Malaysian Online Journal of Instructional Technology (MOJIT) Vol.. 1, No. 1, pp 20-29
August 2004
ISSN: 1823-1144

Still Pictures and Audio: Second Class Multimedia Elements?

Wan Mohd. Fauzy Wan Ismail
Center of Instructional Technology and Multimedia
Universiti Sains Malaysia, Malaysia
fauzy@usm.my

Abstract

Broadband is now being presented and discussed at great length as it is presented as a popular solution for delivering multimedia materials. Presently in Malaysia, the infrastructure to support the use of high quality video streaming, high definition animation and immersive environment is limited and very expensive. Numerous educational systems are often hard pressed to cater to the high bandwidth required for delivering multimedia courseware materials. Often overlooked are the simple still graphics and mid quality mono audio signal which would require a much lower bandwidth to access or deliver. The end result of the use of these "lower" level multimedia elements would be similar in terms of information delivery and processing when compared to their "high" level counterparts. This paper will examine the benefits of this "lower" level multimedia in terms of information dissemination and processing. The Pavio's Dual Coding Theory and the Mayer's Multimedia Learning based on the Sweller's Cognitive Load Theory will be revisited in relation to the use of still pictures accompanied with audio signals. It is hoped that educators and institutions without access to broadband facilities will be encouraged to present their course content and information through a viable, practical and low bandwidth approach.

INTRODUCTION

One of the major goals of the Malaysian education system is to inculcate the desire to be a life long learner and this has emerged by the increasing number of working adults who are keen on pursuing further education. The goals may differ from one individual to another but the commonality among them is the desire to learn more and updated and upgrade their knowledge. Technology today has made this goal attainable through various communications medium ranging from simple paper-based correspondence to high quality real time interactive video conferencing. Most of the education systems in Malaysia are now opting for interconnectivity through various networking systems either supported by the Ministry of Education (MOE) or through their personal operating budgets. The connectivity may be through a dial up modem facility or through dedicated lines. Educators are now being inundated with marketing promotions for broadband facilities either wired or wireless. The main themes of these advertisements are broader bandwidth to allow faster access for high quality digital materials.

Often forgotten are the simpler things in life and this applies to all areas. In the area of education, we should not overlook the fact that more may not necessarily be better. Developers of instructional materials often overlook the technical requirements of a learner to access their materials. The typical learner may not have the latest software version or have access to a high bandwidth network. Though it is often desirable to present the teaching and learning materials as closely to resembling to the actual objects in terms of their physical attributes, the technical requirements to handle all the information may not be supported by the learner's computer network. Therefore the beautifully rendered graphics, multimedia, and surround sound audio files may require long periods for a learner to download.

REALISM IN VISUALS

In the research for visual learning, the amount of realism desired has a curvilinear relationship to learning (Dwyer, 1978). Too much or too little realism may affect achievement adversely. The fundamental differences found from the series of illustrations in Figure 1 are the degree of realism. Though the real object may not be able to be captured pictorially, the degree of realism for the object may range from a very abstract level to a degree of relatively realistic depiction of the object as found in a photograph.

