To See or Not To See: Layout Constraints, the Split Attention Problem and their Implications for the Design of Web Lecture Interfaces

Robert Mertens
Virtual Teaching Support Center
University of Osnabrück
Germany
robert.mertens@uni-osnabrueck.de

Gerald Friedland
Department of Computer Science
Freie Universität Berlin
Germany
fland@inf.fu-berlin.de

Marc Krüger
Research Center L3S
University of Hannover
Germany
krueger@L3S.de

Abstract: The split attention problem and its implications for the design of e-Learning material have been subject to a large number of empirical studies. How results of these studies can be applied to the design of web lecture interfaces is, however, still an issue of debate. Findings are at first glance contradictory and it seems that it is impossible to establish a general guideline. This paper analyses a number of studies and shows that different presentation forms such as slide lectures and chalkboard lectures exhibit significant differences which might be the reason for the above mentioned confusion. The paper discusses how these differences affect the reception of web lectures based on these presentation media. It also presents two examples of web lectures interfaces where the layout of the presentation interface has been designed in such a way that several layout-related problems, including split attention, are avoided.

Introduction

Most live lectures in natural or engineering sciences are comprised of two essential elements, the lecturer and some kind of supporting media. In web lectures, these elements are usually replaced by either a video or an audio recording of the lecturer plus a representation of the supporting media. The representation of the supporting media ranges from drawing-inspired media in Lectern II (Joukov & Chiueh 2003) or E-Chalk (Friedland et al. 2004) over electronic slides in eTEACH (Moses et al. 2002) or virtPresenter (Mertens et al. 2004) to grabbed screens like in the ProjectorBOX (Denoue et al. 2005) or the TeleTeachingTool (Ziewer 2004). While the necessity to display these supporting media in a web lecture interface is evident, the question whether to use a video display or audio only to convey the presence of the instructor is an issue of debate. The reasons against using a video of the lecturer are that it uses up space and that it might lead to the problem of split attention which reduces learning efficiency when students have to concentrate at two different things simultaneously. A video image, on the other hand, conveys a wealth of additional information like gestures and mimics of the lecturer or the audience when a question is asked. In other words, it helps to transmit the dynamics of the lecture. If video is not displayed, viewers get deprived of an essential part of information that is available to the audience of the live lecture (Glowalla 2004).

This paper first analyses didactical and psychological research results concerning the split attention problem in educational media and web lectures. The next two sections discuss whether and how the split attention problem applies to PowerPoint-based slide lectures and to chalkboard lectures. They also present approaches that are designed to tackle the split attention problem and that solve the problem of wasting screen space with a video that is not always important to viewers.

Research results from a didactical point of view

When two streams of mutually referring visual information are presented to learners in parallel, learning results can decrease due to the effect of split attention as shown in (Moreno & Mayer 2000). Similar effects have been reproduced by (Brünken & Leutner 2001). On the other hand, the combined presentation of audio and visual information has shown to reduce the cognitive overload (Brünken & Leutner 2001) and can increase learning efficiency (Mayer 2001).

The implications of these findings for the design of web lecture interfaces seem to be clear. Video images that are in part duplicating what happens on the supporting media should be avoided. This recommendation can, however, be challenged by a number of practical considerations, experimental results in the field of web lectures and experience reports.

Experimental studies have shown the importance of a video image to assess the validity of what is said by providing clues about the speaker (Dufour et al. 2005). These clues like posture, gestures, or mimics can convey non-verbal information about the confidence of the speaker at certain points or irony. It has also been shown by a number of studies that displaying the lecturer's gestures has a positive effect on learning (see Kelly & Goldsmith 2004 for an overview). In a comparative study of lecture recordings with slides and a video image versus recordings with slides and audio, (Fey 2002) has reported that students show higher motivation in the video condition. An acceptance study by (Glowalla 2004) has lead to similar results. Results from (Kelly & Goldsmith 2004) might lead to the assumption that this effect is connected to the display of the lecturer's gestures as a means of communication.

In a short-term study, (Fey 2002) found no significant difference in learning outcomes when lecture recordings with slides and video were compared with lecture recordings with slides and audio only. She does, however, suggest a long-term study to evaluate the effect of higher motivation in the video condition. In another comparative study, (Glowalla 2004) has shown that students prefer lecture recordings with video images over those without and over live lectures because they feel less distracted. Listening to the lecture recording without video was felt by the students to be as boring and ineffective as the live lecture.

