daydreaming, even after watching television for as little as thirty seconds (Barry, 1997). As viewers become relaxed, the left side of their brain, which tends to process information and is more analytical, becomes inactive, allowing the right side of their brain to process information both uncritically and emotionally (Dixon, 1981). In such a state, viewers are susceptible to emotional manipulation just when they are forming ideas and attitudes about the viewed material and content (Barry, 1997). Krugman (1970) argues that while the brain’s response to print is characterized by fast brain waves and mental activity, the brain’s response to television can be described as passive and is largely made up of slow brain
waves. What can be concluded from this information is that viewers tend to become emotionally involved with television and that images stream into the psyche bypassing the logical part of the brain (Krugman, 1970).

Much of the existing research on multimedia learning—which may include the use of text, spoken word, video, audio, animation, and graphics—appears to be performed only on a few specific types of media designs, such as the combination of the spoken word and animation. The author anticipated some interesting and insightful results from a comparative study of three different multimedia designs that shared the same content, focused on video playback, and included the latest innovations for video streaming such as embedded triggers in streaming video.

Theoretical Perspective

The theoretical perspective for this study is found in Mayer’s (2001) well-established cognitive theory of multimedia learning. Mayer’s work is based on a learner-centered approach, which begins with an understanding of how human cognition works and focuses on the use of multimedia to enhance human learning. Some of the key findings and techniques leading to effective practice and models that relate to the present study include:

• Spoken narration combined with an onscreen visual guide that does not split the attention of the learner, but in fact can
enhance the experience in certain instances.

- Sudden onset of pictures and animation is more effective for learning than static pictures alone, presumably by directing the learner’s attention and focus to specific elements in the visual display.

- The sudden onset of having visual elements appear produces the same learning enhancements as an animated presentation. Thus, the procedure of flashing appropriate parts of the pictorial information as they are described in the spoken narrative is as effective as a full animation. (Craig, Gholson, & Driscoll, 2002)

Researchers have focused on how various design features relate to human information processing, such as comparing and testing designs that place a light load versus a heavy load on a learner’s visual processing channel. Allan Pavio (1986), for example, developed a human cognitive theory termed “dual coding”, in which he makes a case that the human brain is divided into two cognitive subsystems, one for processing nonverbal objects and information, and one for processing language. The research done by both Mayer and Pavio provide a theoretical base for the design of the present study in which the same video content is presented using three different models and tested for learning effectiveness.

Purpose of Study
Video and other media forms in multimedia environments can be effective for learning, especially in designs aimed at a constructivist notion of effective learning. Research has been done using specific designs and components and comparing multimedia to traditional learning approaches. More studies that test specific variables and components of similar designs that are conceived based on the same video/audio playback component, but presented in different forms to teach the same subject matter, are needed. The present study was designed to help address this specific knowledge gap.

This type of research is important as different forms of computer media, like streaming QuickTime video, now offers designers the ability to quickly create multimedia environments within these specific formats (by embedding triggers in the video stream itself). This technology enables the content creator, such as a video editor, to create a multimedia playback environment, which is a fundamental shift. Typically the process of post production where content is created through video and audio editing has been limited to a single screen, but these creators can now include other media components such as pop-up windows in their work. While designing the multimedia models for this study, the author ensured that one model would utilize steaming QuickTime video with embedded triggers that launched pop-up study guide windows during the presentation.
A key goal in this study was to determine if increased interactivity is more effective for learning (based on measures of comprehension and retention) when compared to similar models with less interactivity. Based on previous research and media design theories, specific study questions emerged, such as: (1) Is a video that bypasses the analytical part of the brain, but connects to the emotional part, a better or worse teaching tool?; (2) How much interactivity combined with that same video is too much?; and, (3) How do different types of media designs and components compare within the framework of teaching a single subject?

The present study is based on the testing of specific variables and components of different designs that attempt to teach the same subject matter, which in this case is the non-traditional content of “Dreams and Dream Interpretation.” Each model shares the same video and audio clip, but uses different presentation forms. It is hypothesized that testing three different multimedia designs all based on the same content, and specifically sharing the same video component, should help to identify which specific media components enhance, or detract from, the learning processes as measured by tests of comprehension and retention of specific content.