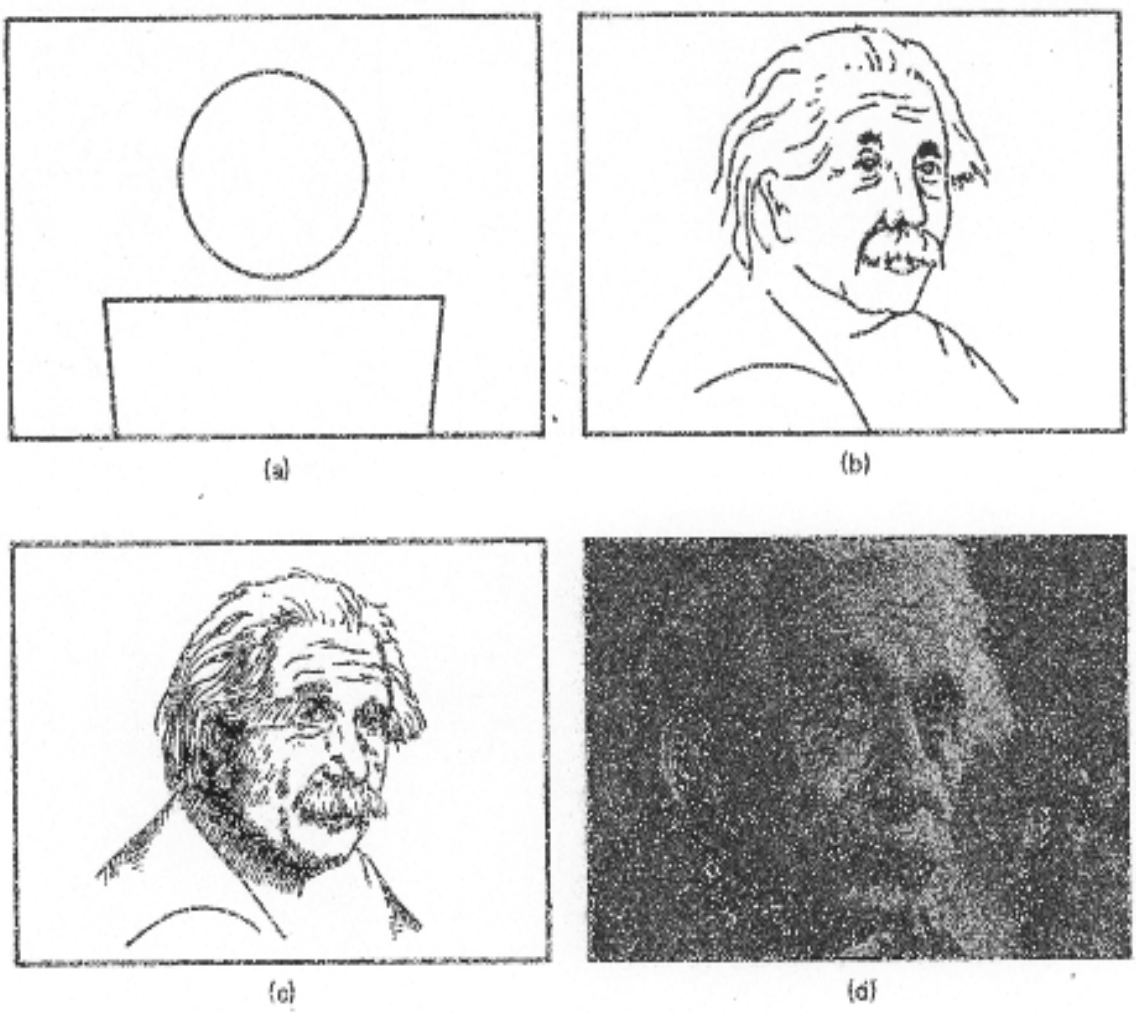

Figure 1: Realism in Visuals (Heinich et al., 1999)

Dwyer (1978) found in his review of visual research the amount of realism desired has a curvilinear relationship to learning. Too much or too little realism may affect achievement adversely as illustrated in Figure 2.