(Glowalla 2004) also tracked eye movement of students during watching lecture recording. The empirical study shows that students spend 70 percent of the time watching video and only 20 percent looking at the slides. In the audio only condition, students spend 60 percent of the lecture time looking at the slide. Taking into account the findings from (Fey 2002) discussed above, it does not seem as if the increase in time watching the slides has a positive effect on learning outcome. On the other had, it can be concluded, that recordings with video captured the students' visual attention 90 percent of the time while recordings in the audio only condition only managed to do so 60 percent of the time.

At a first glance, these result from research on split attention in learning in general and the use of videos in lecture recordings in particular seem to be contradictory. A closer look at the research objects involved reveals, however, that the media used in the studies by (Moreno & Mayer 2000) and (Brünken & Leutner 2001) exhibit different properties than slide based lecture recordings. In their studies text was used in combination with video animations. In the study by (Moreno & Mayer 2000) the text and the video based animation have to be followed more or less synchronous and both media had a high information density. This leads to a cognitive overload which in turn causes the split attention problem. These results can not be reproduced in the studies conducted by (Fey 2002) and (Glowalla 2004). Both studies focus on lecture recordings in which slide shows are used. Mostly, the slides and the video do not contain a large amount of information. The most information is presented by words of the lecturer. As a consequence the visual information delivered by video and slides is not sufficient to cause an information overload.

- Based on this analysis, two basic guidelines to avoid the split attention problem can be formulated:
- Cognitive overload should be reduced by keeping the amount of visually presented information low.

Visual media should be presented in a way that does not require users to process related contents on different display regions simultaneously. The following two sections analyses slide and chalkboard based web lectures, discuss how they are affected by the split attention problem and how it can be avoided.

Slide Based Lectures

The way in which presenters interact with the presentation media in slide based lectures like PowerPoint presentations differs profoundly from the way in which for instance a chalkboard is used. In most cases attentional focus on the media and on the lecturer happens in distinct phases as schematically depicted in Figure 1. When lecturers talk, they usually do not animate their slides at the same time. When they trigger an animation or a slide change, they usually do not speak until the animation has finished. An animation or a slide change draws the attention of the audience. When this effect has ceased, the audience concentrates again on the words of the lecturer.
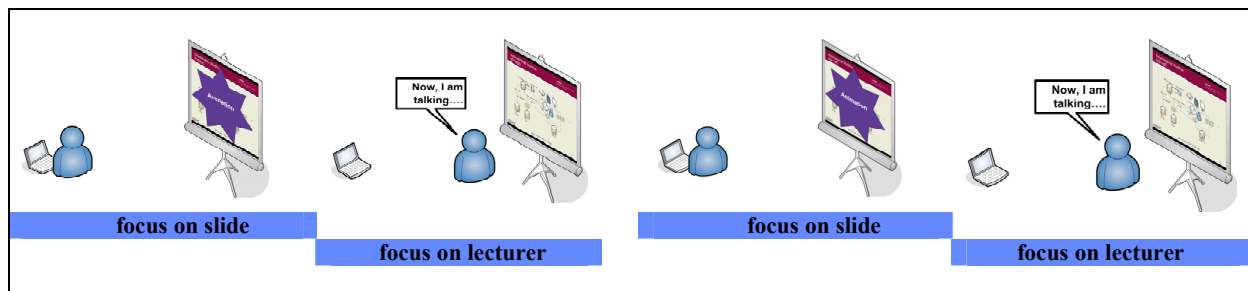

Figure 1: Alternating focus pattern of attention in slide lectures

There is also no physical connection expected between the lecturer and any changes on a slide. It somehow feels unnatural if writing magically appears on a chalkboard. When animations happen on a slide, however, nobody expects the lecturer to move his or her hands over the screen. Another difference that separates slide shows from classic chalkboard lectures is that slides can be animated in order to focus the attention of the audience to a specific part. The role of gestures is thus in many cases taken by slide animations.

This means that not only attentional focus on the lecturer and the presentation medium happen asynchronously but also, that the slide requires visual attention whereas the lecture requires auditory attention. Furthermore, the information presented visually is mostly closely related to the information presented on the auditory channel, at least if the lecturer is doing a good job. If related information is presented synchronously as auditory and visual input, there is no split attention problem but, to the contrary, a congruency effect that helps the audience to process it (Larsen et al. 2003).