Research Questions and Hypothesis

The central research questions for this study are: “How well do each of the three multimedia designs perform as learning tools?”, and “How do these three different designs compare
to each other when presenting the same content?” The goal is to study the effectiveness of specific types of multimedia elements and designs by testing subjects for both retention and comprehension from the presentation. The researcher’s first hypothesis is that the video model with a single screen of video playback (Model B) would emerge as the most effective overall. The researcher’s second hypothesis is that Model B, that combined single screen video playback with the embedded triggers in an enhanced version, would provide a significant increase in retention and comprehension. This second hypothesis is based on the assumption that the learner’s connection with a traditional video segment is largely emotional / affective, and would improve when combined with the carefully placed pop-up study guides that reinforce the material being presented.

Research Design

The research project used quantitative methods to evaluate the three different multimedia models in a series of controlled study sessions in a computer lab environment. In each session, only one of the three models was presented to a group of university undergraduate communication students. These participants all shared similar academic achievement scores, a high level of computer literacy, and a declared major in communications, as well as belonging to the same early-twenties age group. The controlled environment included headphones for each student so as not to be
distracted by other participants, and teaching assistants who guided each participant to their workstation and closely monitored the lab (Figure 1).

Each model’s design and effectiveness was tested and measured by a single common method based on a written test, measuring each learner’s retention and understanding of the material presented during a session of one of the three models. This test was evaluated and refined during the pre-study, and the final version had ten questions that measured retention, and four of the questions also measured comprehension of basic concepts presented in the material. Each participant’s model was pre-selected based on the study’s need to create sample groups of the same size for each model.

Figure 1. Study session in computer lab

Model A. Traditional Linear Video
A 2-minute segment of video and audio was played back via a QuickTime movie window on the computer screen. The segment was an edited version of an interview with an expert on the subject of dreams and dream interpretation. It included an on-camera interview, audio by the interviewee, and a series of related graphic slides that broke up the segment and reinforced the concepts discussed. The viewers had the ability to control playback functions via the standard QuickTime player controls, such as start, pause, rewind, and play similar to a VCR (Figure 2).

The multimedia learning issue raised by Model A is how effective traditional linear video is for learning and comprehension when compared to more enhanced media forms (Models B and C). As previously stated, it is widely believed that through the act of experiencing a video playback similar to a television, the cognitive processing that takes place can be very powerful, since it tends to be more of an emotional experience rather than an analytical one. Television has a perceived credibility that helps the viewer to believe in the content being viewed.
Figure 2. Traditional Linear Video Stream Model A: The QuickTime movie was played back in the viewer on the desktop sized to 320 X 240 pixels. The video screen could enlarge to twice its normal size before suffering from pixilated video.

Model B. Directed Linear Video

Model B used the same video segment as the first model; a QuickTime movie was played back in the same format and used the same player. However, Model B included the major enhancement of embedded triggers in the video stream. These carefully placed triggers opened two separate static study guide windows at specifically timed points during the video stream, and included Java script coded window consoles for the location of the pop up windows on the screen. These windows were carefully designed enhancements that incorporated text and graphics relating to the content in the
The multimedia learning issue raised by this model is whether the enhancement of triggered timed events in a traditional video playback environment is more effective overall (first hypothesis). Since the viewer was forced to use both the analytical and emotional parts of the brain, do the triggered study guides increase learning and retention by enhancing the powerful emotional experience of viewing video? The second hypothesis is that the specific combination in Model B would provide for increased retention and comprehension (over the other two models).

Figure 3.

Directed Linear Video Model B: The QuickTime movie was played back in the browser window. This screen shows the two triggered browser content pages to the right.

Model C. Non-Linear User-Controlled Video
Model C used the same video and audio segment as its content, but was played back in a custom viewer created with Macromedia’s Flash MX software. The actual video playback window is similar to the QuickTime players used in Models 1 and 2, but added a text field under the video screen that ran in sync with the video (similar to closed-captioning on a television set). Links placed on the side of the video player allowed viewers to access the same study guide windows as in the second model, as well as related web pages that included interactive dream dictionaries (Figure 4).

The multimedia learning issue raised by this more user-controlled model was whether or not the textual enhancements impact learning and retention. Many theorists, including Mayer (2001) and Craig et al. (2002), believe that when content is presented in a range of media forms, and when the users can control their experiences, a deeper level of learning can occur. In contrast, the author anticipated that this model would be the least effective, and believed that the user experience would become unfocused based on two factors. First, the viewers would not connect emotionally to the content, and second, their focus and attention would be fragmented by the video, audio, and text all being presented simultaneously.
Participants

The study was performed at a competitive private university in the mid-Atlantic region of the United States. The pool of participants was made up of freshmen and sophomore students currently enrolled as Communications majors at the university. This group shared similar levels of education and scholastic aptitude test scores from 1180 to 1350. Each model was tested with a sample group of 20 to 25 students, randomly selected from a total pool of 87 students, in order to further ensure comparability. The controlled environment included headphones preset at stations for each student to use for audio playback (see Figure 1). Each participant tested one
model only, either A, B or C, and once a participant had completed his or her respective model, he or she was not allowed to test another model.