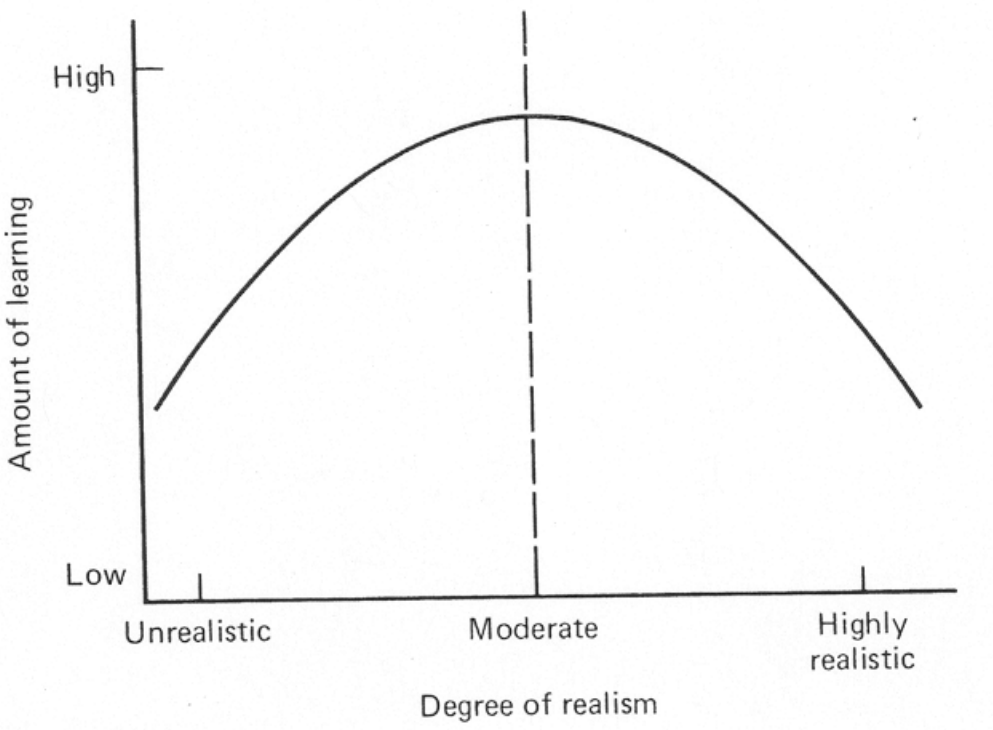

Figure 2: Relationship between the Amount of Learning and the Level of Visual Realism (Heinich et al., 1999)

From the graph in Figure 2, it can be implied that more does not necessarily be better. Therefore educators have to be aware of this relationship and should consider the alternatives for the use of detailed graphics versus a moderately detailed graphics for illustrating their lessons. The current trend for educators presently is to develop and produce multimedia courseware. These digital multimedia materials are to help illustrate complicated and abstract concepts as an aid to help students learn.

The technology available to educators for the production and development of multimedia materials has advanced to a high degree. It is now possible to reproduce real life objects to resemble the original objects almost flawlessly. But the delivery and accessibility technology available for most learners have not caught up to with the production and developmental technology in terms of cost and infrastructure. High speed broadband facilities are still very expensive at the moment and difficult to find in most of the educational institutions in Malaysia. Save for a small number of formal institutions, most schools still rely on dial up technology which is limited to a bandwidth of

56K a second and may be more expensive to utilise exclusively in the long run for a 24 hours, seven days a week setup.

This situation is further complicated when rural schools try to access the Internet facilities through dial up systems as their connection may require extra telephone access fees due to their remote location. Therefore an alternative solution should be proposed to help reduce the cost of courseware accessibility.

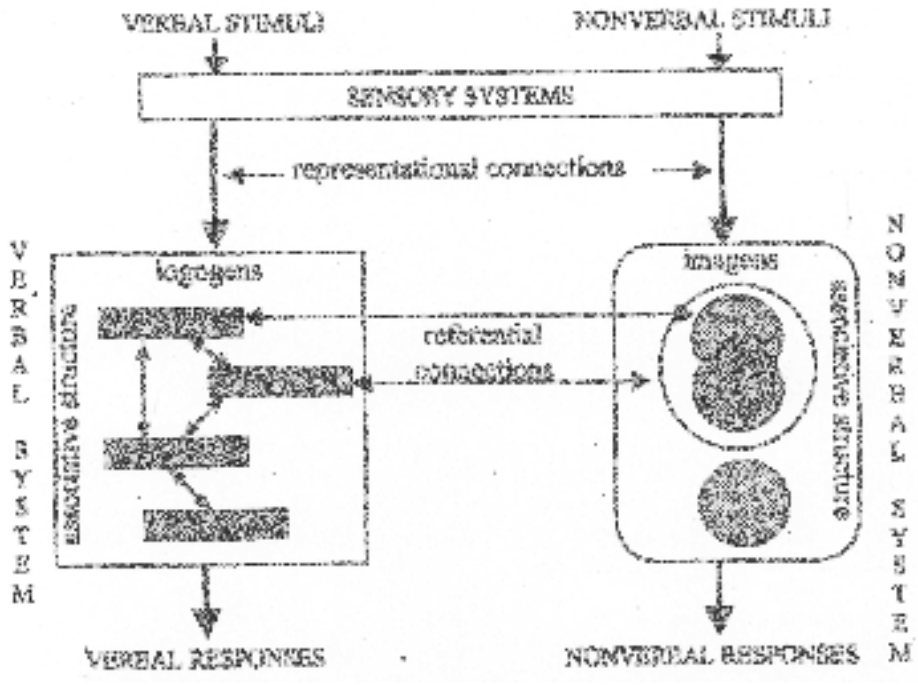

Figure 3: Pavio's verbal and Nonverbal Dual Coding Model (Pavio, 1986)

Pavio's study of mental representation provides another perspective in support of using visuals to represent information. His dual coding model (Figure 3) is referred to as the non-verbal system and signals such as text information is then processed in verbal systems. The representational units found in the verbal system are called the logogens. The representational units in the nonverbal system are known as imagens and are believed to be more holistic. Therefore information that is delivered to a learner has to be filtered, processed and categorised into two sections that is verbal and nonverbal system.