The absence of a media based split attention problem should, however, not be taken as a reason for static interface layouts. While the video of the lecturer and the slide displayed do not compete for attention, they still compete for space. The same holds true for a number of other visual elements like navigation aids or in some cases corporate logos in most web lecture viewing interfaces. Navigation elements are usually quite useful when they are needed but take up valuable space during the rest of the time. One solution to this problem is the use of different windows for presentation and navigation as found in the TeleTeachingTool (for a description of the TeleTeachingTool see Ziewer 2004). This approach is accompanied by the drawback that one application is distributed over unconnected and in most cases even overlapping windows on the screen. Switching between these windows is hampered by the fact, that window managers such as the windows desktop environment treat these windows as totally unrelated objects. Unless users arrange these windows on the desktop in an appropriate fashion, this causes navigation to be especially cumbersome when search targets are not found at first trial but several short passages have to be reviewed. Different presentation modes such as the slide or video full screen mode in DVDconnector (Jantke 2005) are another way to tackle this problem.

A modified version of the virtPresenter viewer interface presented in (Mertens et al. 2004) combines different screen modes with the possibility to modify the size of each visual element (slide, video and slide overview) of the interface. As shown in **Fehler! Verweisquelle konnte nicht gefunden werden.**, the interface contains four buttons to switch to either video or slide full screen, a preset overview or a view previously defined by the user.
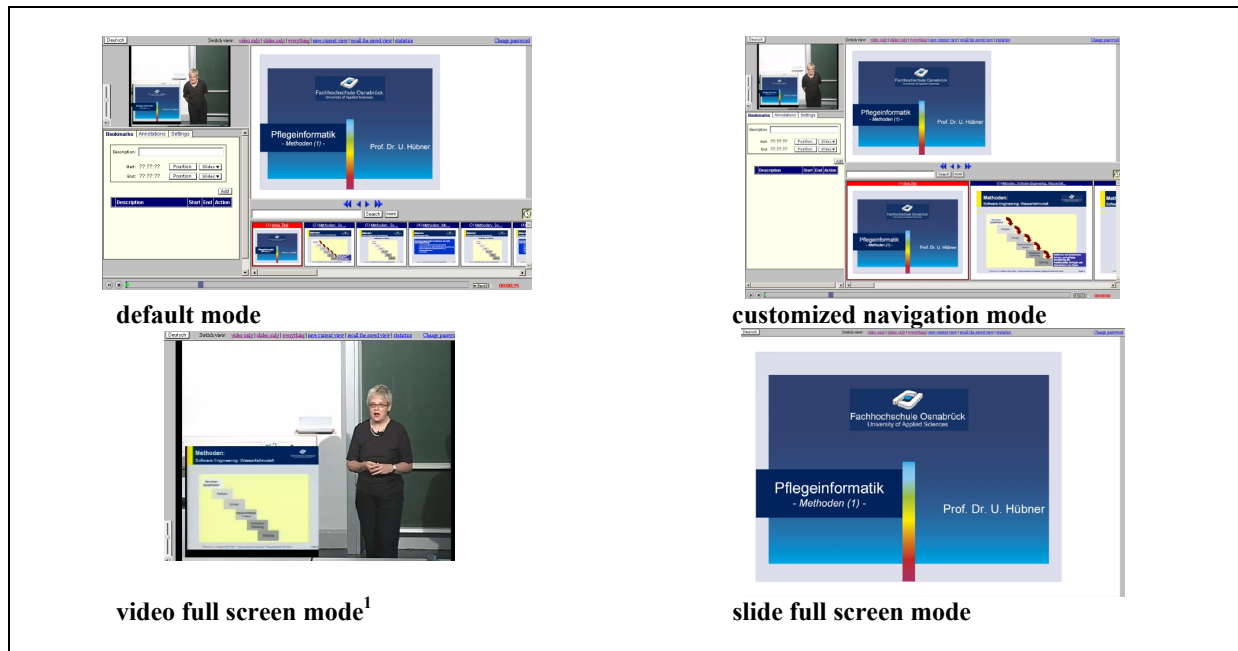
default mode

customized navigation mode

video full screen mode[1]

slide full screen mode

Figure 2: presentation modes in virtPresenter (video, slide, default and customized navigation view)

The scalability of the slide overview component solves a problem described in (Hürst & Götz 2004). The contents of slide overviews are usually displayed quite small which makes them hardly legible. This drastically reduces the information provided by the overview. In the virtPresenter interface, the overview can be enlarged until text on the slides can be read clearly. Since the slides are stored in SVG, a vector graphics format, fonts can be displayed in high quality while the amount of data necessary to store a slide can be kept small.

