A framework for the lightweight augmentation of webcast archives

by

Clarence Chan

B.Sc., The University of Toronto, 2005

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

Master of Science

in

The Faculty of Graduate Studies

(Computer Science)

The University Of British Columbia

(Vancouver)

October, 2008

Abstract

We propose a framework for augmenting archives of webcast lectures at alow benefit/cost ratio that finesses the issue of costly video post-production, while still significantly enhancing the quality of the webcast. We argue that lightweight augmentations such as alternate slide tracks, video re-orderings via timeline pointers, and simple groupings of related video and slide content, will allow webcast producers to not only deliver an adequately polished version of the original webcast, but also to re-use and re-purpose archived webcasts for different audiences and different perspectives on the same content.

A prototype application for viewing these kinds of augmented webcast archives was developed, and a user study was conducted to determine the benefits of augmenting an archived webcast lecture in such a way. Our results suggest that augmentations make the webcast itself easier to navigate, and improve comprehension of material enough that selective viewing and navigation of the augmented lecture is comparable to simply watching the lecture in its entirety without augmentations.

Table of Contents

List of Tables

List of Figures

Acknowledgements

First and foremost I would like to thank my supervisor, Dr. Kellogg Booth, for his guidance and mentorship. None of this work would have been even remotely possible without his wealth of knowledge. His helpful advice and insightful criticism were invaluable throughout all stages of my research and writing. I would also like to thank the second reader of this thesis, Dr. Charles Krasic, for taking the time to read and comment on my work.

I am grateful to everyone at the IDRG and LSD research meetings for their helpful feedback on my user study. I am also thankful to Dr. Ron Rensink and his bright students at the Visual Cognition Lab for granting me access to their experiment subject pool.

I would also like to express my thanks to Dr. Ron Baecker, Peter Wolf and the rest of the ePresence team for granting me permission to use several screenshots of the ePresence system and website in this thesis. I should also mention that I am grateful to them for showing me the ins and outs of webcasting and HCI when I was just an undergraduate student, and I hope this work is reflective of their teachings.

The research reported in this thesis was supported by The University of British Columbia and by the Natural Sciences and Engineering Research Council of Canada under the Discovery Grant program and the Strategic Network Grants program as a component of NECTAR, the Network for Effective Collaboration Technologies through Advanced Research.

I also wish to express my appreciation towards certain people in my personal life. I've often heard that things like acknowledgements should be timeless and universal, to avoid running the risk of sounding ephemeral, irrelevant or even humorously embarassing. I don't believe in that. Simply put, there's no time like the present to make a record of what things are important to you, to have something to look back upon when the days go idly by.

To Adrian, for showing me the finer points of cycling, and helping me smash a tooth. Alissa, for always nagging me to do the right thing. Barry, for his wit that far exceeds the teachings of academia. Carlo, for being a bro at just the right moment. Margarita, for reminding me I can just be myself. Mike, for reminding me I have nothing better to do tonight. Sandra, who keeps me up to date on all things 8. Sarene gets Chapter 2, but only because she asked. Sherman, for getting me ice cream after helping me fix a tooth. To W., because Geo is such an awesome (and allergenic) pillow.

Lastly, I want to thank my family. To Baba, Brandon, Spencer and Ryan, for even at the end of all things, I can always rely on your unconditional support and love. Finally, to Mama: because if nothing else, I want something of mine to speak to the fact I never stopped loving you.

Chapter 1

Introduction

1.1 Motivation

The existence of the Internet facilitates the distribution and dissemination of information in digital form. Since the beginning of the 21st century, online distribution of various multimedia content has become significantly easier. Many different technologies have mitigated the problems of storing large amounts of digital media, and of delivering media from point to point in a timely manner. Because of these advances, video webcasting has become an effective and commercially viable method to stream digital media for a number of different purposes.

Webcasting is the term used to describe the delivery of video content in ways that are distinctly different from traditional media communication tools such as television and remote video conferencing, which predated the high-speed Internet era. While the idea of delivering video to remote locations is common to all of these tools, webcasting is unique in that most webcasting systems provide a framework to support and supplement the basic video stream around which the media viewing experience is centered. While webcasting systems vary widely in their nature, they generally share the same basic paradigm of delivering a real-time video presentation or lecture to multiple remote viewers and supporting their viewing of the presentation. Support typically includes (but is not limited to) features such as electronic slides, text chat and some limited interactivity amongst the video presenter and the remote audience.

Despite the promise of webcasting technologies, the real-time experience of viewing a webcast is seldom optimal due to a great number of issues. These range from predictable technical constraints (inherent delays in multicast broadband streaming of large amounts of video) to the unpredictable flow of a live presentation delivered by a human (the speaker may make mistakes in his presentation or otherwise deliver content in a manner that is different from what was expected or rehearsed). If we approach this problem from a user experience perspective there is not much that can be done to make the webcast better that does not involve solving a number of fundamental engineering and real-time production issues.

However, if we turn our attention to recordings of webcasts that are hosted on servers and can be viewed at any time, then we can see that there are usability challenges that can indeed be addressed. A recording of a webcast, which we refer to as an archive, is a permanent copy of the webcast. The challenges faced in archive production and viewing are somewhat independent of the technical challenges of high-bandwidth media streaming.

With webcast archives, there is much room for improving the user experience, without having to worry about the technical limitations or human constraints. While there are still the standard network issues of having to deliver webcast media from one point to another, the technical problems behind archive viewing are not nearly as problematic as with live viewing because there is no real-time media delivery guarantee that needs to be met. Given a copy of the previously delivered presentation, it is theoretically possible to edit and augment it so that it can be delivered on-demand in the manner in which it was "intended" to be received, thus compensating for the problem of human mistakes and errors made in the actual presentation.

This augmentation can be done in a number of ways that are relatively lightweight compared to professional re-mastered releases of digital video. An example of this might be to re-organize chunks of video/slide content or to add supplementary text notes to a presentation. By providing the appropriate interface functionality to access this re-purposed material, users would be able to see a modified presentation at minimal extra effort to the producer of the content, relative to the effort for digital video editing of the presentation in a professional studio.

An appropriately structured interface would allow users to seamlessly navigate the new information spaces represented by a modified archive. Furthermore, as long as the modifications are not permanent and do not require significant production overhead, we can keep the original version of the webcast as an alternate track as additional context for the "optimized" presentation. We can further extend this idea and have multiple viewing tracks of information, all based on the same basic webcast, existing as different viewing contexts for different users. For example, we can imagine sets of specialized slides and slide timings tuned for viewers unfamiliar with the discussion material, as an alternative to the existing material that is tailored towards more experienced viewers.

If we assume that a webcasting system is equipped to let webcast producers modify and augment webcast archives to be navigated in such a manner, it becomes possible to evaluate the benefits of these augmentations from a user's perspective, compare these benefits to their production costs and thus decide on the utility of a system designed to process lightweight webcast augmentations. The chief focus of this thesis is to explore this particular issue of webcast archiving.

1.2 Definitions

In our discussion of webcasting we use several key terms throughout our thesis as defined below:

1.2.1 Definition: Stream

A video stream is a representation of video content that is delivered from a server to a client (or clients) over the Internet. The video frames in a stream are sent by the server, and viewed sequentially by the client as they arrive, without having to wait for the entirety of the video to be transferred. A video stream may have no sound, or it may have one or more audio streams embedded within it.

The stream may represent a pre-recorded video that is delivered over the Internet, or it may be a live video feed that is delivered and played out in real time (minus any network delays in video transfer, artificially induced or otherwise). Streams are delivered point to point between a client and a host, and a host may stream out a video to many different clients simultaneously. It is also important to note that a stream is not a media file, and so does not have any semantically defined beginning or end.

The concept of a video stream is fundamental to our discussion of webcasting. Streams are the basic building blocks of webcasts. Video streams comprise the majority of a webcast's content from a data perspective, as they consume the most bandwidth of any media in a webcast, much more than audio, text or slides. We do not discuss streaming in any further detail in this thesis. There is a great deal of literature and research on the mechanics behind video streaming. The subject of our thesis deals with the application and usage of streams in webcasts rather than the low-level implementation details of video streaming. We instead focus on the higher-level concept of webcasting, which is defined next.

1.2.2 Definition: Webcast

A webcast is a video stream that may be accompanied by any number of parallel media supplements, such as accompanying slides, additional audio tracks, or transcripts of the video. These parallel media supplements are synchronized with the video. There may be multiple simultaneous viewers receiving the webcast at the same time. The webcast material is streamed out to all recipients at once in parallel, a process commonly referred to as multi-casting.

Whether the source material is a live or pre-recorded stream, a webcast is delivered in real time and has a single point of entry for viewing. At any given time, all viewers see the stream and supporting media at the same position (which may vary depending on network delays). An important limitation of a webcast is that it cannot be navigated back in time to view earlier material while it is being delivered. For all intents and purposes the webcast is a 'live', ephemeral event. Semantically speaking the webcast starts and concludes at given times, and is not available outside these times.

We will discuss webcasting in some detail in our thesis insofar as it relates to our discussion of the handling of video and other webcast media in post-production, but we remind the reader that live webcasting is not the main focus of our thesis. We talk about live webcasting in our research mainly so we can describe archives, the objects resulting from the post-production of webcasts.

1.2.3 Definition: Archive

An archive is a webcast that has been recorded and stored on a server to be retrieved by a client on demand at any time. A copy of the video and all supplementary media materials is usually stored on the server. When the archive is requested by a client, all materials are appropriately synchronized with one another and delivered for viewing. By default, the displayed media objects and their synchronization timings are the same as those used in the original live webcast.

Unlike a webcast, which cannot be navigated and can only be viewed in simultaneity with other viewers at the time of delivery, an archive can be viewed by a single client independently and is completely navigable. When the term 'archive' is used as a verb, it refers to the process of recording a webcast and converting it into an archive stored on a server. Used as a noun, it means either a collection of previous webcasts, or (as we mostly use it), a single archived webcast.

Archives of webcasts are the main focus of this thesis. We are particularly interested in how archive viewing may be augmented by using the original media materials of the webcast in different ways from the 'default' behavior we have defined above. This thesis discusses why we would want such functionality in archives, and how this might be made possible through an archive viewing prototype that serves as a proof-of-concept.

Knowing the meaning behind these terms will give us a common frame of reference to discuss the approach we took in our research as outlined in the next section.

1.3 The ePresence system

In our research, we implemented a prototype of an archive viewing application. We did this to see how some of the augmentations to archives we decided upon might be used in a real setting. The concept and design of our application was largely inspired by the ePresence Interactive Media webcasting system [9].

We chose to base our prototype on the ePresence system because its feature set is representative of most modern webcasting systems. The ePresence system is a webcasting package that contains many typical webcasting features that we were interested in working with and building upon. It was also a system whose design and inner workings we were familiar with - the principal author had been previously involved in the development of various software modules within ePresence [10].

The ePresence system is capable of webcasting lectures to remote audiences. A rich set of media channels exist in the ePresence system to supplement the live webcast as well. These include a synchronized presentation slide deck, live chat, and facilitator-managed voice interactivity between the audience and speaker. It also provides an archive facility that lets the producer of the webcast record and upload the archive of the event for on-demand viewing.

1.3.1 Webcasting and archiving features

A lecture can be captured as a live webcast with the ePresence system and streamed out to remote viewers as it happens in real time. Typically, the speaker delivering the lecture also controls the presentation of slides to viewers by navigating to the appropriate slide. This navigation information is sent out to the remote audience. To synchronize the remote slide presentation, the audience watches the webcast using an application running in a web browser that shows the video and the current slide. Text chat and voice functionality, if they are enabled, are also available in the ePresence browser application.

Webcasts delivered through ePresence are automatically recorded on a hosting server, and published as an archive. The video, slides and audio chat stored on the server can be accessed at any time for further viewing. The slide titles and timings are automatically logged during the webcast, so they are identical to those used in the original webcast. In the current version of ePresence, archives can be augmented in various ways. These augmentations include the specification of chapters to divide the webcast into logical sections, adjustment of slide timings relative to the video and some degree of customization of the webcast viewing interface to suit the material presented.

An ePresence live webcast potentially provides much more content than is currently used for typical post-production of an ePresence archive. One limitatoin of the current ePresence system is that it only provides one simple model for editing and re-organizing slides. It is not possible to re-use individual slides for other uses outside of their immediate existence in the single slide deck accompanying a webcast.

We are interested in the re-purposing of archives for a variety of uses. An ePresence webcast has many additional tracks of information and media in its various text chat and audio channels. Accomodating their inclusion in some form in an augmented archive gives the archive additional meaning and context when properly executed. A generalized model for piecing together different kinds of media in an archive "mashup" would be useful in recognizing the value of these additional tracks of information. We were interested in finding novel ways to re-use and present slides. We decided that the basic framework of ePresence's slide model was one worth looking at and expanding upon. This thesis describes our exploration.

1.4 Research Contributions

This thesis formalizes some of the concepts underlying webcast archiving in order to gain a clear understanding of the current paradigm of webcast viewing and navigation. This formalization gives us a model on which we can construct a prototype. The value of this process is that in laying the foundations of our model and testing some of our assumptions through our prototype, we can refine both the model and the prototype as needed in future work. This also helps us in making sense of webcast viewing from a user's point of view, so that it is easier to speak of improving the usability of the webcast by identifying where the system succeeds and fails in terms of the theoretical model.

Our work also contributes new ideas to the field of webcasting by demonstrating the potential benefits of alternate information contexts. The work is a first step towards expanding the accessibility of webcast material to different audiences, as well as introducing new paradigms for navigating multimedia material. Through a user study we have gained a preliminary understanding of how such material might be used, making it possible for us to comment not only on how to improve the general usability of webcasts, but also on the viability of presenting customized information through one or more edited webcast tracks.

1.5 Research Methodology

In working towards our goal of finding effective, lightweight augmentations for webcast archives through study of the ePresence system, we first analyzed the issue of webcast usability from a broader perspective. We assembled a list of major commercial and academic webcast systems and compared them to each other to determine the functional attributes common to all systems, as well as analyzing how each of these functionalities are differently realized in individual systems. This analysis and other background research on webcasting is presented in Chapter 2.

We incorporated this knowledge into a conceptual navigation model that describes the structure and flow of a viewable webcast archive, starting with atomic media units. In addition to laying a theoretical baseline for discussion of webcast archives in general, our model gives us the flexibility to describe lightweight augmentations to archives in terms of how the viewed webcast is structured. This lets us build a rich archive navigation framework comprised of the basic archive model's primitives. This in turn allows us to describe alternate archive information tracks in a way that builds upon the current paradigm of a single video and slide track in a webcast. This discussion is the central focus of Chapter 3. A preliminary prototype illustrating some of the concepts introduced here was built on top of the ePresence webcasting framework as a proof of concept. We describe this in Chapter 4.

The final component of the research is an exploratory user study we conducted with a prototype of the ePresence multi-track software that we created. The intent of the study was two-fold. Drawing on the broad range of webcast systems we had studied earlier, we wished to validate a number of hypotheses on archive usage and viewing in its current form, independent of the modifications of our software. In addition to this analysis we also investigated the effect of the navigation and content delivery augmentations in our multi-track software on users' ability to comprehend and absorb webcast material. The description of the experiment, our discussion of the results, and concluding remarks on the ramifications and consequences of our findings, are provided in Chapter 5, followed by a brief discussion of possible future research directions in Chapter 6.