This would then imply that the use of visuals would be beneficial for a developer to include in the multimedia materials and that verbal information with its associative structure will be invoke and the resulting consequences would help a learner engage with the information presented. For example, a learner may be presented with a visual of a cat and accompanied by either the text for "Cat" or a verbal narrative sounding out the word "Cat". The user will then internally invoke the associative structures found in both verbal and non-verbal system. This process would then engage the student to associate the visual of a "Cat" and the word "Cat or its pronunciation as delivered in the form of an audio file and played back to the learner.

COGNITIVE OVERLOAD WITH MULTIMEDIA

With relation to the use of both visuals and audio elements typically found in multimedia courseware, Pavio's Dual Coding Theory advocates that a user will benefit most if the visuals and the accompanying audio information is closely related to allow the users associative systems to aid the learners learn from the materials. But there are other considerations that have been brought forward by researchers in the field of multimedia use for learning and development.

In the design of instructional materials, the main goal is to present information as efficiently as possible. For digital materials one of the major concerns while utilising these materials would also include time and memory requirements, as it would influence accessibility times. The contiguity principles of cognitive theory states that 'On Screen' printed words should be placed near the corresponding graphics. Due to the technological advancement on the use of audio for online materials, it is now practical and feasible to use audio due to advancement of compression and decompression (Codecs) techniques and formulas. According to Clark & Mayer (2003), there is now considerable research evidence that presenting words in an audio form rather than as 'On Screen' texts results in significant learning gains. This is illustrated in Figure 4, as there is a psychological advantage in using both visual and audio material, as the incoming information is split into two separate cognitive channels – words in the auditory channels and pictures in the visual channels. This form of presentation would then reduce the cognitive requirements of a learner if he or she has to concentrate both words and pictures in the visual channel as illustrated in Figure 4, which would then contribute to a cognitive overload for the learner.

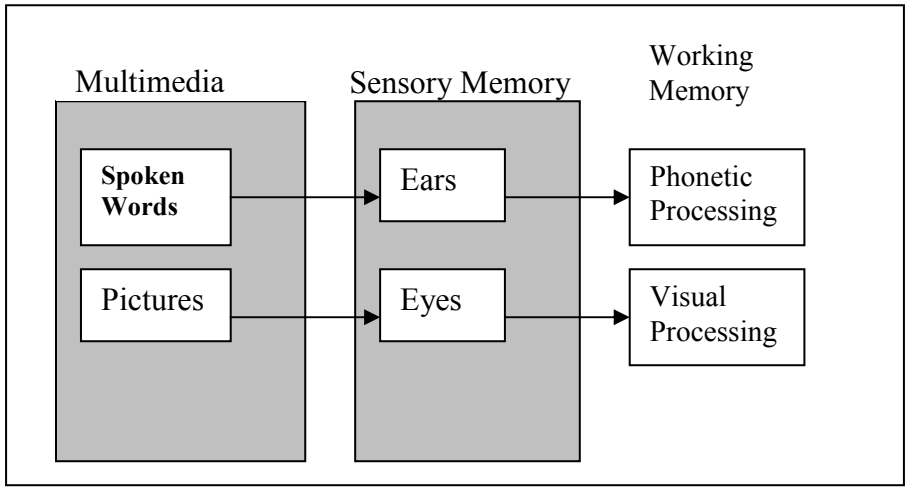

Figure 3.Visual and auditory channels signal processing by the human memory (Clark & Mayer, 2003)