Electronic Chalkboard Lectures

Lectures held in front of a blackboard can be captured and transmitted in two ways: Either as a video of lecturer and board, or as a set of strokes and images captured by an electronic whiteboard which are rendered with high quality in the remote computer. In order to record or transmit classes, it has become common to use either standard internet video broadcasting systems or software that records and transmits stroke based information.

The advantage of using state-of-the-art video broadcasting software is its availability and straightforward handling. The downside of encoding board data into a video format is a bandwidth inefficient storage. Videos are usually encoded frame by frame. This results in the stroke data being converted from vector format to pixel format. Even though motion compensation accounts for some redundancies, vector format storage is not only several magnitudes smaller, it is also favorable because semantics is preserved. After a lecture has been converted to video, it is for example not possible to delete individual strokes or to insert a scroll event, without rendering huge parts of the video again. Another disadvantage concerns the way most codecs work. Mostly lossy image compression techniques are used that are based on a DCT or Wavelet transformation. The output coefficients representing the higher frequency regions are mostly quantized because higher frequency parts of images are assumed to be perceptually less relevant than lower frequency parts. This assumption holds for most images and videos showing natural scenes where a slight blurring is perceptually negligible. For vector drawings, such as electronic chalkboard strokes, however, blurred edges are clearly disturbing.

---

[1] The slide shown in this picture was encoded in the video as part of an evaluation project. It is not necessarily present there.
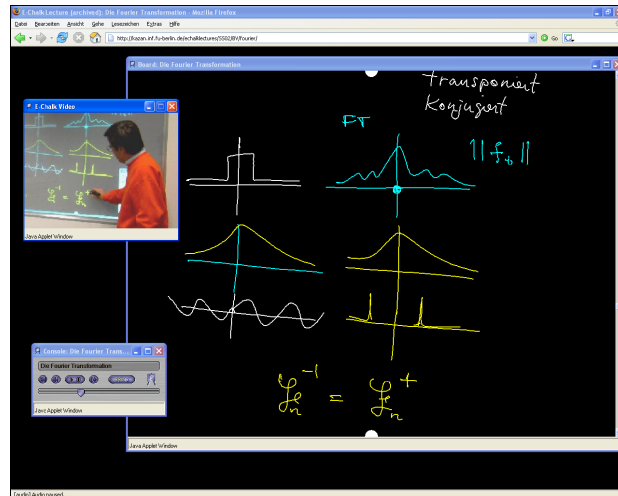
Figure 3: Example of split attention in a remote lecture held with an electronic chalkboard: The board image is transmitted independently of a small streaming video showing the instructor.

Pen tracking devices, on the other hand, capture strokes that can be transmitted and rendered as a crisp image: The strokes can be further processed, for example, using handwriting recognition software. However, when only the board image is transmitted, the mimic and gestures of the instructor are lost and, as explained above, the replay sometimes appears quite unnatural when drawings appear from the void. For this reason, many lecture recording systems do not only transmit the slides or the board content but also an additional video of the instructor (compare Figure 3). However the issue of split attention arises because we have two areas of the screen competing for the viewer's eye: the video window showing the instructor and the board window.
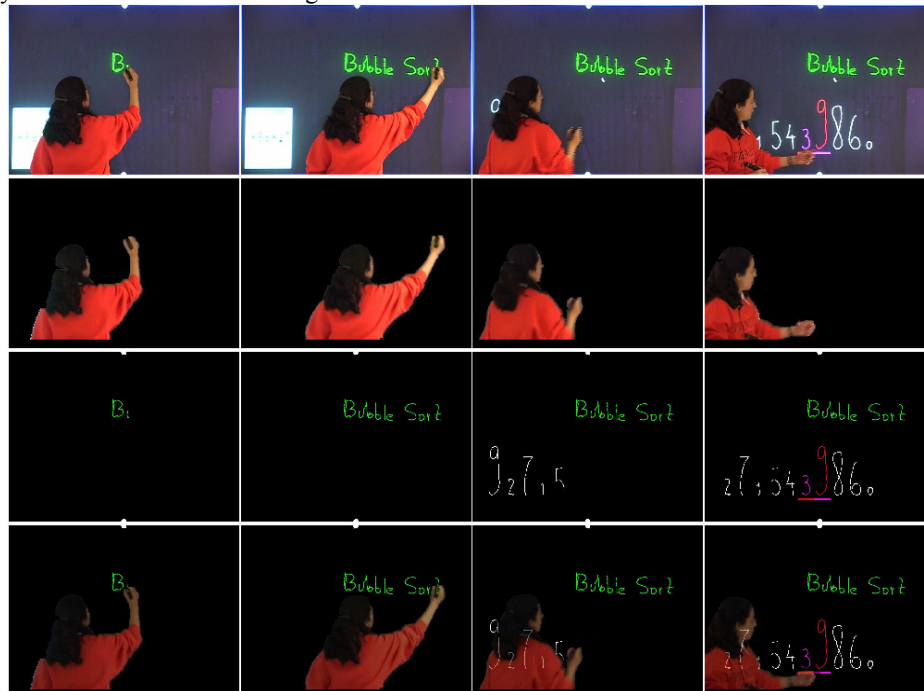