Chapter 2

Related Work

This chapter is divided into two major components. The first describes a number of academic and commercial software systems that are either webcasting packages or have elements of webcast functionality built into them. We compare and contrast these systems to gain a sense of what features are most commonly supported to assist in subsequently developing a generalized model of webcast archive navigation based on these common features. The second component of this chapter provides background information on research relevant to webcast viewing and archiving. We relate how particular developments are significant to the issues we tackle in our research on webcast archiving.

2.1 Webcasting

Knowing the history of webcasting is important to understanding how the webcasting space evolved to its current state today. The inspirations for much of the functionality in modern webcasting systems can be seen in the early works that pioneered this technology.

2.1.1 Early history

The term "webcasting" is a portmanteau of "broadcasting" and the "World Wide Web". It refers to the broadcasting of video media over the Internet. Fittingly, the main aim of webcasting was initially to stream video media live over the Internet to remote viewers [36], analogous to how television stations deliver live broadcasts to viewers at home via cable.

A number of systems classifiable as webcast packages [21] have been described as following a push-based content delivery model, much like the way that radio or television stations work. As servers of media content, radio and television stations "push" their content and dynamically announce the availability of their content to their clients, who are the listeners/viewers at home that passively receive content through the channels they are subscribed to, rather than actively reaching out to find specific programming that suits their needs.

Most research projects have taken webcasting in a different direction than the "information portal" model suggested above. Many early webcasting systems were used as educational tools for distance learning and/or professional development in a corporate setting. These include the Berkeley Internet Broadcasting System [33], Sun Microsystems' internal Forum system [23], and the ePresence System [9], which has seen regular usage at institutions such the Northern Ontario School of Medicine and the University of Toronto. All of these systems were designed to stream lectures live on a regular basis to viewers not collocated with the speaker.

Today we see webcasting technologies employed for this specific purpose at various institutions of higher learning [9] [7] [4]. Many of these institutions' systems offer archives of their lectures for viewing after the fact, along with the accompanying slides. This suggests there is indeed merit to providing previously webcast media on-demand, because it may serve a number of different purposes to viewers that a live stream cannot [36].

Many early systems in the webcasting space were also designed to promote remote colloboration. Sun's Forum system, as well as Flatland [39] and the TELEP system [24], were designed for lightweight point-to-point delivery of video and slide presentations from speakers to their audiences. Importantly, these systems were designed to support interaction between a speaker and audiences on remote desktops, so that one would not have to be present at a lecture hall or classroom to participate in an event. This also made these systems a good solution for inter-office presentations at the workplace.

These remote collaboration systems were among the first in the domain of streaming video technologies to provide any kind of organized system for interaction amongst the viewing audience, as well as with the speaker. Features included voice and text chat, moderated queues for audience members to ask questions of the speaker, and limited whiteboard functionality. All of these systems promoted the idea of audience inter-awareness, to give a greater sense of latent interaction and engagement in the webcast as a collaborative activity. Examples of awareness manifesting itself in these systems included the idea of raising virtual "hands" to pose questions, and the ability for audience members to provide video or static pictures of themselves for others to see.

This rich history of webcasting laid the groundwork for the standard functionalities that can be found in subsequent commercial systems. Many of these features were eventually replicated and refined in other systems, such as ePresence. Some of the newer systems also include new capabilities that were previously unseen in earlier works. In the next section we document and classify these features to see where our ideas for augmenting archives coincide with existing concepts in webcasting.

2.2 Relevant features

We start our discussion of modern webcasting technologies by dissecting these systems into several key concepts that can be related back to our research goals. They are:

- The use of slides to communicate information from the speaker to the audience
- The use of collaborative tools to connect audience members to each other and to the speaker
- The customization of webcasts and archives on various levels, from appearance to actual informational content
- The indexing and organization of webcasts on both the macro- and micro- levels

We survey the current webcasting systems and analyze how the different functions provided by each system line up with the four features just identified.

2.2.1 Slides

Many webcast systems allow for the display of presentation slides in conjunction with a webcast, or with the playback of an archive. However, some of the earlier webcast systems such as BIBS only provided streaming video and thus had no support for viewing slides, except as independent downloads in the archive of the lecture. Forum and Flatland were among the first systems capable of showing slides in a dedicated section of the screen beside the video during a webcast; we believe this is immensely useful for the following two reasons:

- One cannot typically see slides in a simple video stream of a lecture, unless the speaker is doing a presentation for a local audience and the slides happen to be captured on camera for the stream. Even then, details in the slides are unlikely to be legible in the video. Being able to show the slides directly allows a remote audience to see the slides in full resolution.

- If slides are made available for viewing in this manner, it cleanly divorces them from the video. The video can simply be a way of hearing and seeing the speaker without necessarily having to be a proxy for viewing the slides. This lets the information in each individual channel stand on its own merits, and the flow of the slides is not interrupted even if they are not captured on video.

As seen in Table 2.1, most modern webcast systems, including ePresence, allow the speaker or a trained facilitator to control the display of slides in the webcast. In all these systems, the viewers can view the video and slides simultaneously during a webcast, albeit in the same application window. The majority of systems, including Accordent and ePresence, convert presentation files into JPEG format before delivering them to webcast viewers. They either require that the slide images be uploaded to the server before the presentation, or that the presenter use specialized software on his laptop to convert and push slides from his presentation file dynamically in the webcast as he presents them to a local audience.

If the presenter does not wish to install presentation software for this purpose, both Sonic Foundry and ePresence feature the ability to capture slides from a VGA source. This is useful if the speaker is already presenting the slides to a local audience with a projector, as the VGA source can be split off and fed into the webcast with the right hardware. The number of different ways for delivering slides in a webcast that have been developed speaks to their importance as an information track.

2.2.2 Collaborative tools

Tools for collaboration and two-way communication between speaker and audience have been present in some limited form in most webcasting systems. However, they are generally quite limited in scope as the main purpose of a webcast is to deliver information in one direction, from speaker to audience.

Text communication

Text communication has been implemented in various forms in modern webcasting systems. The text chat model that ePresence uses allows all remote viewers to converse in the webcast viewing application with one another

	A	B	C	D	E	F	G
Accordent	3	yes	yes	no	yes	yes	no
Breeze	1	yes	yes	yes	yes	yes	yes
Elluminate	1	yes	yes	yes	yes	yes	yes
Tegrity	2	yes	yes	yes	no	yes	no
Sonic Foundry	1	yes	no	no	no	no	no
Internap	3	no	no	no	no	no	no
Limelight	3	no	no	no	no	no	no
Webcastgroup	1	yes	no	no	no	yes	no
OBC	3	yes	yes	no	no	yes	no
BIBS	2	no	no	no	no	no	no
Forum	1	yes	no	yes	no	yes	no
TELEP	1	yes	yes	yes	no	yes	yes
ePresence	3	yes	yes	yes	no	yes	no

Table 2.1: Availability of features in different webcasting systems during live streaming mode. Legend: A = # of video formats, B = synchronized slides, C = text chat, D = voice chat, E = whiteboard, F = Q&A / polling features, G = multi-P2P videoconferencing.

	A	B	C	D	E	F	G	H	I
Accordent	yes	yes	yes	4	yes	no	no	yes	yes
Breeze	yes	yes	yes	4	yes	no	no	yes	yes
Elluminate	no	yes	yes	3	no	no	no	yes	yes
Tegrity	yes	yes	no	2	yes	yes	no	no	yes
Sonic Foundry	yes	yes	yes	4	no	no	yes	yes	yes
Internap	yes	yes	no	2	no	no	no	yes	no
Limelight	no	yes	yes	2	no	no	yes	yes	yes
Webcastgroup	no	no	no	2	no	no	no	yes	no
OBC	no	no	yes	2	no	yes	no	yes	no
BIBS	no	yes	no	2	no	no	no	no	no
Forum	no	no	no	1	no	no	no	no	no
TELEP	no	no	no	1	no	no	no	no	no
ePresence	yes	yes	yes	3	yes	yes	no	yes	no

Table 2.2: Availability of features in different webcasting systems during on-demand / archived webcast mode. Legend: A = searchable webcasts, B = searchable libraries of webcasts, C = customizable skins, D = # of platforms E = annotation/keyword support, F = embeddable video, G = custom banners for related content, H = multiple fidelity webcast viewing options, I = Closed caption or alternate language tracks.

in real time, and a trained facilitator monitoring the conversation forwards questions directed to the speaker. Other systems such as Elluminate, Webcastgroup and OBC provide surveys and / or dynamic polling features to elicit user opinions and gather feedback for the speaker in real time, while Online Broadcasting Corporation's Q&A Manager handles and moderates question and answer sessions between the audience and the speaker.

There has been some research on ePresence that deals with text communication. Fono's work on backchannel interactions [17] beyond the initial webcast further builds on the concept of having multiple, useful information threads run parallel to the webcast proper. Data tracks such as conversations extend the webcast's usefulness as an archival tool beyond its ephemeral existence as live material.

Audio chat

The ePresence system has implemented a Flash-based voice communication module to support conversations between remote webcast viewers and the speaker, by means of a moderated queue system [10]. Any viewer who wishes to participate in a voice chat must "raise their hand" to be selected by a moderator, who then patches the viewer through to the speaker and any other previously approved viewers in a multi-way voice chat. Because the voice chat is real time and the webcast's video stream is typically 10-15 seconds late in arriving due to network delays, the viewer's video is turned off during voice chat to eliminate any possible confusion to the viewer.

Extra audio information allows viewers to not only absorb information through the primary tool of presentation (the video and slides), but through an interactive audio channel as well. This work is also of interest to us, not so much for its live interaction capabilities, but because the audio conversations are another channel of information provided by ePresence, an auxiliary channel that provides valuable information in its own right while not incurring the high costs and delays of large-scale video delivery.

Tegrity and Elluminate both feature voice chat systems similar to that of ePresence. However, in the case of Elluminate, all viewers are included in the voice chat by default and the moderator must de-select people to mute them, to encourage participation in discussions. This is part of Elluminate's greater focus towards peer-to-peer conferencing, where the feedback loop between audience and speaker is purposely reinforced.

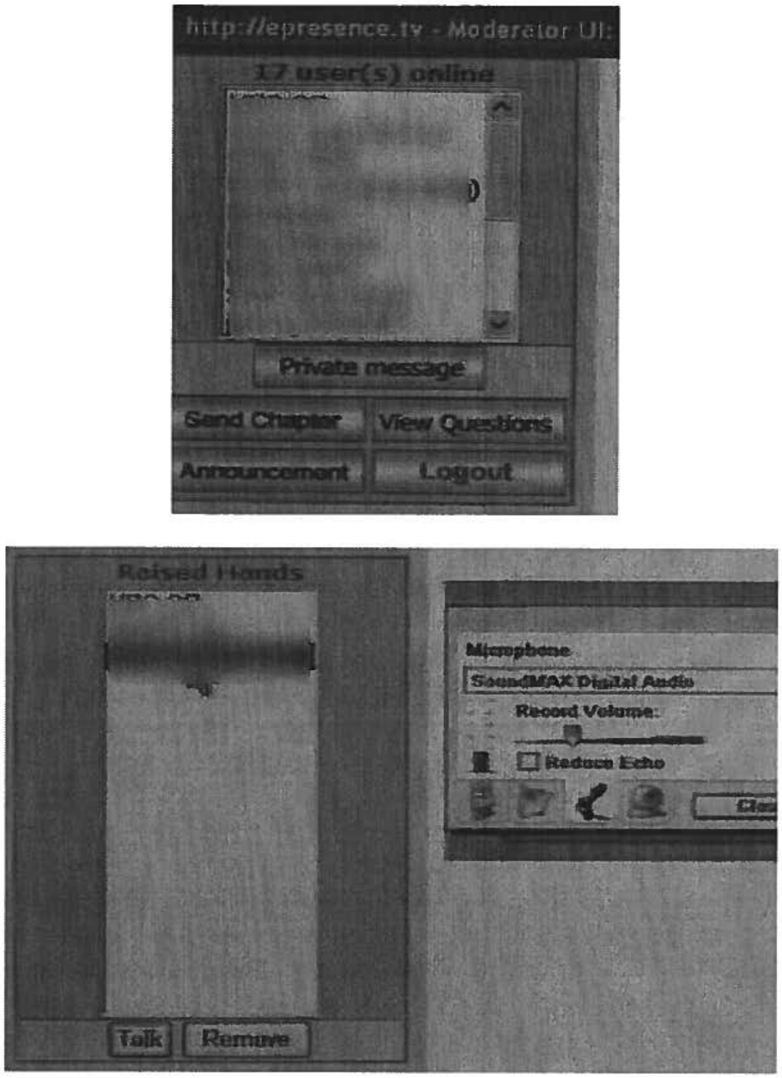


Figure 2.1: A view of the ePresence voice moderation queue. The moderator may select which users to add to a real-time voice conversation with the speaker, and adjust each user's volume level as needed. Screenshot used with permission.

Video conferencing

Video conferencing allows webcasting systems such as Adobe Connect Pro [1] and Elluminate [2] to operate on a more intimate collaboration level. These video conferencing systems feature some similarities with the academic systems we mentioned earlier, like TELEP and Flatland. While all four of these systems possess the ability to deliver full-fledged webcasts, comprised of a single speaker and many remote audience members, they also possess the ability to deliver actual web video conferences. All participants, be they speaker or viewer, can participate with just a webcam. Taking this idea further, Adobe Connect Pro features "breakout" rooms, which allow participants to converse with one another in a window completely independent of the existing web conference to not interefere with the main discussion.

This peer-to-peer collaboration aspect is markedly different from other webcast systems because it puts all participants on equal footing, be they speaker or audience member. Awareness of other participants is of high priority in these systems: their idea of collaboration revolves around supporting communal meetings rather than formal speaker-audience presentations.

2.2.3 Customization/post-production

The ability to customize the look and feel of archives in some capacity is available in most webcasting systems. All systems can be branded with custom colors and logos according to the specifications of the publisher hosting the archive content. We believe that being able to provide this kind of flexibility in the way archives are displayed is a useful precursor to our goal of providing customized archive content for different audiences in different contexts.

Archive skins

In the ePresence system, when publishing an archive, one can choose from a set of prepared "skins" that determine the archive's look when viewed on-demand. These skins vary chiefly in the screen location and size of the video. For example, the default appearance of an archive is a 320x240 video accompanied by a slide pane to its right, but one can choose instead to have the video on the right at a size of 640x480 pixels, and the slides on the left. If no slides were used in the presentation, the publisher can choose to use a skin that has the video occupy the entire space of the viewing window. Sonic Foundry offers similar options to adjust the basic layout and colors used in its archive viewing window.

Accordent features many more ways to customize archives. When publishing an archive through Accordent, the video and slides can be juxtaposed, resized and placed in any of four different regions on the screen. Logos, custom text and other media can be inserted as well through Accordent's PresenterPlus publishing software. Systems like Limelight and Internap take the idea of logos and text a step further, and feature support for banners and customized advertisements for different viewing audiences.