Data Collection

Participants filled out a brief survey before they were tested about their respective model. This information allowed the collection of relevant data on each subject, and included questions regarding each participant’s computer experience, learning style, television viewing habits, and other related issues regarding the study. This survey data was used to refine the analysis of the study findings presented in another study, which compared different categories such as gender, computer experience, self-proclaimed learning styles, and television viewing habits. The final survey question asked if the participant had any previous experience with the content being presented. If this final question was answered yes, the data from that individual was excluded from the study (see Appendix A for complete survey).

Once the participants were properly set up at their respective viewing stations, they viewed their pre-selected model. They then took the written test during which they could not gain access to any of the information from the presentation. The test was the primary data-gathering tool for the model’s effectiveness (see Appendix B for written test).
In November 2004 a series of test sessions were conducted to work out various logistics of running the study, evaluate and refine the written test as well as data collection practices, and create a standard set of procedures and guidelines. Then in February 2005, six separate study sessions were facilitated, gathering data from a total of 87 participants spread evenly over the three models. From the data collected, a database was created that included each participant’s survey information, and their accompanying test scores. From that database, a series of 17 spreadsheets of the participants’ self-proclaimed learning style categories, coupled with the test scores calculated by percentage of correct answers with a margin of error .05%, was created. It is from that data and spreadsheet series that the following study findings and charts have been created. The few participants that had technical problems, or did not fit into the study demographic sample, were not included in the study findings.

Findings

Model A represents the Linear Video Model, based on the video and audio playback of the 2-minute segment on dreams. Model B is the Directed Linear Model, with the same dreams video and audio segment played back on the computer, but also includes a series of embedded triggers in the video stream that open study guide pages at specific times during the playback. Model C is the Non-Linear User Controlled Video Model, where the video and audio is played back in a model that includes a text field, and a series of links are provided to
various interactive websites and study guides. The \((n = 87)\) refers to the total number of participants whose data was collected in the study.

Figure 5 shows the overall scores for each computer-based video model, with Q1 through Q10 denoting the ten questions used to determine each model’s effectiveness. In seven of the ten study questions, Model A either tied or surpassed the scores of the other two models overall. The exceptions to this trend were Model B in question 5, with a slight lead over the other models, and Model B in question 9 with a significant lead over the other models. It is also interesting to note at this point, the interactive multimedia Model C lagged behind in eight out of the ten questions.

![Overall Scores](image)

**Figure 5. Models A, B and C overall scores**
After reviewing the study’s overall data, the findings indicate that author’s hypothesis #1, which stated model A would be more effective than model C, was indeed supported by the study’s findings (t-statistic = 2.498, p-val = 0.0154). But the author’s hypothesis #2, which stated that the additional pop-up study guides would increase the effectiveness of Model B, was not supported by the study findings (t-statistic = 0.0872, p-val = 0.9308).

It becomes clear in this study that the model solely based on traditional video playback, Model A, appears to be at least as effective as Model B and more effective than C. So in the end the simple act of experiencing content presented on a single screen provided the best results for retention and comprehension, while the other components added to that single screen model lessens the effectiveness with few exceptions. Important to note here is that retention is more widely measured from the written test, while questions measuring comprehension was limited. The more sophisticated Models B and C, in almost every case, are seen to be not only less effective, but also appear to be linked with decreased retention and comprehension, apparently by overloading or distracting the viewer.

An analysis of the data across several demographic dimensions did not yield statistically significant differences (p=0.05). The data were separated according to gender and self-reported preferred learning styles, with no significant
differences were found [3].

Overall, the lack of improvement in both Models B and C could be linked to Pavio’s dual coding cognitive theory, where the additional features of these models apparently overloaded one of the two types of brain processing, in this case the ability to process language, as most of the extra information provided is text-based (Pavio, 1986). Similar results have occurred in other studies, as researchers at the Educational Technology Expertise Centre at the Open University of the Netherlands conducted a multimedia instructional study and found replacing visual text with spoken text resulted in lower retention scores (Tabbers, Martens & Merrienboer, 2004).