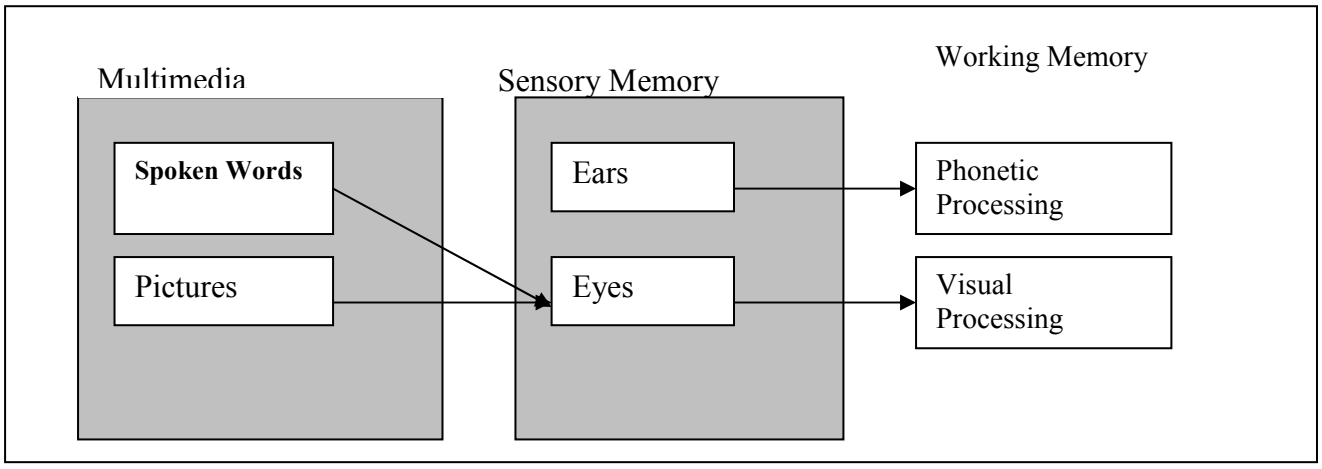

Figure 4: Overloading the Visual Channel with Presentations of Written Text and Graphics (Clark & Mayer, 2003)

When a learner experience a cognitive overload the instructional impact of the learning materials would be negative and the learning would then be less efficient. Clark & Mayer (2003) research findings are based on Sweller's Cognitive Overload Theory as an extension of the Cognitive Theory of Learning. The rationale for the theory suggests that when words are presented as 'On Screen' text in multimedia presentations, this process conflicts with the way the human mind works. People have separate information processing channels for visual/pictorial processing and for auditory/verbal processing. When learners are given concurrent graphics and 'On Screen' text, both information signals must be processed initially in the visual/pictorial channel. The capacity of each channel is limited, so the graphics and their explanatory 'Onscreen' text must compete for the same visual attention. Clark & Mayer (2003) present a Cognitive Learning Theory principle that is called the Modality Principle. Their research strongly suggests that learners are able to learn more deeply from multimedia lessons when words explaining concurrent visuals are represented as a speech rather than as 'On Screen' text. This would then support the use of visuals accompanied by audio narration.

VISUAL QUALITY AND CONSIDERATIONS

For graphic files, besides the actual image screen size, the strongest influence is the choice for the resolution of the graphic as well as the number of colors used to represent the visuals. This is illustrated in table 2.

Table 2: Visual Characteristics of the Common Graphic Types

Graphic Type	Resolution	No. Of Colors	Advantages	Suitability
GIF	8 bit	256 from 16,000 colours	Transparency capable of animation	Good for large areas of uniform colors e.g. banners, icons, plain background
JPEG	24 bit	Millions	Follows a standard by the Joint Photography Expert Group	Good for photographic images with lots of colors
PNG	32 bit	Millions	Transparency capable of animation	Targeted at users with the latest browsers and mobile devices

The main concerns of an educator are to encourage and facilitate a learning process. Therefore it is proposed that when multimedia materials are being used to present information and courseware via networks and the Internet, the materials presented may be in a simpler form such as still graphics and mono audio files or narration. Researches in the use of visuals combined with audio are well established and have concluded that materials rich in details may not always be the best (Dwyer, 1978; Pavio, 1986; Clark & Mayer, 2003). The use of static illustrations accompanied by spoken text will achieve the desired effect and goals for instruction and learning but requiring less bandwidth.

AUDIO QUALITY AND MEMORY REQUIREMENTS

The frequency and intensity of sound waves determine the volume or how "loud" the sound wave is and how distinct or "clear" the sound waves are. All sound waves are in an analog format and a computer system only recognise and process digital signals therefore all this analog waves have to be converted into a digital format. When digitising and recording audio signals from an original source, such as microphones, or from previous recordings, the signals go through several processes. The final quality of the digital audio files relies on several factors such as resolution, sampling rate and stereo or mono separation. Digital materials are ultimately stored in a binary form of ones and zeroes that are commonly called bits. Resolutions represent sets of these bits in clusters to a value of 8. The higher value of the resolution sets, the "richer" the digital sound signal, therefore a digital audio signal utilising a 64 bit set would be technically be much better when compared to a signal using an 8 bit set.