Post-production: supplemental information

A concept more interesting to us than the aesthetic customization is the idea of adding more information to an existing archived webcast. This is an important concept which we believe can be taken in interesting directions by current webcasting systems. Some systems have the ability to modify and edit archives of webcasts to include more information, but they have yet to reach the level of sophistication we envision in our research goals.

Sonic Foundry, Accordent and ePresence all feature software for modifying and uploading archives of webcasts to a hosting server. They include the ability to insert chapter headings, re-order or re-time slides, and embed small video clips in the main presentation. It is possible to edit existing tracks of information with these systems, but they do not feature ways of adding entirely new content parallel to the original webcast.

Nonetheless, there are some novel examples of such functionality. Elluminate has a built-in process for including alternate audio presentation streams into a webcast or archive, while other systems such as Limelight and Tegrity include the option of inserting closed captions for the archive's video. These supplemental tracks of information can be used for a variety of purposes, such as to provide translations of the content into multiple languages, or for interested parties to provide audio commentary on top of the webcast. In a similar vein, the work of Munteanu et al. [29] on ePresence introduced automated transcription of speech in webcast archives, which, when provided as a supplementary viewing track, can be an effective way of re-iterating and clarifying information in the main video (Fig. 2.2).

We believe that a constructive approach to incorporating changes and adding new information is worth looking at. While the systems mentioned here introduce some interesting ways of achieving this goal, we believe that there is much more that can be done in this domain.

2.2.4 Indexing / organization

We comment briefly on methods of indexing and organizing archives introduced by modern webcasting systems. While this is not the main focus of our research, we note the manner in which archives are indexed and organized in certain systems can be seen as an indirect source of information about the content of the archive, and it reveals how the archive is perceived by the producer and / or the viewers. This in turn gives us ideas for how archived content can be augmented via indexing, which serves as an additional information source.

Annotations / keywords

Almost all webcasting systems allow publishers of content to host libraries of archives on a server. These libraries all have search functions and browsable interfaces. To enhance the indexing and categorization of these archives in different ways, some systems have provided tools for making them more easily navigable and searchable.

Certain systems allow either the viewer or the producer of an archive to include annotations and keywords at specific times and positions. Accordent and Tegrity support the annotation of slides by the speaker. These text annotations can be added to slides as supplemental information after the webcast is over. These annotations can be made searchable both inside the archive, and within the greater library of archives the in which the content is hosted. In addition to this functionality, Tegrity follows a "virtual classroom" model that allows both instructors (typically the presenter of the original material) and students (the viewers) to write on top of slides using a Tablet PC.

Keywords can be used in ePresence archives as indexing mechanisms. Either the producer or the viewer can add keywords to the archive as a whole, or within the timeline itself at specific points as a bookmark or highlight. It is possible to sort, categorize and search ePresence archives by these keywords.

2.3 Unified view

While most of these systems appear to vary wildly in the features and types of interactions and methods they support, a common theme binds them all together. Whether these systems provide additional channels of communication to support awareness, or extra contextual meaning to webcasts through tags and customized banners or templates, the base idea of presenting alternate tracks of information to further enrich the webcast viewing experience is present. Even many video hosting websites today recognize the value of integrating user-created content within videos. Youtube [6] and Nicovideo [5] both allow for users and the original video creator to add annotations and commentary at specific times and locations in the video.

However, it is clear from tables 2.2 and 2.1 that the focus of developments in webcasting technology is chiefly on live delivery and less on the archival / on-demand webcasting experience. Many of these systems do not support the kinds of elements we would like to see supplementing or re-defining webcast material, such as closed captions, custom skins or even banners. Systems developed in academic circles such as Forum and TELEP focus mainly on supporting the live webcast experience. The goals of their research were mainly to further the understanding of live collaboration. There is a relative paucity of attention to how archived content is managed, viewed and navigated. Many of these systems offer neither keyword searching nor custom annotations.

It remains to be seen whether or not all these different concepts can be consolidated and defined under a unified model of video webcasting. Our model hopes to draw all these variations into a single theme and create a conceptual system that can handle and present all these different kinds of contextual tracks of information without significant loss of generality.

However, to be able to understand how such a model would fit within the constraints of media management and video editing, we need to discuss some developments in other areas of webcasting and video editing at large.

2.4 Video Editing

The domain of video editing extends far beyond just HCI and computer science. Its roots can be traced back to theatrical film production, and many digital video technologies draw their conceptual designs and feature sets from traditional, analog film editing methods.

2.4.1 Traditional Film production

The subject of film production and editing is much too broad to be covered in its entirety in this thesis. However there are several aspects of film production that are of particular interest to us, as they are relevant to our research in webcast archiving. We are interested in the process of making

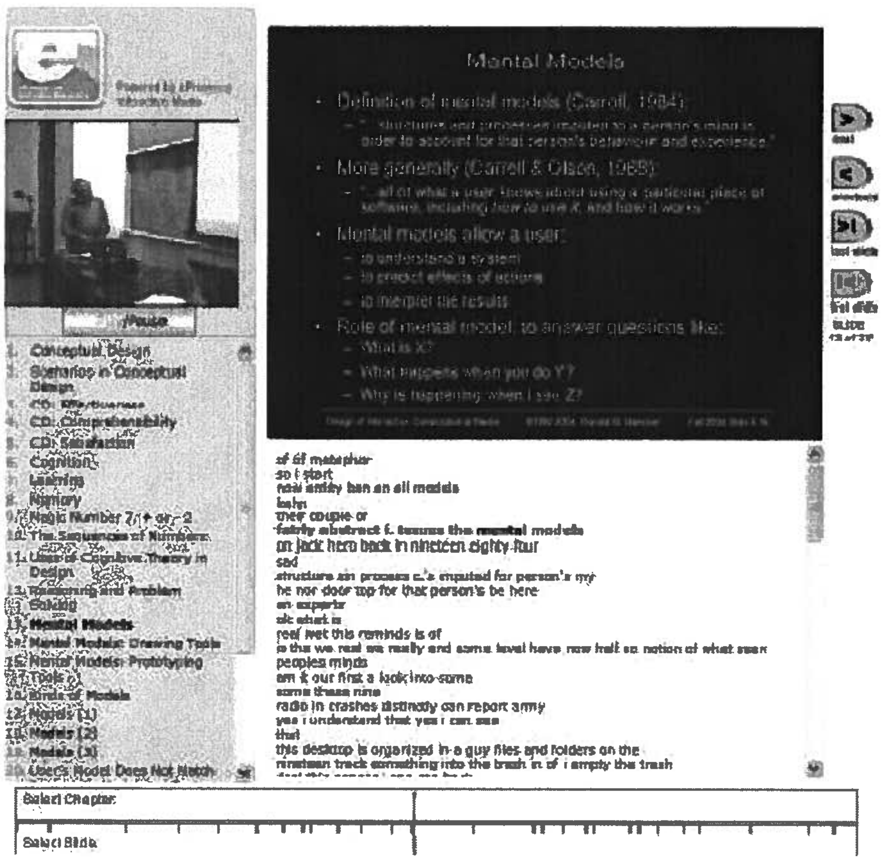


Figure 2.2: The ePresence archive viewer, enhanced with synchronized transcripts. Screenshot used with permission.

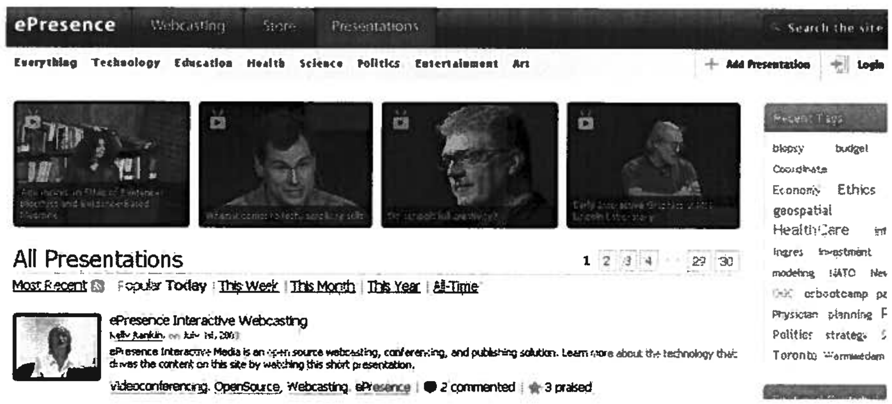


Figure 2.3: The ePresence library features fully searchable archives. It supports searching by various kinds of annotations, such as keywords, tags and comments. These annotations can be inserted by both producers and viewers of content. Screenshot used with permission.

re-cut and remastered versions of films, which touch on several important aspects of media production in general that we hope to reflect in our model:

- judicious re-use of existing media, coming at a high benefit-to-cost ratio
- re-contextualization of existing media with supplementary material and tracks
- re-insertion of initially omitted film content

These techniques have been used to great effectiveness in many reworked versions of classic films, to provide new information and new meaning to movie scenes without outright re-creating the existing material itself.

For example, the 1958 Orson Welles film, *Touch of Evil*, exists in three different versions that draw on the same essentially the same source video material [37]. The sound effects and musical score are tweaked in each version for different effect. One example is the difference in the opening montage between the version of the film scored by Henry Mancini, and the re-edit by Walter Murch. The former is scored with contemporary rock music that aurally dominates the montage, while the latter is suspensefully and discreetly scored with an emphasis on the sounds of cars and footsteps in the movie itself, in accordance with Welles's original vision of the scene as described in his personal notes.

This resulted in two different presentations of the same material. We can almost understand one as being a safer, more commercially viable opening for a mainstream audience, in contrast to the Murch version which generated much more of a *film noir* feel. We can see immediately that without having to dramatically alter the visual content of the source material, it is still possible to provide novel and interesting presentations of old media to different audiences, and that this idea may have merit in other forms and domains.

In the present day, film franchises such as *Star Wars* have produced updated versions of the original movies with deleted scenes re-inserted and new visual effects digitally overlaid in key scenes [18]. Both of these kinds of edits reflect the philosophy of our model: they do not require actual re-shooting or remaking the source film, only changing it in minor ways for maximal effect. Even as more effort is expended in altering scenes using digital production techniques, the base concept of re-mastering film reflects the idea that extra value and previously unseen meaning can be gained by cleaning up imperfections in the original source material. We hope to leverage that kind of power by providing the necessary framework in a webcast environment. Commentary tracks on DVD's serve a similar purpose, providing an alternate voice track as an "overlay" to the original audio, giving incentive to view the original video material more than once for a better appreciation of it.

2.4.2 Modern Video Editing Systems

The benefits of efficient film production are quite apparent when faced with the demanding process of actually splicing together film clips and manipulating separate media tracks in a movie studio. With the advent of modern multimedia computing, many of the remaining issues with the logistical difficulties of video editing, and inherent problems of dealing with temporally navigated media, have been addressed at least in proof-of-concept applications. In Mackay and Davenport's seminal paper on virtual video editing [28], the authors recognized and identified several key bottlenecks in traditional video editing mechanisms and showed how they could be potentially averted in a computer software environment.

They give several examples of software tailored towards different audiences. Their video viewing and editing software, designed for documentary film makers to organize and juxtapose multiple video clips, introduce multi-track videodisc play in a computer environment, as the ability to seamlessly segment and mark video portions in an edit list. In a similar vein, their

EVA system allows for the annotation and tagging of videos of experiments conducted with subjects in real time, while leveraging computing power for analysis and further refinement of observations in archived playback mode.

Mackay and Davenport note that practices in video viewing and editing differ amongst groups, from human factors researchers to documentary film makers and educational software designers. This suggests a need for a varied and flexible set of tools able to create and manipulate hypermedia in different ways for each particular audience. It is interesting to note that in all the examples given by Mackay and Davenport, a common theme of providing supplementary, context-sensitive information exists, be it through additional splicing of tracks, tagging and annotation, or interactive messaging as a backchannel.

2.4.3 Experimental Analysis and Annotation

As mentioned, video has long been used by human factors researchers for analytical purposes. However, their work focuses less on active re-purposing of video materials than it does on passive interpretation of observations. We only mention developments in this domain briefly. Of interest to us are experiment analysis systems that provide rich functionality in "offline" mode, or when viewing an archived video of an experiment. The EVA system mentioned with Mackay and Davenport's work was one of the first to support this kind of practice in a digitized computer setting, as noted earlier.

Baecker and Harrison's VANNA [20] experimental annotation system provides a comprehensive specification for all aspects of such a video system, from data entry and observation annotation, to functional and interface requirements. While most of the systems in this area [25][27][32] fall under the "online", real-time classification according to Baecker and Harrison, they all nevertheless feature the ability to review annotations made during the experiment proper, to link associated comments noted in the experiment for further relationship analysis, and the ability to create a presentation of segmented video clips from an edit list. Even in the domain of experimental recording and observation we can see that video segmentation and recomposition is a useful and desirable function.

2.4.4 Media editing and repurposing

A number of video editing systems have arisen in since the early systems, such as the VAnnotator [15], which supports not only annotations in video but the ability to annotate video with different kinds of metadata through its VideoLens system. Views of the video timeline through various "lenses" provide different information that users have annotated the track with. A track could have one view correspond to technical video information, while another view would contain aesthetic commentary specific to given scenes. In a similar vein, the AntV system tries to support the creation of malleable "video documents" that allow viewers to make sets of annotations and choose which ones are to be displayed in conjunction with the video. Furthermore, the system allows for users to generate their own streams from these documents that combine the existing video documents with user-created text and images.

Another system of note is Davis's doctoral research on Media Streams [16], which supports the re-organization and repurposing of video media through an extensive visual language that not only serves as an intuitive iconic representation of media concepts and operations, but as a fully generative language that is capable of describing complex relations and semantic hierarchies between both re-purposed video content and the annotations that supplement them.

The Family Video Annotator [8] introduces a simplified framework for non-experts to annotate home videos. The system is complex enough to allow for a free-form hierarchical tag model to accomodate different kinds of annotation schemes across different families, but it also tries to simplify the process by suggesting tags for scenes through a combination of tracking tag usage and semi-automatic scene boundary detection. Other novel methods for simplifying common operations in video tagging and editing include Ramos and Balakrishnan's research on employing fluid stylus interaction techniques to manipulate video. The intuitive mechanics of the system allow for natural pen-based gestures to represent common video editing actions such as linking, tagging and segmentation.

Myers et al.'s Silver [14] system attempts to streamline the video editing process by introducing multiple views and classification of video and audio content in Outline, Transcript and Subject modes, so that the more tedious aspects of video editing will be made simpler with a greater number of context-appropriate views. They found that these kinds of simplified views made the process of video editing much easier. The digital video library they used provided metadata such as shot and clip boundaries and transcripts, which made it posible to create these alternate views. We can see that both the concept of multiple views of the same content and metadata annotations, are key to making video a tangible and flexible form of media for repurposing and recomposition.

Further work in this area has mainly focused on automating the process of video segmentation and annotation, which for example could be used to automatically detect scene boundaries in video [38]. Most work in this area stems from Zhang et al.'s foundational work on the topic [41]. These tools have also been applied in specific domains such as the broadcasting of entertainment and sports media, where automation has been used for the generation of video summaries [26] and 'highlights' by identifying "key" scenes [35].