At the same time, this initial data seems to conflict with parts of Mayer’s research, specifically with his findings that the sudden “onset” of pictures and animation (such as “pop-up” windows) prove to be more effective for learning and produces the same learning enhancements as an animated presentation.

![Overall Scores Compared](image)
Looking further at how each model scored overall, it is apparent in Figure 6 that Model A scores higher for questions 1, 2, 3, 7, 8 and 10. Model B comes in as the second most effective with higher scores for question 5 and 9. All three models tie on questions 4 and 6. This data implies that Model A may be the most effective overall, with Models B and C both coming in a close second, producing similar results for all but question 9.

It is interesting to note that all three models follow a similar pattern until question 9, when both Models A and C decrease in scores, while Model B scores rise significantly. Question 9 is based on the retention of certain objects presented in the video. In this instance Model B’s triggered study guides that appeared in sync with the video playback indicate a dramatic increase in retaining content from memory. Since this study guide uses text combined with images, perhaps the non-verbal cognitive processing channel, as discussed in Pavio’s dual coding theory, helps the learner to retain this information.

These initial findings lead to more questions than definitive answers. Why does the simple video Model A appear to be successful overall? One reason might be that given the content and amount of information presented, Model A’s simple linear design is more directly related to the ten assessment questions, while the enhanced designs of Models
B and C, which present additional forms of information, fail to increase the retention and comprehension of the shared content presented in the video, resulting in lower scores overall.

The makeup of the study’s sample population should be taken into consideration, as the demographic makeup of the participants may skew the data in different ways. The majority of the study subjects were women at 61%, versus the men at 39%. This leads the author to question whether more male subjects in the study might affect the results for each model.

Age may also have been a factor in these results. It would be interesting to run the same study and designs with younger participants. Perhaps younger students are more adept at multitasking on a computer with several windows and applications open at once, and hence the more complicated models might prove to be more effective with this group. Or perhaps younger students are not so adept at multitasking, but actually are “skimming”—letting their attention and focus bounce from multiple tasks. Further research may yield some answers about age-related use patterns.

Another factor that may have impacted this study’s findings is the environment in which the sessions took place. The open lab with two rows of workstations may have a negative impact on any of these models. Does a lab environment create a certain level of distraction, or, possibly, discomfort? A future study might probe what would happen if the participants
tested their respective model on their home computer.

Conclusion

For most of the questions based on memory and/or comprehension of this specific content, this study found that a standard video segment in a single screen, with spoken word, audio and graphics slides (Model A) appears to be the most effective overall, which supports hypothesis #1 predicting that the video model with a single screen of video playback would be most effective overall. While the results of Model B produced a significant boost in retention and comprehension in one of the ten questions, it still was not effective enough to support or disprove the author’s hypothesis #2, which theorized that the embedded triggers in the enhanced model would provide a significant boost in retention and comprehension. Model B’s results were surprising and differ from the conclusions of previous research of both Mayer (2001) and Pavio (1986).

This study’s findings suggest that a media design that is able to control the focus of a learner’s attention to one specific stream of information, a single-stream focused approach, may be the most effective way to present media-based content about dreams. It is worth noting that given the limited scope and parameters of this study, more research is needed in order to better understand the present findings on a single stream focused approach to learning.
One possible modification to the present study is an enhanced single-stream focused approach with the use of embedded triggered windows in a single window of information that compliments the content. This modified single-stream focused design reinforces Mayer’s (2001) theory that spoken narration combined with an on-screen visual guide does not split the attention of the learner, but in fact can enhance the experience in certain instances. It may be interesting to change both the frequency and nature of the triggered pop-up study guides in future studies. Among the study guide design changes could be the presentation of visuals only, rather than text and visual combinations that were used in this study, as well as increasing the frequency of the guides by incorporating more triggers in future models.

An interesting and useful variation of this study could also include the addition of an audio-only segment model, which may indicate how much of the content is absorbed, understood, and retained from simply listening to the spoken word / audio. Another useful design change could be to recreate the same model designs but with different content, such as a history or science segment. Also the study demographic pool is a variable that could easily be changed. Conducting the same study with younger students, and/or older participants, could show trends of how various aspects of the three models perform with sample groups outside this study’s limited demographic pool.
Model C scored the lowest overall, and was also the most complex. The design was based on the same video segment, including a text window in sync with the spoken voice, and provided interactive links to study guides as well as other interactive components. This model’s apparent ineffectiveness is likely due to an overload of the learner’s cognitive processing abilities, presumably overwhelming the learner’s visual processing channel. Model C’s reduced scores for memory and comprehension may be caused by the distraction of the participant’s focus due to the model’s design. The idea that our mind’s processing channels can be overloaded and lose much of their functionality is not new. Much research has been done on cognitive load theory, based on the assumption that the mind’s working memory has limitations. Consequently, results as seen in Model C would be typical (Mayer & Anderson, 1991).