Figure 4: The remote listener gets the segmented image of the lecturer overlaid semi-transparently onto the dynamic board strokes stored as vector graphics which is shown in the last row. (Upper row: original video, second row: segmented lecturer, third row: board data as vector graphics. In the final fourth row, the image of the lecturer is pasted semi-transparently on the chalkboard and played back as MPEG-4 video).

An example of an electronic chalkboard lecture recording system that solves this problem is the E-Chalk project (Friedland et al. 2004). In the E-Chalk project, the video image of the lecturer is cut out from the video stream (Friedland et. al. 2005), separating it in real-time from the steadily changing background. The segmentation is done in real-time without any manual intervention using the video variant of the SIOX algorithm (Friedland et. al. 2005). The image of the instructor can then be overlaid on the board, creating the impression that the lecturer is working directly on the screen of the remote student. Mimic and gestures of the instructor now appear in direct correspondence to the board content. The superimposed lecturer helps the student to better associate the lecturer's gestures with the board content. Pasting the instructor on the board also reduces bandwidth and resolution requirements. Moreover, the image of the lecturer can be made opaque or semi-transparent. This enables the student to look through the lecturer. In the digital world, the instructor does not occlude any board content, even if he or she is standing right in front of it. In other words, the digitalization of the lecture scenario solves another split attention problem that occurs in the real world (where it is actually impossible to solve). Last but not least, the more compact display of the content needs less space and resolution requirements are put down. A feature highly desirable when playing back the lecture on a small handheld device (such as an Apple Ipod).

Conclusion and Further Research

The word "Multimedia" stands for the combined use of different content channels to deliver information that addresses several senses simultaneously in order to provide a pleasant and rich experience. Due to its potential to make learning easier, more convenient, and more effective, multimedia is applied in education very often. Empirical results show that this is the right path to follow. Sometimes, however, the rich experience is reverted to a perceptual overload: The issue of split attention arises when too much information is presented in the wrong way. In this case, the biggest strength of multimedia, namely the easy delivering of a huge amount of information becomes the biggest weakness. We investigated exemplarily two types of media from a practical point of view: Recorded slideshow presentations and remote chalkboard lectures.

Although slide lectures are not really affected by the split attention problem, an analysis has shown that the synchronous presentation of all visual components of the interface still poses a number of problems. Two approaches to tackle these problems have been discussed. Rescaling of visual components has shown to be the better suited approach.

The analysis of chalkboard lectures has shown that split attention is a massive problem. Replaying a chalkboard lecture while showing the lecturer in separate window makes the viewer alternate between the chalkboard content and the gestures and mimics of the lecturer that are presented in the second window. Let alone the visual switching reduces the amount of time the student is able to concentrate on the content of the lecture. We presented a solution that, although technically non-trivial, seems to improve the situation beyond the real world. The instructor is presented in front of the board, but without ever occluding the board content.

It seems that although the split attention problem is known to us for decades and has been cited very often, there is no empirically proven theory on the proper design of multimedia lecture recording systems. In other words, the biggest challenge is still vastly unexplored. The question is not whether to integrate a certain kind of media or not. The question is: How do we create seamlessly integrated multiple media without a cognitive overload. Therefore our basic guidelines give a practical way based on a literature research.

The term multi implies more than the addition of an audio track to a sequence of images, or the combination of digitized pictures with a set of text paragraphs to form an electronic book. Content that uses different media should be presented in an integrated fashion, so that the resulting combination forms more than the sum of its parts without ever producing information overload.