2.5 Information retrieval and hypermedia

We briefly comment on relevant work in the field of hypermedia. Hypermedia is another portmanteau, this time of "hypertext" and "multimedia". In its base form, hypermedia structures are conglomerations of multimedia objects. The structure is not necessarily arranged or viewed sequentially, but rather is held together by a network of hyperlinks that tie everything together through relevant references and cross-references (just like the notion of hyperlinks between pages on the Web). While the actual term, "hypermedia", was introduced in 1965 by Nelson [30], the ideas behind hypermedia predate the usage of the term. Its roots can be traced back to Vannevar Bush's conceptual Memex system [13]. Bush envisioned Memex as a mechanism for tying together data and media, and arranging it in an easily navigable, semantically linked web of related information, which he compared to the thought processes and organizational patterns of the human brain.

Significant developments in hypermedia were achieved with the HyperCard system, which enabled rapid prototyping of a variety of multimedia applications with non-linear progressions (serving much the same purpose that richer environments such as Flash do today), ranging from presentations to electronic documentation, information kiosks [34] and learning environments [11]. These are all domains in which multiple threads of content tailored for different uses and audiences would be useful. Our webcast model draws on this basic notion of hypermedia. Zellweger proposed a methodology for designing tailored paths through hypermedia that would help solve the problems of user disorientation and information overload [40], which is important when users wish to selectively browse hypermedia for what is relevant to them.

The ideas of re-using content and introducing additional content in branching hypermedia paths have been addressed a number of times. The

Dante + WSDM approach to structuring hypertext for visually impaired users acknowledges the usefulness of having supplementary information such as semantic annotations and tags, so as to improve the accessibility of the media [31]. The idea of content re-use and redundancy can also be seen in Harumundanis and Rowland's work [19]. They designed a system for encapsulating documents in XML for quick and opportunistic referencing. They applied it in a constructive setting to documentation of software that was still in development, where the malleable nature of beta release software and incomplete documentation requires quick editing and re-ordering of existing materials.

Bederson and Good's work on zoomable user interfaces as tools for creating novel slide show presentations [12] can perhaps be seen as an extension of the idea of hypermedia to a specific domain. Their slide show software is constructed to support multiple paths and hierarchies on top of flat presentations, with an appropriate user interface to navigate and visualize it. The idea was that in de-coupling individual slides from the path and structure of presentations themselves, a distinction could be made between content and interpretation, leading to the capability of customizing presentation content for different audiences as needed by simply providing a different interpretation (path). Our belief is that a webcast, much like a presentation, could also benefit from similar treatment. Alternate paths and slides in webcast archives would let different audiences absorb and understand content according to their needs. The system should be able to accomplish this while still re-using materials in different contexts as required.

2.6 Archiving: the need for a formal model

A comprehensive review of the relevant literature reveals that much work has been done in the areas of live webcasting, video production and presentation of media, but there has not been a strong body of work that ties all of these concepts together to demonstrate how webcast archives can benefit from valuable ideas and concepts found in each of these fields. The next chapter describes a webcast archive model that attempts to formalize the kind of archive functionality we envision.

Chapter 3

Theoretical Model

In this chapter we construct a theoretical archive model to leverage webcast media materials beyond their immediate existence as simply replications of original live webcasts. We develop the model and justify it through descriptive examples, and we present arguments for the meaningful re-use of multimedia in a variety of different contexts.

3.1 Background

The scope and availability of different kinds of media within a webcast have become much greater and varied in their presentation (as seen in the previous chapter), but the webcast itself is still typically treated as a single synchronous unit of multimedia, with only a single context through which the content of the webcast is interpreted. For example, most webcast systems that have archives with both video and slides will have the slides inexorably tied to the video through slide timings, reinforcing their synchronicity. Transcripts and chat logs are similarly tied by very specific timings relative to a webcast video.

This is further reflected in the archival format of the webcast: any form of editing, be it the insertion of higher quality slides, re-editing of the video, or re-editing of the slide timings, is ultimately a singular re-intepretation of the live webcast (and is more often than not just a straight reflection of it). The archive exists in this static form and is not subject to re-use or modification.

As discussed earlier, there are many situations of webcast usage that may benefit from the re-use and modification of its media channels. While there has been work that examines the utility of extra media channels, there exists no interface that treats these different forms of media as abstract objects, so they may be inserted into webcasts and interacted with in a consistent manner.

In creating our model, we make the following three assumptions:

1. The video component of an archive is special because changing or

adding to it is more costly than with other media, so we do not change it.

2. There is value in finding cost-effective ways of editing and re-purposing existing webcast materials for different uses, so we allow manpulation of the materials that accompany a video.
3. There is value in adding new content to archived materials in a cost-effective and lightweight fashion, so we allow the addition of many types of media *except* video.
4. There is value in preserving the original archive, so we always keep it available, but allow for new interpretations through alternate versions.

A fundamental belief with the model is that it is more appropriate to understand a webcast as a collection of separate media tracks. Just as video production tools split film into multiple video and audio tracks for editing (i.e. changing timings, inserting special effects and overlays), we can think of a webcast archive as having video and audio tracks, as well as any number of accompanying abstract media tracks (for example, slides or synchronized transcriptions).

3.2 Analysis of assumptions

It is important to analyze and justify each of the assumptions we have made about our archive model in order to understand how it works.

3.2.1 Assumption 1

The video component of an archive is special because changing or manipulating it is more costly than with other media, so we do not change it.

This assumption is important to understanding why our model focuses on lightweight manipulations of archive materials. We must remember that not all media can be re-purposed or edited with the same expediency, and video media is the most difficult to work with.

To actually create and distribute video in the first place comes at considerable cost. For example, the transmission of a set of static slides at predefined intervals pales in comparison to streaming the video accompanying those slides. Streaming video in real time is a sufficiently hard problem that various protocols have been developed simply to maintain video integrity and basic quality of service at varying bitrates. Video is also difficult to edit and requires professional tools that are both expensive and time-consuming to use, whereas more lightweight media such as slides require nothing more than an image editing or slide editing program.

Our assumption, therefore is that with more expensive media we are not so concerned with editing and modification as we are with simply re-using and re-presenting at minimal cost. Dealing with video requires a different level of attention and care from other multimedia. Our desire to frame webcast archives as an abstract set of separate yet equal media tracks is a difficult goal to achieve. We are better off acknowledging the fact that some forms of media are inherently harder to use and re-purpose than others. By backing off a bit from the ideal of generic methods of handling media for video, we can then achieve a webcasting model that is adequately equipped to deal with these special cases, but is sufficiently generalized to handle all other forms of media in a unified manner.

3.2.2 Assumption 2

There is value in finding cost-effective ways of editing and re-purposing existing webcast materials for different uses, so we allow manpulation of the materials that accompany a video.

We believe that there are many good reasons to edit and re-purpose existing archives for other uses. These reasons will become apparent in the scenarios we use to illustrated our work in Section 3.3. These demonstrate howpeople could benefit from a system based on our proposed model of archive modification. Before we discuss these scenarios, however, we further examine the motivation for this assumption.

A study by He et al. [22] gave a broader look at how video presentations should be structured for optimal on-demand viewing. They analyzed video training lecture usage over a period of several weeks and developed a set of recommendations for post-production of webcasts. The authors suggested that structuring lecture videos to deliver key points quickly could prevent user interest from dropping off quickly. They also suggest that providing meaningful slide titles and labels in the table of contents would be effective in providing additional context to on-demand viewers. This additional context would simplify navigation: since people tend to jump to different locations in video lectures (not unlike the process of "skimming" through a paper), better text headings would give a clearer sense of what information could be found at various points in the video.

We believe that these kinds of modifications are beneficial for the same reasons mentioned by He et al., and that the goals of our model fall roughly in line with the guidelines they suggested. While some of the suggestions by He et al. may be too difficult or intractable for the kind of lightweight system we envision, they nevertheless give motivation for looking into easier ways of accomplishing the same thing at lower cost.

3.2.3 Assumption 3

There is value in adding new content to archived materials in a cost-effective and lightweight fashion, so we allow the addition of many types of media except *video.*

In explaining this assumption we again turn to a suggestion from He et al.'s paper [22] . They argued that providing quick summaries, annotations and time-compressed audio clips are effective mechanisms for quickly delivering the "gist" of a lecture. These additions to the archive could perhaps to be used as supplemental tools to the main video, and creating them does not incur the same kind of cost as including new video material.

We believe that viewers who find these kinds of additions useful will benefit greatly from them at relatively little cost to the producer of the content. Because the original archive still exists, it then becomes possible for people to view and absorb the content of the archive as it was originally presented, with the option of understanding and interpreting it in a new context with the additional material. The material is made accessible at several different levels with these lightweight additions.

While the suggestions of He et al. vary in the amount of effort required in post-production, they all fall in line with our view that providing alternative sources of data, as well as alternative methods to access the same data, is helpful in delivering a richer experience to on-demand webcast viewers.

3.2.4 Assumption 4

There is value in preserving the original archive, so we always keep it available, but allow for new interpretations through alternate versions.

Adding new material to an archive need not necessarily violate the integrity of the original content. If a proper archive editing system is in place, one should be able to add parallel information tracks or create altered versions while maintaining the original version. It is useful to have the original for a number of reasons. So long as the original exists, it is possible to create any number of derivative works, as the source material can always be referenced. It is also useful to have the original archive so that one can draw comparisons with derived versions, as well as to see and understand interpretations of the source material in different contexts.

3.3 Re-purposing of webcast media

In this section we outline a number of situations that characterize the kinds of needs that could be met with a webcast system designed to promote archival media re-usage. With each example we first state a current (sub-optimal) practice related to webcast production and archival, and then present a theoretical scenario in which these problems could be addressed with in the context of the archive framework we will present.

Our examples focus on the usage of webcasts and archives in an academic setting. One use of webcasts is for professors to deliver lectures in real time to remote audiences, as well saving these lectures as archives for later viewing by students. Our scenarios describe a variety of situations in which a webcast is archived after the professor has delivered the original talk. There are various stakeholders in our scenarios: they include professors, teaching assistants, and students. Students can be further sub-divided into those who have watched the original webcast and those who have not. Each group has a different stake in the outcome of the final archive.

3.3.1 Scenario 1

After delivering a webcast, Professor Adams releases the archive for public viewing. There is a huge demand for a translation of the slides into other languages, and someone has offered to do audio translations as well. The professor wishes to include these alternate versions of the slides and audio into the archive. However, he wants to make sure they are included as separate tracks, parallel to the slides and audio that were used in the original webcast.

Our model is designed to support this kind of situation, where there exists a need for alternate sources of information in an archive. In our model, it would be possible to create alternate, translated slide and audio tracks for this express purpose, while still preserving the original slides and audio as they were delivered in the webcast. Our assumptions concerning the model are justified in this situation. Editing existing materials (like slides) gives us a cost-effective way of producing "new" material for a different audience.

Furthermore, we can see the importance of producing this material in a non-destructive manner, as we wish to preserve both the altered and original versions of these tracks so that viewers have a variety of choices.

3.3.2 Scenario 2

Teaching Assistant Bob feels that some points in the lecture his professor has just delivered via webcast require some clarification. He has spoken to students in his tutorial and they have stated they would like some extra explanation of various sections in the lecture. Bob wishes to edit the content of the webcast before releasing it as an archive but is not sure how to go about doing so, as he feels that editing the entire webcast may be too resource-intensive. He feels that the easiest way to get points across that the professor missed is to include some extra information in the archive, which would explain and expand on certain statements in his lecture.

In a situation like this, inserting extra video footage explaining the points would be too costly in terms of time and resources, and could possibly break the flow of the existing lecture. It would be better in this case to be able to attach parallel tracks of information at key points in the lecture, such as audio commentary, or supplemental slides. This way, Bob can insert his own explanations into the lecture without overwriting or supplanting his professor's original comments. If a model and system is designed properly, these extra tracks will not interfere with the original archive, but instead exist as optional media to support and inform the main video.

3.3.3 Scenario 3

Professor Cody is reviewing a webcast she has just delivered, and has noted that later sections of the webcast are redundant. In the webcast, she reviews information that has already been covered better in the beginning minutes of the webcast, while other sections that should build on previously introduced concepts do not provide adequate context with respect to these concepts for discussion. She wishes to re-tool the webcast in such a manner that these problems are addressed.

In this situation, we note that the capacity for both the re-use and revision of webcast material is high, as background material earlier in the webcast, such as slides or video covering important concepts, could be re-referenced in an augmented archive. Sections covering new concepts that build on older ones could stand to benefit from having some form of redirection to guide the user towards material she would need to know. This material is already available within the webcast so there is indeed value in re-using those portions. There is no new material required; we can gain much out of the relatively small effort put into revising the order of the content.

Additionally, this revision of the material makes it possible to not only review old information, but also omit redundant material with proper redirection as well. There is indeed a strong case for properly identifying and compartmentalizing different media components and sections of the webcast proper so that they may be efficiently referenced and re-used, and omitted as necessary without great effort or detriment to the overall consistency of the webcast.

3.3.4 Scenario 4

Dave, a student, has watched a live webcast of a computer science lecture and wishes to access the archive later to review the material presented. However, he wishes only to review specific parts of the webcast that pertain to the quiz he is studying for. Meanwhile, Ellen is viewing the material for the first time in the archive, having missed the live version. She wants a comprehensive overview as well as the full set of code examples presented in the original webcast. Fred, on the other hand, is from a different faculty and is watching this webcast for the second time because he is unfamiliar with some of the terminology presented. He hopes to find something that will give him a smooth introduction to the prerequisite material for the lecture, or otherwise refresh her memory of certain basic concepts.

It is often the case that a given lecture will have a multitude of different viewers with different backgrounds anddifferent reasons for watching the webcast. This means that each person ascribes his or her own personal value rating to any given part of a webcast, and may in fact place value not only on the information in the webcast itself, but the manner in which it is organized or presented. Just as different tracks of information serve as different means of delivering information, we believe that our assumptions concerning the importance of being able to modify individual media components in post-production is well justified.

The capability to augment or contextualize these individual components as necessary is a worthwhile consideration in delivering a media archive to multiple audiences. To think that a single media archive is sufficient to satisfy the needs of multiple audiences is perhaps naive. It may be the case that material will have to be re-purposed and re-presented with slightly different approaches tailored towards each specific audience. Thus we again see the merits in our assumptions; it is important to be able to re-present these materials in different ways at minimal cost.

Other Uses

Thus far our examples have focused on reasons for developing a unified webcast archive model, through which the original author of the content (and associated parties) can manipulate the recording. However, we note that this model need not be limited to author-driven revisions of the material. As hinted earlier, different users have different needs and different expectations with regards to a given webcast. We can imagine a system where these same users have the power to re-organize and re-contextualize archive content for their own purposes. Such a system might enable users to:

- add annotations or extra slide tracks in parallel to an existing webcast as a method of note-taking and prioritizing the importance of material in the webcast
- Re-order or omit portions of the video and slide tracks in a manner that reflects the importance of the different topics in the webcast to the user
- Share all of these different revisions with other users for collaborative discussion and feedback

With a model that supports easy re-use of different types of materials under a single abstract notion of an information track, users would be free to create new instances of these kinds of annotational webcasts for personal or shared use. This kind of model could be beneficial in empowering users to shape and define archived content. It would allow them to build their own understanding of the archived content and share it with others. The initial assumptions we made in justifying and developing our model are crucial in supporting this kind of media archive re-use.