One conclusion drawn from this study is that a well-produced video segment can be a powerful learning tool that provides users with a rich and rewarding experience when incorporated in the right multimedia design. As new technology influences educational multimedia designs, what will remain important is the ability to control the learner’s or end-user’s focus. When watching a video, either on a television or computer, audience reaction is directly linked to how people are engaged during the event. A successful multimedia model using video as the primary design component is one that creates an environment that connects the content with the viewer on an emotional
level, includes related information woven into the experience in a visual way, and is careful not to overload the viewer with too much information.

References

Glossary

Multimedia: A content presentation form that augments written text with the spoken word, other audio sources, animation, pictures, and video, all of which utilizes different levels of interactivity.

Metatrack: The track in a computer video stream that incorporates triggered events.

Metadata: The computer software component that creates a triggered event.

Embedded Triggers: The ability to have events placed into a stream of video, such as the launching of a browser, opening a website, opening a graphic, etc.

Linear: A traditional segment of video that has a set starting point, plays back in real-time, and then ends.

Enhanced Linear: A traditional linear segment of video that includes triggered events in the video stream.

Non-Linear: A model based on the ability of a user to choose various media components, and hence create a unique individual experience, which may take on aspects of a traditional linear fashion, but incorporates user control features so the experience will most likely not follow a clear linear path.
Note: The Flash-based computer video model C used in this study is available for free via the links page on his web site [www.davidreiss.com]. The site also contains media samples and other related work.

Appendix A: Survey

DREAMS Computer Video Study

Participant Name:

Age (circle one) - M / F

Major

(circle one) - Freshman Sophomore Junior Senior

Workstation: Date:

Model: (circle one) - A B C

SURVEY QUESTIONS:

Computer Experience & Knowledge: (circle one for each type)

Windows: none limited comfortable proficient

Apple/MAC: none limited comfortable proficient

Learning Style/Preferences:
What type of information do you learn best from? (please circle ALL that apply):

- Reading
- Writing
- Visual
- Lectures
- Class Discussions
- Hands-on workshops
- Traditional TV programs such as documentaries
- Computer-based video & audio
- Animations
- Static Visual & Graphic presentations like PowerPoint Presentations
- Interactive Computer-based multi-media
- Other

Television Viewing for Entertainment per WEEK (circle one)

0-5hrs 5-10hrs 10-15hrs 15-20hrs Over 20 hrs

Television Viewing for Learning per WEEK (circle one)
Appendix B: Written Test

PLEASE ANSWER THE FOLLOWING QUESTIONS FROM MEMORY

1. What part of our mind is working when we dream? (Circle One)

2. What is the basic ingredient of dreams? (Circle One)
   a. Desires b. Day’s Events c. Past Events d. Imagination

3. How can dreams give you insight or helpful information? (Circle all that apply)
   • Looking at the characters in your dream and relating back to your life.
   • Looking at the theme of your dream and relating back to your life.
• Asking a question before you go to sleep.

• Discussing what you want to dream about with a friend.

• 4. What is the language of the brain? (Circle One)
  

5. How is contrast used when discussing dreams? (Circle all that apply)

• Compare similar things.

• Compare different things.

• Arrange things to show similarities.

• Arrange things to show differences.

6. How is the term “correlate” used when discussing dreams in this video? (Circle one)

• To gather and compare unrelated things.

• To gather and compare related things.

• To gather and compare different things.

• To gather and compare similar things.
7. What does our mind try to do when dreaming? (Circle all that apply)

- Escape our day’s events.
- Make sense of our day’s events.
- Attempt to predict the future.
- Relive past experiences.
- Re-experience that day’s events.

8. When you program your dreams what helpful things can they do for you? (Circle all that apply)

- They can help you find a job.
- They can you learn new things.
- They can help solve a problem.
- They can help you answer questions you have in your life.

9. Identify some of the specific symbols presented in the video portion only. (Circle all that apply)

- Grass
10. If you can’t remember the specifics of a dream, what else can give you useful information? (Circle all that apply)

- Nothing
- Location
- Plot
- Theme
- Colours