Sampling rates refer to the process by which determines the computer take "snapshots" of the analog waves and duplicates these "snapshots" to determine points to represent the original analog waves. These "snapshots" which are taken in rapid successions for each second of time will determine the accuracy of the digital audio signals when compared to the original analog source. Table 2 illustrates how the sampling rates of an analog signal. Stereo or mono separation is a process where parts of the signals are separated to represent signals which a learner would "hear" either with his left or right ear. This separation would then give an illusion for a normal binaural learner as though he or she is listening to the original audio signal. Therefore a complete stereo signal would then double the size of the file, as the original file size has to be made into two audio signals to represent the "left" and "right" audio signal. This file size would be much bigger in a surround sound audio situation found in immersive environments such as virtual reality systems.

Resolution, sampling rates as well as stereo or mono separations will then ultimately determine the actual size of the resulting analog wave signal after processing. Table 2 illustrates the corresponding file sizes for one minute of a digital audio file based on the technical considerations made when digitising an analog signal to the digital format. These files sizes would then be directly affected and influence by the memory requirements of a delivery system for a learner to access.

Table 2: Memory Requirements and Audio Playback Quality (Vaughan & Tay, 1996)

Sampling Rate	Resolution	Stereo or Mono	Bytes for 1 minute of Audio	Comments
11.025kHz	8 bit	Mono	0.6615 Mb	The lowest you can go for usable results. Very dark and muffled
	8 bit	Stereo	1.323 Mb	Very few advantages for the stereo capability
	16 bit	Mono	1.323 Mb	A thinner sound but very usable
	16 bit	Stereo	2.626 Mb	About as good as a TV set
22.05KHz	8 bit	Mono	1.323 Mb	A decent sound but very usable
	8 bit	Stereo	2.626 Mb	A very popular choice for a reasonable stereo sound.
	16 bit	Mono	2.626 Mb	An ok choice for speech.
	16 bit	Stereo	5.292Mb	Lower than a CD quality but still full and present for stereo sound
44.1KHz	8 bit	Mono	2.626 Mb	An appropriate trade of for recording a mono source
	8 bit	Stereo	5.292Mb	Achieves highest playback quality for most Windows Sound Cards
	16 bit	Mono	5.292Mb	High quality recording for mono sources
	16 bit	Stereo	10.584 Mb	CD- quality recording and recognised standard of audio quality.

CONCLUSION

As any information that has been converted into a digital format, the technical considerations that have to be decided and determine would be the quality of the final file, file size requirements and delivery/accessibility capabilities. As a developer of instructional materials, two variables are within his or her control; quality and file size as the recipients has some control of the delivery/accessibility technology. The learner would then access any material provided by interacting with bandwidth delivered by the Internet service provider. A learner may have a computer which is connected to the Internet by either a dial up facility which is limited to handling 56K of information a second or by leased lines with a higher bandwidth which is capable of processing information up to 100 Mbits a second. But presently even with the marketing blitz to attract consumers to the benefits of broad bandwidth connections, most school systems have yet to move to broadband mainly due to operating costs.

From the information processing models and theories presented in this paper it can be summarised that the use of visuals accompanied by audio recordings can strongly influence how a learner would learn as suggested by Pavio's Dual Coding Model. The learner's associative structures found in both the verbal and nonverbal systems would engage the student's information process and thus make the presentation and learning more effective. But there are several considerations that educators or courseware developers have to be aware such as learners experiencing cognitive overload from the information presented as suggested in Clark and Mayer's Cognitive Overload Theory.

Therefore by taking into account the technical difficulties in terms of accessibility and costs faced by a user it is recommended that educators or courseware developers use static visuals accompanied by mono audio narration. The benefits of presenting information in this format are not reduce when compared to highly realistic motion visuals accompanied by high quality stereo audio narration which may even hinder the user's learning processes due to cognitive overload. Additionally the reduced technical requirements for storing and downloading mono narrated static visuals would reduce download times for a user as well as the cost to access the information when utilising an ordinary dial up system.

REFERENCES

Clark, C.R. & Mayer, E.R. (2003). *E-Learning and the Science Of Instruction.* San Francisco: Pfeiffer, John Wiley & Sons.

Dwyer, F. (1978). *Strategies for Improving Visual Learning.* State College, PA: Learning Servisces.

Heinich, R., Molenda, M., Russel, J.D. & Smaldino S.E. (1999). *Instructional Media and Technologies.* (6th Ed.). New Jersey: Prentice Hall.

Pavio, A. (1986). *Mental Representations: A Dual Coding Approach.* Oxford, England: Oxford University Press.

Vaughan, T. (1993). *Multimedia Making it Work.* (3rd Ed). Berkely, California: Osborne McGraw-Hill.