3.4 Model Specification

Our model is predicated on the principle of re-use and revision of previously recorded webcast media. The ultimate point of such a system is to allow for a previously recorded webcast lecture to be augmented in post-production in an efficient and flexible manner. The expectation is that if a set of simple operations for moving, excising or supplementing individual units of media within the webcast were provided, it would be possible to insert additional tracks of information parallel to, or directly within the webcast, as well as re-organize and re-contextualize the existing webcast material.

The other important aspect of this model is that it is not intended to support "heavy" post-production manipulations of the webcast media. Specifically, the model is not designed to support direct editing of the original source video; "derived" works exist mainly as re-organizations of the video, with any number of parallel non-video information tracks overlaid atop of the existing re-organized material.

To this end, our model employs a basic two-tier resource model to describe a webcast archive. The underlying tier of basic webcast media components are stored as a media "library". This library includes everything from the raw video and audio, to the slides, supplementary images and text provided in the webcast.

The second tier contains meaningful aggregations of individual media components that can be well defined using a set of logical operations. These operations can combine components in a specified temporal order. The "original" webcast that these components comprise can also be described in this manner at the second tier, as can any derived, re-organized versions of the original webcast.

The flexibility of our model provides us with a generalized method of recursively defining groupings of primitive media objects. These groupings can be treated as media objects themselves, and can be manipulated and aggregated in much the same manner, so it is possible to build complex media structures from just a few base rules that govern their behaviours and attributes.

3.4.1 Tier 1

We define the first tier as an un-ordered collection of media components, each of a primitive type. To limit the scope of our webcast model, we identify a few basic primitive types that comprise most webcasts.

- Video: the video(s) that are recorded as part of the webcast archive are stored and indexed, each as their own unit. We note that multiple videos may be recorded of a given webcast (i.e. to provide different angles or cuts) and so we account for this possibility in the model.
- Deck of lecture slides: each individual slide is stored as its own object. If a slide deck is imported from an existing presentation, all slides are treated as individual objects, but an aggregate object representing the

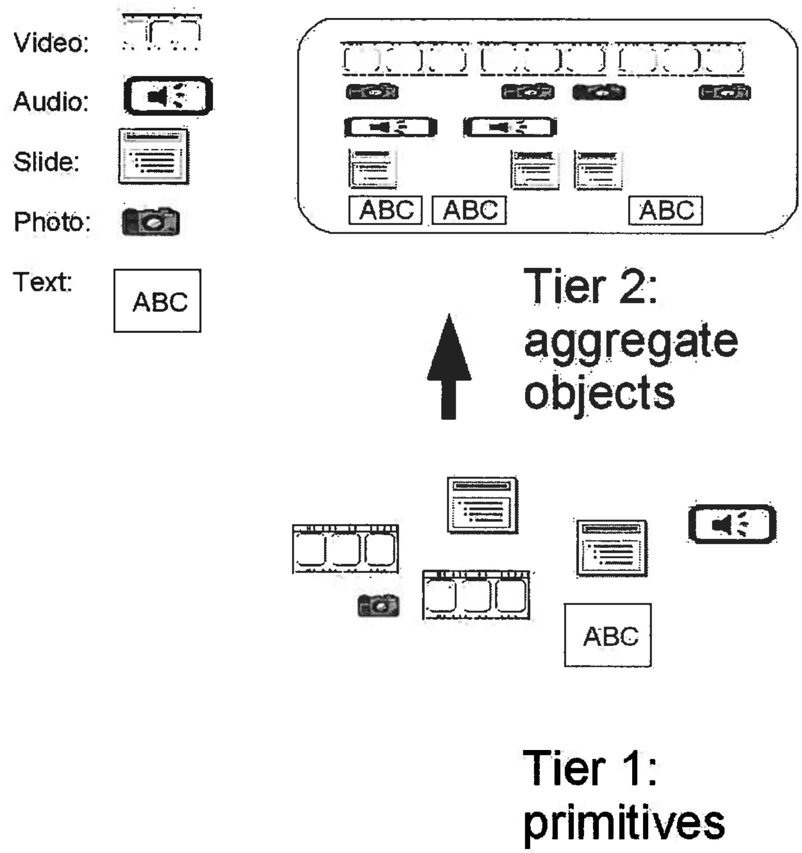


Figure 3.1: Video primitives and arrangements of media tracks grouped according to a tiered model.

entire presentation is created automatically as well. This aggregate object is a second-tier object, which we will discuss in greater detail further on in this chapter.

- Audio: Additional audio tracks are often available in webcasts as either commentary or speaker-to-viewer interaction during the live webcast (as per ePresence's VoIP functionality). The recorded audio stream is stored and indexed.

- Chat logs: transcripts of the text conversation between viewers (and possibly the speaker / speaker's representative) during a webcast can be stored as an indexed track of information. We can treat individual messages as objects; as in the case of the slide deck, an imported chat log will all individual messages to the library, as well as creating an aggregate object representing all messages.

We should note that the above primitives are typically not created independent of one another for the purposes of a webcast. Rarely is a webcast archive constructed in such a bottom-up fashion; furthermore, it is the recordings of existing live webcasts for later use in which we are chiefly interested. However, it is still necessary to define the atoms of the system as we believe simple operations on these atoms will be commonplace, effective means of media manipulation.

Thus, we can imagine that a properly specified webcast system can record a live webcast and segment these different media components properly. We can see in the original webcast itself, that these different components are represented as separate *tracks* presented in parallel, each giving context to another while existing as separate entities. These atomic elements are unmodifiable but can be indexed and referenced into at different points, which brings us to the next tier of the model.

3.4.2 Tier 2

It is at this tier that the basic structure of a webcast can be fully defined. Given the primitives of our webcast model, we can define specific operations for combining and referencing individual media components in such a manner that they can be combined to re-create the original webcast. More importantly, these operations can introduce additional functionality and contextual information to the archive's content that did not exist in the original.

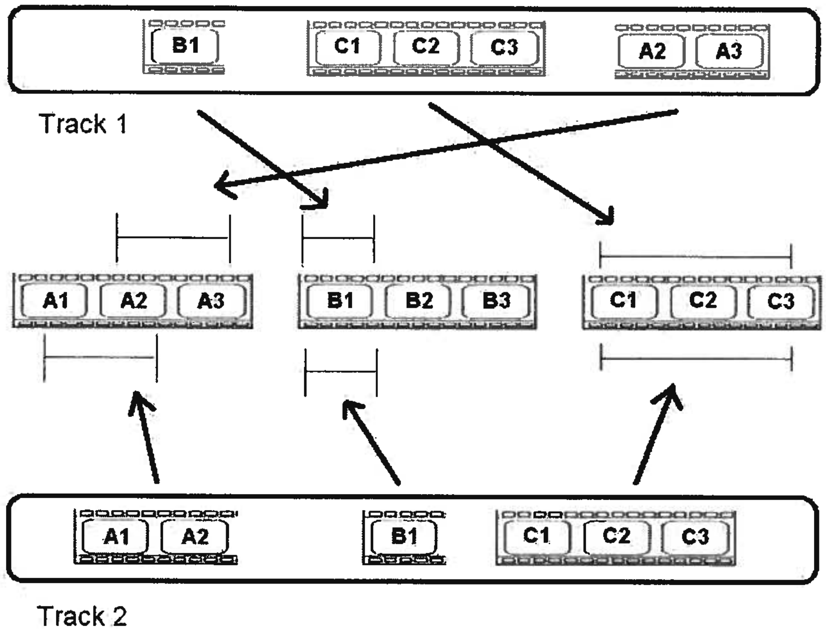


Figure 3.2: Two different video tracks, 1 and 2, drawing on the same source material but with different orderings and timings.

Tracks

As mentioned earlier, we define a track as a non-empty ordered sequence of media objects of the same type. The order determines the order of presentation of the media objects when viewed in a webcast. A single video on its own may be considered a track, as can a set of ordered slides from a presentation. Similarly, an ordering of different videos from the archive may also be considered a track.

This ordered set may possibly have timings associated with each element in the sequence. For instance, within a track, the first video may be designated to play for 5 minutes before switching to the second video, which plays for 3 minutes itself before proceeding to the third. More commonly, a slide deck in a webcast may have a defined order of presentation where each slide is displayed for a set period of time, in order to coincide with the contents of the associated video.

A slide track either has timings for every slide or it has no timings at all; in the case of the latter, the slide track would simply be displayed along with the video, but would have to be manually navigated to view different slides.

For media objects that have an inherent temporal progression (video, audio and chat), it is possible to define intervals of play for each clip. A track is able to reference a set interval for a given clip; i.e. play only the middle two minutes of a five minute video clip. It is also possible to reference multiple intervals within a single clip to be played back to back, or separately at different points of the track.

Thus, if a webcast were recorded with video and timed slides, we can "re-construct" the original webcast in an archive in terms of our model:

1. Two tracks are provided: a video track and a slide track (The video track may have been recorded with audio within, or a separate audio track may have been provided).

2. The video track is comprised of a single video clip: the original recording in its entirety, by itself, with no interval cuts. The interval of play, in fact, can be considered to be the entire video from beginning to end.

3. The slide track is comprised of all the slides uploaded for the live webcast. The ordering and timing of the slides in the archive is the same to that used in the original live viewing.

If we go beyond simply replicating an existing recording, the definition of an archive allows us to easily define webcasts with different slide timings, different orderings of slides, or a different "re-cut" video, as in Figure 3.2.

Tracks do not contain media content, they are simply a set of references to different media objects, not unlike pointers to memory in programming languages. Provided we have all the appropriate media files stored somewhere, a webcast descriptor that implemented this model would simply point to the locations of the desired media, along with playback intervals and timings, rather than store any of the actual media within the descriptor. Depending on the method of implementation, there may be a minimum length to a video interval as the playback of streaming media can possibly introduce delays longer than the minimum length of the video interval, making synchronization difficult or intractable.

Technical description

For simplicity, we can imagine each track to be a sequential array of pointers to individual objects; this is important to understanding the relation of individual video clips to the contextual elements (slides, audio, notes) that give it additional meaning. When an entire webcast recording is imported, it would make sense for a software implementation of this model to segment the video according to slide transitions and timings; in this manner, using the default viewing of the archive, each video clip has exactly one slide associated with it in a clean cut manner.

The slide track aligned with this video track would contain an array of ordered slides, with the slide in each individual cell corresponding appropriately to the matching element in the video track. This allows for context to be easily attached to individual, singular objects by simply putting additional elements in the appropriate cells (see Figure 3.3). Associating a clip with more than one slide can also be accomplished by breaking the clip into as many intervals as there are slides to be associated with the original clip, and combining each interval with its respective slide.

Each video segment defines its own interval timing (the length of the clip contained). We define a combination of media elements in sequential order that occupy a single cell as an *aggregation.* Video aggregations take on the combined length of all clips, and any single slide or audio clip associated with those elements presents itself for the duration of the aggregation.

An audio clip that is shorter than its corresponding video cell will play itself once during the video clip, but we define a "repeat" operation that allows for the audio to be repeated if it does not fill the duration of the video. In the case where audio clips exceed the length of the video clip, one may choose to segment the audio clip appropriately by hand and play the remainder in the next adjacent cell, or, alternatively, the clip can be segmented and played entirely in the single cell, with certain segments selectively chosen for playback (we describe this in further detail shortly).

In the case that an aggregate slide or audio object is aggregated within a video cell, all of the primitives in that aggregate object play in sequence with the video, with no breaks in between, though within that object we allow for temporal spacing between clips to be defined and modified to fill the duration of the corresponding video. "Repeat" and "randomize" functions can be attached to these aggregate objects as well, to either repeat elements or randomize the order of their play for the length of the video.

One extra option exists for the "playback" of multiple elements within an aggregate object. Each element can be set to either present itself when the track is played, as per the normal behaviour of the system, or can be omitted from actual playback. For example, if there are 10 slides in a cell but the author only wishes to have 5 of them presented for the duration of that cell, she may omit 5 of the slides from actual playback and have the preceding slide in the track (be it from an adjacent cell or the same one) take its place. The model also allows for empty space to exist in place of omitted elements.

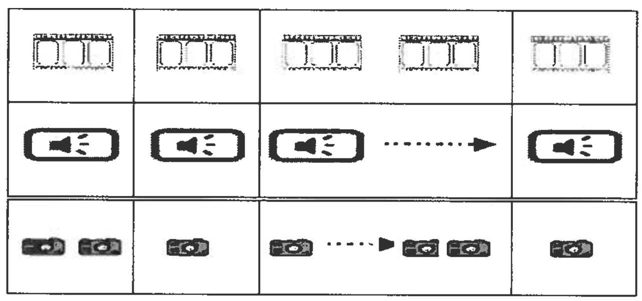
The option for omitting elements from the default playback scheme allows for users of this model to group together many elements in an aggregate object and selectively present those which she deems important. There may be cases where there are simply too many elements to be presented within that timespan, and a smart implementation of the model would give the viewer an option to manually view the extra content within that cell should the need arise. For instance, if the amount of audio commentary exceeds the video, or if chat log messages recorded in that timespan present themselves at a speed too fast for people to read, it may be prudent to present only a portion of the elements and have the viewer selectively decide what else is relevant to her interests through manual browsing during the webcast.

Track concatenation

An important function underlying the technical description is that of media concatenation. We define concatenation as an operation that combines two elements of the same media type (i.e. two video clips, or two slides) in a sequence to form a single track. When one concatenates two video clips together on a video track, those two clips will be presented in the order specified by the concatenation operation. This holds true for any kind of media, timed or otherwise.

When two media elements are concatenated they are treated as a single, new element, these two media elements are the "left" and "right" operands of the concatenation operation that forms the new object. For example, if we concatenate two video clips and view the resulting clip, we would see the "left" clip first followed by the "right" clip. It is possible to place multiple elements of a given media type into a single track. This is done by stringing together many individual elements in sequence by successive concatenations.

For time-based materials, such as video and slides set to specific timings, the resulting object from the concatenation has the combined length of the original elements. When non-timed elements are concatenated, the resulting object does not have any timings associated with it.



1

Figure 3.3: Each cell of a track can contain a variable number of individual media primitives (video, slides, audio, etc.) to be webcast in conjunction with one another. Cells with multiple elements are merely the result of a concantenation operation between multiple primitives. The dotted arrows indicate blank space, if the media clips contained do not fill the maximum time length of that particular cell.