References

Brünken R., Leutner D. (2001). Attention Splitting or Attention Focussing? Empirical Results Concerning the "Split-Attention Hypothesis" in Learning with Multimedia. In *Unterrichtswissenschaften. Zeitschrift für Lernforschung.* Volume 4

Denoue, L., Hilbert, D., Adcock, J., Billsus, D., & Cooper, M. (2005). ProjectorBox: Seamless presentation capture for classrooms. In Richards, G. (Ed.), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2005*, 1986-1991.

Dufour, C., Toms, E., Lewis, J. & Baecker, R. (2005). User strategies for handling information tasks in webcasts. *Proceedings ACM CHI 2005*, April 2-7, 2005, Portland, OR. 1343-1346.

Fey, A. (2002). Audio vs. Video: Hilft Sehen beim Lernen? In *Unterrichtswissenschaften. Zeitschrift für Lernforschung*, Volume 4

Friedland G., Jantz K., Rojas, R.: *Cut & Paste: Merging the Video with the Whiteboard Stream for Remote Lectures*, Technical Report B-05-19, Department of Computer Science, Freie Universität Berlin, May 2005.

Friedland G., Jantz K., Rojas, R.: SIOX: Simple Interactive Object Extraction in Still Images, Proceedings of the *IEEE International Symposium on Multimedia (ISM2005) in Irvine (California)*, December 12-14th, 2005.

Friedland G., Knipping L., Tapia E., and Rojas, R.: Teaching With an Intelligent Electronic Chalkboard, *Proceedings of ACM Multimedia 2004, Workshop on Effective Telepresence*, 16–23, New York City, NY, October 2004.

Kelly, S. D., & Goldsmith, L. (2004). Gesture and right hemisphere involvement in evaluating lecture material. *Gesture* 4:1 (2004), 25-42.

Glowalla U. (2004). Utility und Usability von E-Learning am Beispiel von Lecture-on-demand Anwendungen. In C. Steffens, M. Thüring, L. Urbas (Ed.), *ZMMS Spektrum. Vdi Verlag Gmbh, Entwerfen und Gestalten. 5. Berliner Werkstatt Mensch-Maschine-Systeme*, Volume 18**,** Düsseldorf (Germany)

Krüger M. (2005). Pedagogical Views to Lecture Recordings (eLectures) In *i-com, Zeitschrift für interaktive und kooperative Medien*, Oldenbourg Wissenschaftsverlag, Volume 3, P. 56-60

Hürst, W., & Götz, G. (2004). Interface Issues for Interactive Navigation and Browsing of Recorded Lectures and Presentations. *World Conference on Educational Multimedia, Hypermedia and Telecommunications(ED-Media) 2004*. 4464-4469.

Jantke, K. (2005). Innovative Technologien für das e-Learning: Für und Wider. In: *Proceedings of the Workshop on e-Learning 2005*, HTWK Leipzig, 11.-12. July 11-12, 2005. 9-30.

Joukov, N. & Chiueh, T. (2003). LECTERN II: A Multimedia Lecture Capturing and Editing System. *proc. IEEE International Conference on Multimedia and Expo*, Baltimore MA, July 2003. Volume 2. 681-684.

Larsen A., McIlhagga W., Baert J.& Bundesen C. (2003). Seeing or hearing? Perceptual independence, modality confusions, and crossmodal congruity effects with focused and divided attention. *Perception & Psychophysics*, Volume 65, Number 4, 1 May 2003. 568-574.

Mertens, R., Schneider, H., Müller, O. & Vornberger, O. (2004). Hypermedia Navigation Concepts for Lecture Recordings. *E-Learn 2004: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*. Washington, DC, USA: 1.-5. November 2004. 2480-2847.

Mayer R. E, *Multimedia Learning*, Cambridge University Press, Cambridge UK, 2001, ISBN 0-521-78239-2.

Moreno R., Mayer R. (2000). A Learner-Centred Approach to Multimedia Explanations: Deriving Instructional Design Princilpes from Cognitive Theory. In *Interactive Multimedia Electronic Journal of Computer-Enhanced Learning*, 2.

Moses, G., Litzkow, M., Foertsch, J. & Strikwerda, J.(2002). eTeach- A Proven Learning Technology for Education Reform. *IEEE – Frontiers in Education*. November 6 - 9, 2002, Boston, MA. 267-274

Ziewer, P.(2004). Navigational Indices and Full Text Search by Automated Analyses of Screen Recorded Data. *World Conference on E-Learning in Corp., Govt., Health., & Higher Ed. 2004*. 3055-3062