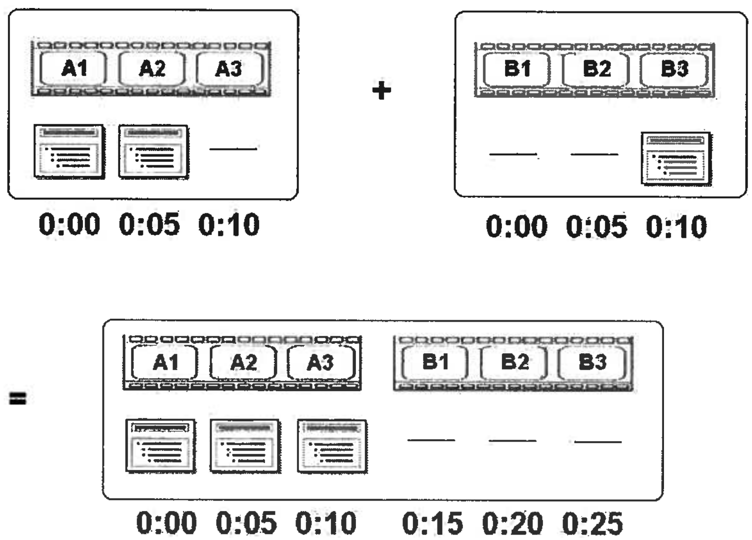


Figure 3.4: A naive approach to concatenation of sets of tracks. When the video and slide tracks are combined without the appropriate "padding", the slide that used to correspond to video segment B3 is no longer in the correct relative position.

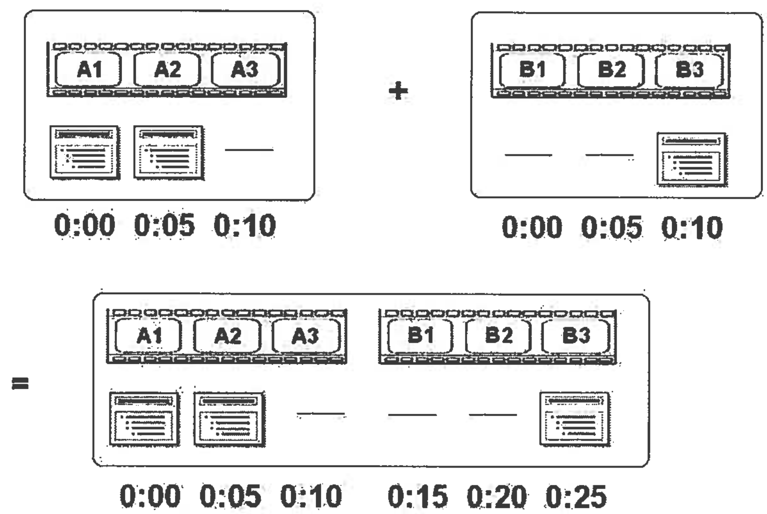


Figure 3.5: The use of padding to appropriately concatenate sets of tracks. When the video and slide tracks are combined, the space in between slide tracks is appropriately filled with empty elements so that all slides remain properly aligned to their respective video segments.

Concatenation of aggregations

A useful functionality is the ability to concatenate groups of tracks together at once. For example, if one is given a slide track, an audio track and a video track, and these tracks have all been synchronized together, it should be possible to concatenate this group of tracks with a separate group of slides, audio and video that have been synchronized with each other.

While it may be tempting to define a group concatenation operation as simply doing simple concatenations on each of the pairs of video, audio and slide tracks, there are certain issues with this definition. For instance, if the "left" video and slide tracks do not share the exact same time lengths, then the "right" video and slide tracks will fall out of synchronization with each other as seen in Figure 3.4.

To solve this issue, we state that a group concatenation operation automatically introduces "padding" to the end of each of the left side elements, which extends the length of the left side element with as much as time as needed to keep the right side elements in relative synchronization with each other (Figure 3.5). The padding used is merely an empty media element that displays nothing. This ensures that elements on the right side still stay aligned. If there happen to be non-timed slide tracks included in each group, they are simply concatenated together without any effect on the timed tracks.

Multiple, parallel tracks

Archives are not limited to containing a single track of each media type. Any number of these tracks can exist in parallel. For example, one may want to define multiple slide tracks that each have different timings, and/or are different subsets of the master slide deck, and associate all of these different slide tracks with a single video track. We can also do the reverse and associate many differently cut or re-ordered video clips with the same slides. As mentioned earlier there are many good reasons for providing this kind of track flexibility; a reasonable implementation of the archive model would allow for the easy switching of slide tracks while viewing the associated video, and vice versa.

We define a parallelization operation that takes two separate groups of tracks and and instead of concatenating them, places each of the individual tracks into the archive and makes them run parallel to each other. This operation takes two parameters, namely the two sets of tracks, and lines up all tracks of each media type. This is a simple yet important operation to creating archives with multiple tracks of information within them.

3.4.3 Complex structures

Main tracks and track chains

We define each webcast archive to have a "main" video and slide track, which is the default set of tracks played in synchronicity when the archive is opened for viewing. In our description of tracks above, we also allow for the existence of multiple tracks of the same media type, which essentially exist in parallel to the main track. We now introduce the concept of track chains, which act as threads of context that allow for tracks to reference content as alternative branches to take.

A chain can be thought of as a diverging point in a track; we formally define a chain to be a link between two slides. It is a path originating from one slide to the other. The only restriction is that the endpoint slide must be on a non-timed track. We make this important restriction to highlight the fact that a slide on the endpoint of a chain does not follow temporally from the other, but is rather a contextually related slide that can be viewed as an aside to the originating slide.

One use of chains would be to link a particular slide on a timed track to the first slide of a non-timed track. As an example, a navigational view of this model would present a "default" progression of slides for a given archive, but would allow the viewer the option to follow a chain on certain slides. The slides that serve as chain origin points would point to slides on other non-timed tracks. These endpoint slides and their tracks could contain extra information related to the originating slide, or background material referenced earlier that is relevant to that slide. This would allow the user to browse related slide material while the video track continued to play.

The software would remember which slide was the origin point of the chain, so that the user would have an easy way of returning to the main video and slide track, thus maintaining one's positional and topical context within the archive. This kind of implementation would allow not only for people to look "back" into an earlier point in the webcast, but also as a foreshadowing tool to see where material will be used in future topics. We develop this concept more in our discussion of the navigational model that accompanies this structural model.

We note that while the chain principle resembles the concept of putting multiple media primitives into an aggregate object, chains serve a different purpose. Multiple elements inside a single cell only have immediate temporal

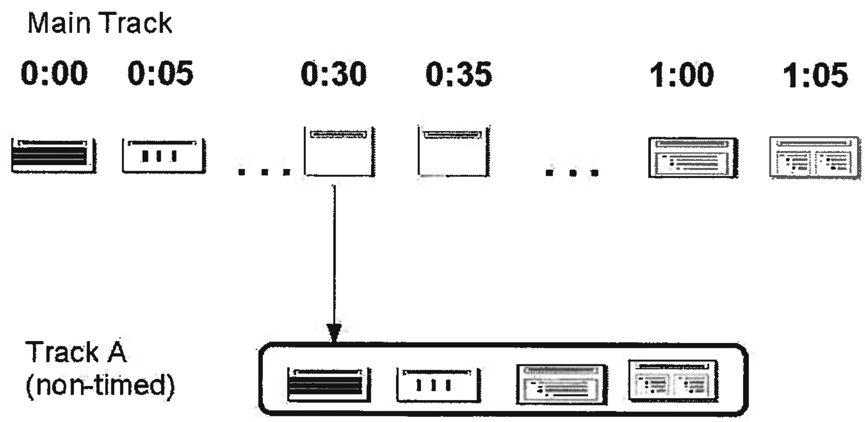


Figure 3.6: The use of chains to link one slide on the main track to four 4 different slides on another. The 4 slides on the also appear on the main track, but at temporally distant points; the chain link makes it possible to quickly access a convenient grouping of the four.

and contextual meaning with one another, whereas the use of chains allows for relations between slides in different tracks altogether. This allows viewers to draw a clear connection between sets of temporally distant materials that may be indeed relevant to each other.

Figure 3.6 demonstrates how a chain connects a given slide to four other slides that appear on the main track, by making references to them on a new, non-timed track. If, for example, material discussed on the originating slide referenced background knowledge from a previously presented slide or slide, but also required understanding of future topics discussed later in the webcast in the second-last slide, then one could follow the chain to quickly find copies of these specific slides in a non-timed track while watching the video. This means that the user does not need to jump ahead or backward in the main track to find each individual slide; all the information is conveniently gathered in the chained track. Because the track containing the four slides is not timed, users are free to browse these slides at their leisure while still maintaining a reference to the slide on the timed track from which they initially entered the chain.

Another example of the use of chains is found in Figure 3.7. A given slide may have related followup slides in the library that are unreferenced

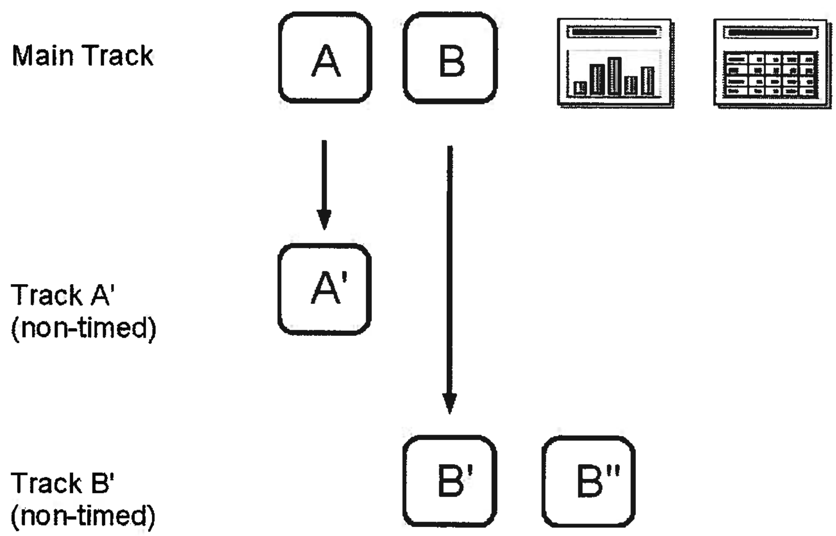


Figure 3.7: The chains coming off the main timeline track point to tracks containing alternates of those slides; in this case none of these alternates appear elsewhere on the main track, and can only be accessed via chains.

elsewhere; or there may exist alternate slides based on the original, but created for different audiences who may be watching (for example, novices unfamiliar with lecture terminology may choose to use a simplified slide). The chain functionality connects all these alternate slides together. A typical implementation would allow these alternates to be optionally viewed while watching the webcast. We note that unlike the first example, these alternates are not pointed to elsewhere on the main track and thus are linked to directly in the library. This alternative material could not be browsed without the existence of chains acting as an affordance to identify the points of divergence.

Multiple chain membership, references with chains

Figure 3.8 demonstrates some more complex relationships with chains. Individual slides may have chains pointing to more than one track. Different slides may chain to the same non-timed track, or to different slides within that track. Originating slides for chains need not be from a timed track; if there is related material that needs further elaboration on a non-timed track one should be able to point to that material using chains, just as with timed tracks. This allows for more complex relationships between slides. The model is flexible enough so that slides that depend on multiple references can be connected to all the related material at once; at the same time, a single slide can be linked to by many different slides elsewhere. This is in line with our concept of finding multiple uses of the same material.

3.4.4 Further manipulation operations

This webcast archive model could be easily adapted to support various forms of track editing and manipulation in a software implementation. Operations such as inserting, copying, and deleting portions of tracks would be of great use to an archive editor trying to polish up a webcast. For constructive functions like insertion and copying, we can define both shallow and deep variations of these operations. In Figure 3.9 we demonstrate the difference between shallow and deep references; the portion of Track B that is being copied and inserted into Track C will consist of a single pointer to a specific location in Track B in a shallow copy. A copy function that is one level "deep" can be used to duplicate the original slides, while not de-referencing track pointers that are being copied (in this case, the one pointing to track A's material); alternatively, one can choose to make a copy arbitrarily deep so as to dereference all pointers as seen in the final example.

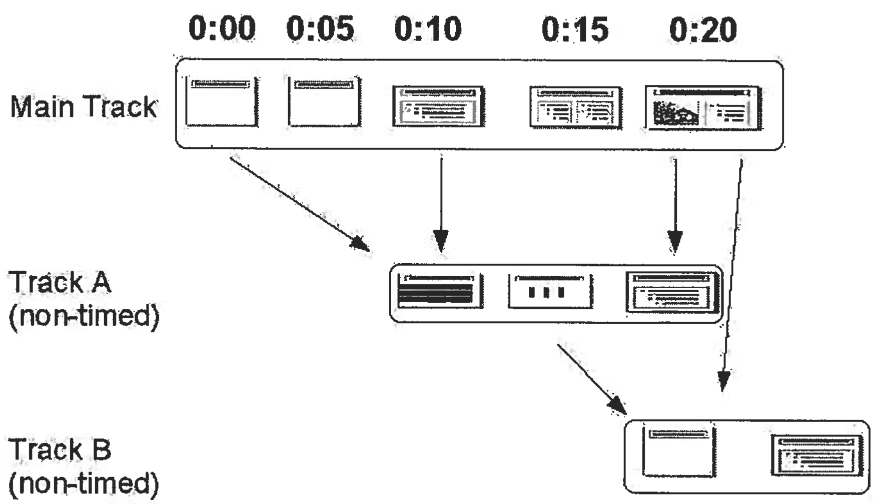


Figure 3.8: An example of how multiple chains between slides work. Two slides on the main track link to form a chain starting with a slide on Track A (an example of linking many slides to single slide on another track), while another slide on the main track is linked to a later slide on Track A, as well as one on Track B (an example of linking one slide to many slides on different tracks). We also demonstrate that slides on non-timed tracks can have chains leading out from them as well, as can be seen from the chain link between slides on Track A and Track B.

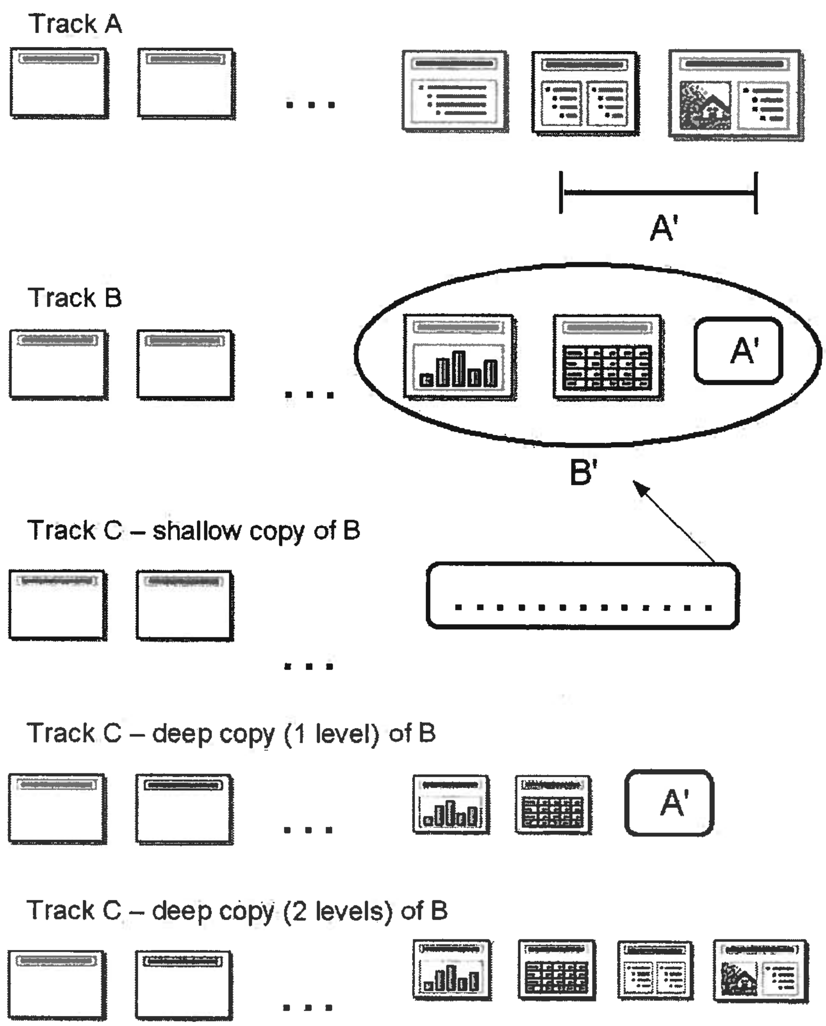


Figure 3.9: Examples of deep and shallow copy functions of varying depth.

Similarly, operations on chains can have deep and shallow variations as well. If one wishes to tie together slide elements on one chain to those on another chain, that person could either a) reference the first chain directly or b) duplicate the first chain in its entirety. Depending on the situation, the first method may be more useful for chain elements being discussed in the same context (different lecture slides referring back to the same key topic), while the latter method would be useful in allowing the existing chain material to be optionally expanded upon without violating the integrity of the original chain.

3.5 Navigation Model

While the underlying structure of these webcast archives are deep and complex in many ways, the theoretical framework for navigating archives that we have developed attempts to simplify and project the essential features of the structural model. We wish to strike a balance between offering rich archive functionality to users, and making the mental model behind the navigation easy to understand and grasp. Our hope is to offer a navigation model that offers consistent behavior and makes it possible to browse webcast archives without being hindered by unintuitive controls, or losing one's place in the main video and slide tracks while looking at other materials (such as alternate slides) in the webcast.

3.5.1 Derivation from structural model

In developing a navigation model we chose to establish a model that would hide some of the structural implementation details that might otherwise confuse or disorient viewers. However, we wished to do so in such a way that would make slide and video tracks to appear as much as simple continuous entities as possible, rather than as aggregations of links and chains. As such, we can think of the navigation model as a "projection" transformation applied on the structural model proper.

To make an important clarification, our basic navigation model still employs the concept of multiple tracks of video, audio, slides and possibly other forms of media. Formally, we define the system as having:

- One or more video tracks, each presented in a linear form, that are completely distinct from each other
- zero or more audio tracks

- zero or more slide tracks with timings
- zero or more slide tracks without timings
- Any number of chain links that point to slides on non-timed tracks

While this description may not be significantly different at first glance from the structural model we provided, we note that there are several key differences.

3.5.2 Video and audio tracks

When viewers watch a webcast archive, they are presented with any number of video tracks derived from the structural model as mentioned above. They are all synchronized together regardless of actual length; video play controls on one video track affect all other video tracks simultaneously (for instance, pausing the main video track would make all other tracks pause at the same time, and playing a track would cause all others to play at the same time as well). All tracks are appropriately padded with empty space to give them the same effective length. A typical navigation implementation would only make one of these tracks actually viewable at any given time by default, though it might be useful to have the option of viewing two or more tracks at once for purposes of comparison.

3.5.3 Timed slide tracks

From the user's point of view, all timed slide tracks update themselves in synchrony with the video. By default, they display the appropriate slide for a given time interval in the video track(s). An implementation of this framework would update all slide tracks in the "background" as appropriate regardless of what slide tracks were actually being shown as the main ones. This is not so different from the way video tracks are presented in our navigational model. However, we introduce a desynchronization operation that causes the slides on all tracks to stop synchronizing their slides with respect to the current video time.

The moment the desynchronization operation is carried out, all timed slide tracks essentially "freeze" in place at the current location; navigating the video will have no effect on the slides and vice versa. Some basic slide navigation operations are made available in desynchronized mode; such operations include the ability to go back and forth, or jump to different locations in the current slide track. As expected in desynchronized mode, timed slide tracks are completely independent of each other; for example, going forward by one slide in one timed track only advances that particular track and none other, and navigating to a particular slide in one track has absolutely no effect on any other timed slide track, which means that one can be at a different temporal location in each track while in desynchronized mode.

To make the synchronization aspect of our navigation model easy enough to comprehend, we state that resynchronizing slides with respect to the video tracks simply resynchronizes *all* slide tracks to a particular point in time.

3.5.4 Non-timed slide tracks and chains

All non-timed slide tracks (which we collectively refer to as the non-timed space) operate completely independently of the timed tracks, and one another, with regards to synchronization, as they simply have no associated timings with them. A viewer's location in each of the individual non-timed tracks is always saved, and navigation within the non-timed track space on its own has no effect on the synchronization of the timed tracks. One can carry out desynchronization and resynchronization operations whilst in the non-timed space, and the operations will still have the appropriate effect on the timed tracks while doing nothing to the non-timed ones.

While browsing the non-timed space on its own is completely independent of the synchronization state of the webcast, we note that the special case of going into a non-timed slide track via a chain on a timed track enacts an implicit desynchronization operation. This is done so that users can return to the point in the timed space where they left off.

3.6 Conceptual player

We would like to be able to implement all of these features to their full extent in an archive player, so we describe how an ideal player might look, visually, under this model.

This ideal player would be able to play all video tracks in a given archive simultaneously. It would also be able to display all slide tracks simultaneously, be they timed or non-timed (Figure 3.10). There would exist the capability to enable and disable the views of any tracks as needed.

The functionality of tracks operates as described earlier in the navigation model. Video and audio tracks are all padded to the same length and play in synchronicity with each other at all times. Timed slide tracks update

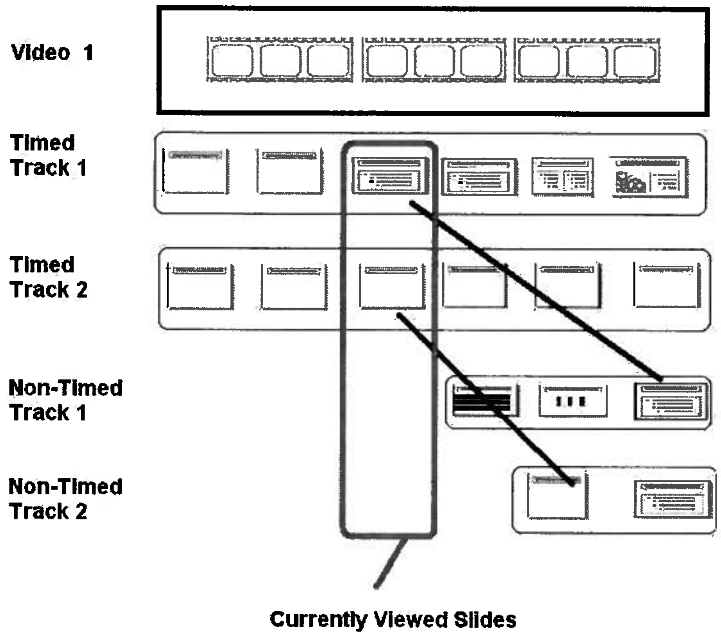


Figure 3.10: A schematic of our conceptual full archive player. All video and slide tracks can be viewed simultaneously, and chain links between slides are made visible.

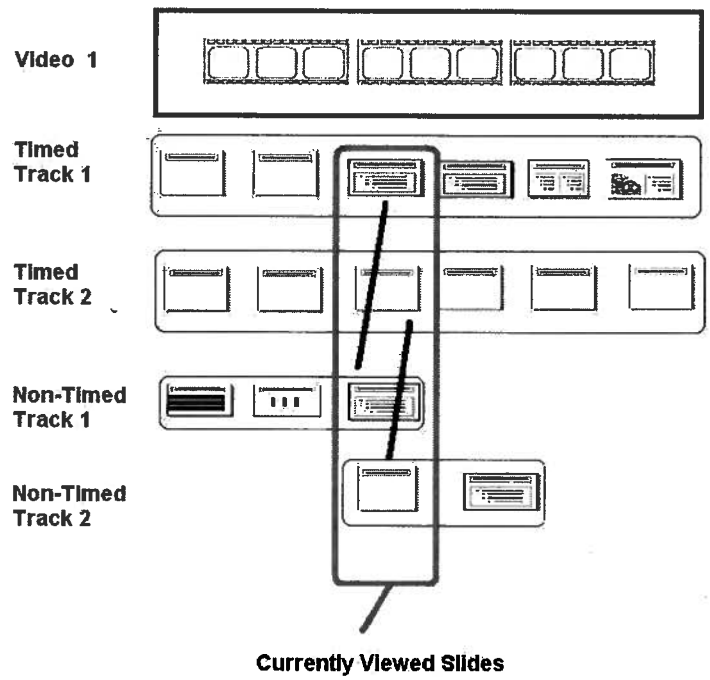


Figure 3.11: The result of following a chain in the full archive player. All elements chained to the current slide are brought into focus and are set as the currently viewed slide for their respective tracks.

themselves with respect to the video, and can be desynchronized individually. For example, if multiple timed tracks are being viewed, they will all advance with the video by default. If one chooses some tracks to be desynchronized, those tracks will be frozen on screen and will not advance with the video.

Chains would be implemented as a way of "lining up" timed tracks with non-timed ones. For example, given a chain that originated on a timed slide A and pointed to a slide B on a non-timed track, making use of the chain functionality would advance or rewind the non-timed track so that the current slide it displayed would be slide B, and would be aligned right under slide A in the player view (Figure 3.11). If timed slides on different tracks had chains pointing to different slides on the same track, extra views of the non-timed track at the different slide locations would be generated as needed when chains are enabled.

While this conceptual player presents an interesting way of watching an archive in its totality, its implementation and user interface challenges are beyond the scope of this thesis. For testing purposes we built a prototype based on a simpler interface model, described below.

3.6.1 Simple player: interface model

In our simplified interface model, we only allow for the display of a single video track alongside one slide track at a time. While multiple tracks are available, they cannot be viewed simultaneously. We use a different way of displaying tracks: instead of being interface elements that can be turned on and off, a set of tabs represents the different "views" of the archive. Each tab represents a combination of a slide track and a video. A simple example is provided in Figure 3.12. There would exist some way of opening a new tab and specifying which slide track and which video track to be viewed in that tab, while making sure that no duplicate tabs (i.e. same slide and video tracks) exist. Following a chain would either generate a new tab containing a view of the existing video and the appropriate slide track being chained to, or would automatically switch the view to an existing tab that already has that combination of media tracks available.

With this model we can see that it is possible to both build complex structures for webcast archives while maintaining usability and simplicity in the final framework presented to users. The effectiveness of making rich edits and constructing novel augmentations to an archive is not lost in the navigation model.

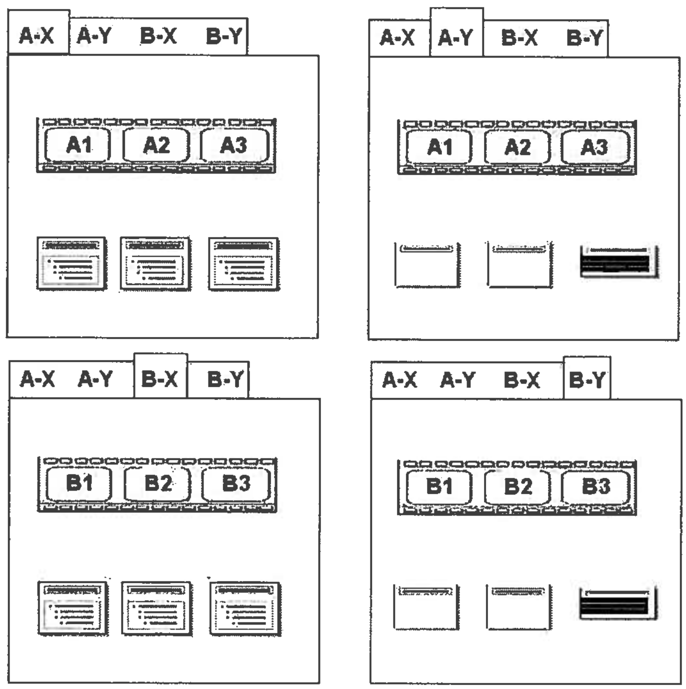


Figure 3.12: A conceptual view of our simplified archive player. In this particular archive there are two video tracks and two slide tracks. Each tab in the window provides a view of a different combination of slide track and video track.

Chapter 4

Prototype

In this chapter we discuss the archive viewing protoypes developed using the theoretical model. We also provide the rationale for specific implementation and model interpretation decisions that we made in creating these prototypes.

4.1 Background

There are many aspects of the theoretical model that was developed in Chapter 3 that are worth exploring via software implementation. The creation and editing of webcast archives, for example, would be an interesting opportunity to explore techniques developed by others in the video editing and production domains. Seeing these techniques in the context of a malleable multimedia presentation such as a multi-track archive would perhaps draw new insights into the benefits and drawbacks of these methods from a cost-effectiveness perspective (which our model has been predicated upon). However, we chose to limit the scope of our work to that of the archive player, and exploring how archive players could interpret the track structures described in our model using prototypes we built.

We approached this problem from the point of view that it would be worth first determining if such an archive model would be of any benefit to people watching webcasts, the main consumers of such a product. We wanted to see how users would respond to archive viewing software based on this model, assuming that the framework for augmenting these archives on the production side already existed. The motivating question is: Given a set of simple, yet informationally rich augmentations to a webcast archive, will users viewing the webcast benefit significantly from these augmentations? In particular, webcast viewing prototypes based on this model could enable to us to see if users can understand the navigation model and be able to use it to acquire more information, more easily.

4.2 Prototype Overview

Our webcast archive prototype is designed for the viewing of previously recorded webcasts. It is designed to read in metadata representing the structure of a multi-track webcast archive and present this information in the form of a viewable webcast, using the media resources whose locations are specified by the metadata. This interface is designed to provide the basic functionality of a typical webcast archive viewer. It introduces a set of controls for navigating and browsing multi-track archives that is in line with the theoretical model discussed in Chapter 3.

4.2.1 Basic controls

Our initial prototype was heavily modelled upon the interface setup used in the ePresence archive viewer. The ePresence viewer (Figure 4.1) features a typical set of widgets designed for browsing webcast archives:

1. A video pane
2. A pane for viewing the accompanying slides
3. A toolbox with a sequential list of all slides
4. A timeline of the video
5. Markers on the timeline representing the temporal locations of slides
6. Navigation arrows to move back and forth between slides

As noted in Figure 4.2, these features (or their equivalents) are present in our first prototype. We also include a miniature set of 3 slide "preview" panes as in the ePresence Producer; the middle pane represents the current slide while the left and right panes correspond to chronologically adjacent ones.

4.2.2 Prototype 0

The first prototype we designed was an attempt to model the link/chain paradigm described in the previous chapter, so that it was possible to easily navigate slide chains added to an archive. We describe below the basic functionality of this software as an archive viewer. It is very similar to the ePresence on-demand in-browser archive viewer. Our prototype was developed in the .NET Framework using Visual C#, and incorporated RealMedia's SMIL framework for video plugins and media management.

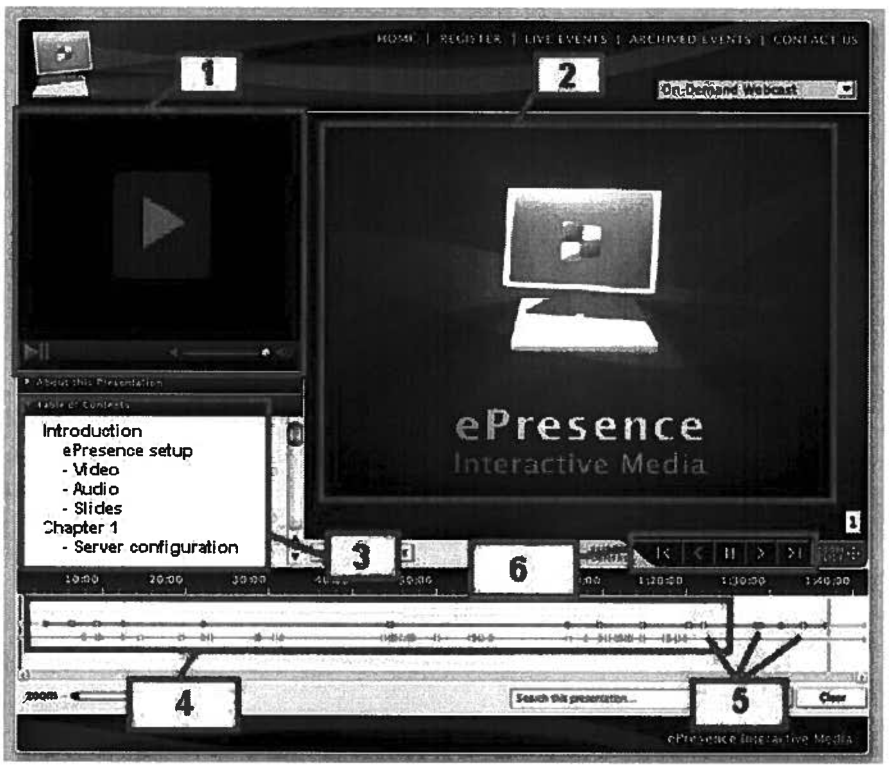


Figure 4.1: The ePresence on-demand archive viewer. Important features are labeled by number and referenced in Section 4.2.2.

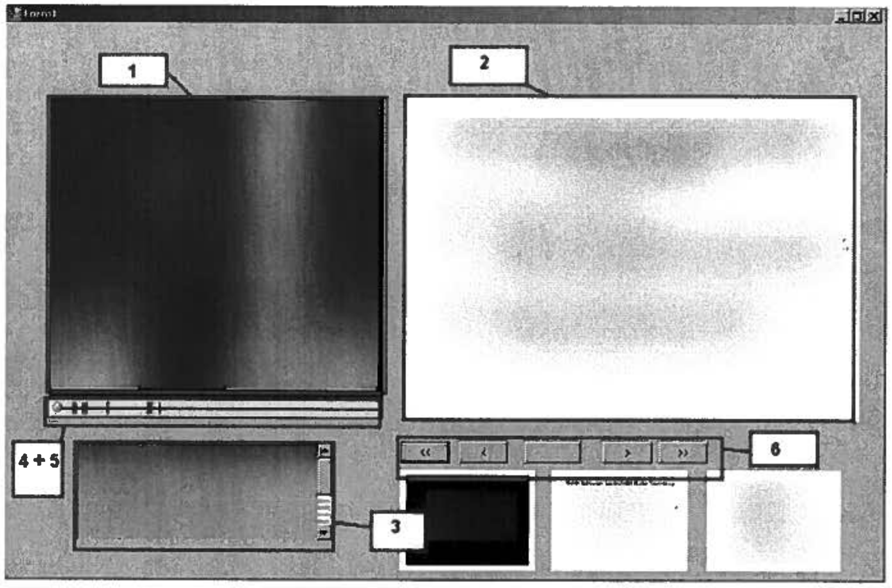


Figure 4.2: Our first prototype of our multi-track viewer, based on the ePresence player.

Functionality

When the archive video is played, the slide pane updates the current slide being viewed according to preset timings in the metadata. Jumping to different locations in the archive by clicking on the timeline updates the slide pane appropriately, as does using the slide toolbox to navigate slides. The toolbox is used by double-clicking on a slide title to jump to the beginning of the video interval corresponding with that slide. Any change in the current slide effects a change in the video and vice versa. It is also possible to browse slides without them being synchronized to the video.

When a viewer uses the “reverse slide” arrow button to return to a chronologically earlier slide, the slide deck ”de-attaches” itself from the video timeline, and users are free to browse the slides using the arrows while the video continues to play from its current position. We refer to this state as desynchronized mode. The view of the current slide does *not* update according to the time markers set throughout the video. In addition to the arrows, users can use the slide toolbox to jump back and forth between slides independent of the video, or drag the video timeline to seek to a particular position without affecting the slide.

Whenever the system has entered desync mode, the resynchronize button appears in between the navigation arrows (Figure 4.3). Clicking this button returns the user to the correct slide relative to the video's current location; additionally, it resets the system to synchronized mode so that the slides once again update according to the time markers. We will see that resynchronization is an important feature when we consider the multiple chain system implemented in this prototype.

4.2.3 Prototype 1

Using an early version of the conceptual chain model presented in the previous chapter, we modified our prototype to contain a basic mechanism for navigating across slides chained together in an archive.

As we see in Figure 4.4, a slide that is the origin of a chain can be used to navigate to other slides using the "Chain" button, which is typically greyed out, but becomes active when the current slide contains chain links to other slides. Clicking the Chain button takes the user to the endpoint slide of the chain, which as seen in our navigation model is a slide on a non-timed slide track. All other slides on this non-timed track can be browsed with the arrow buttons.

The navigation of the chain is designed to allow easy access to other slides on the corresponding non-timed track, yet retain a simple way of returning to the point in the main slide track where users entered the track. The system is assumed to be in desynchronized mode when on the track; to return to full synchronization, one must first use the "Home" button to return to the slide where one deviated from the chain (as well as reset the arrows to correspond with the main timeline track), and then subsequently press the resynchronization button to catch up with the video. The slide preview panes change upon entering chain mode to represent the slides on the corresponding non-timed track.

4.2.4 Chains of chains

For this prototype we experimented with various features concerning chains. Non-timed tracks linked to by chains could have slides with chains branching off of them as well to other non-timed tracks as well. We believed that this would allow for one to browse and descend a hierarchy of related slide materials. However, informal evaluations showed this to be too complicated for users to use effectively, which we discuss below.

We tried linking previous slides in a chemistry lecture archive back to

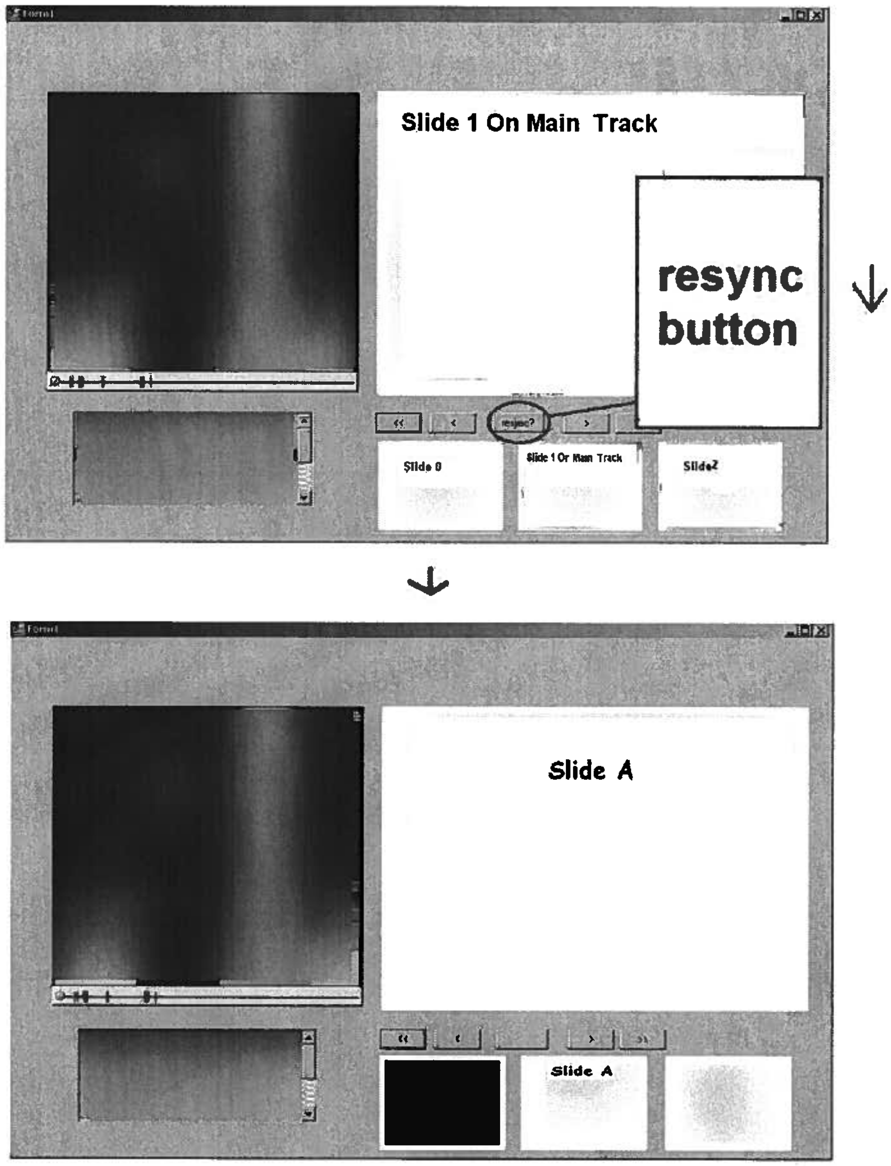


Figure 4.3: The resynchronization mechanism. Notice the slide panes are updated to reflect the actual location of the video.

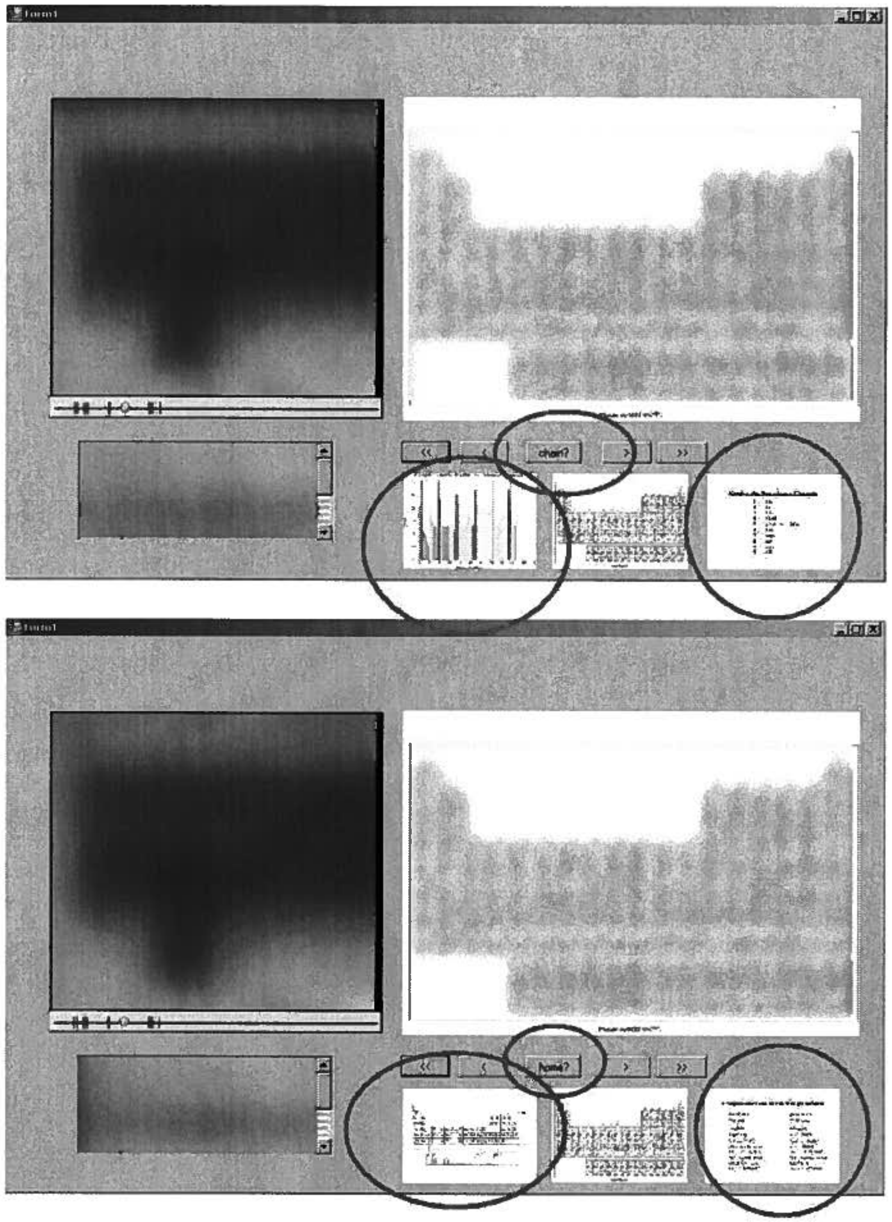


Figure 4.4: The chain button becomes available when the slide has a chain of related slides branching off of it. In this example, this slide of a periodic table has related tables on its branching chains that can be viewed (note the change in the slide previews) when the chain button is pushed.

previous slides with element tables for reference, as well as reference earlier equations at key points in the video. An informal test of the system revealed to us that users found the nested chain system used in the interface to be confusing, as the arrows were unintuitive to use across so many different levels. Users said they could find information they needed just by browsing in the main timeline track manually if they needed to, rather than struggling with two or more chain hierarchies, besides which they did not like the two-button method of having to resynchronize the archive.

Several uses for chains were given in Chapter 3. The model let us create examples of chains used to link related material or to provide alternates of slides. However, our informal tests failed to elicit a positive response with respect to the usefulness of multiply nested chains, as it was easier for the most part to find slides by manually desynchronizing the archive and searching through the slide deck. The commitment to browsing the non-timed track made it more difficult to return to the initial slide by having to leave the track AND resynchronize. Furthermore, because we did not make clear that the chained-to set of slides were actually another physical slide track, users had a hard time grasping the mental model behind our design.

This issue of pressing multiple buttons to "catch up" with the video was seen as a bit of a hassle, and chains were not seen as significantly more useful than individual browsing, especially since one could only enter the chain on specific linked slides (we did not take into account the necessity of viewing previous slides to supplement information being presented in the video itself, which would necessitate easier slide access from any point).

A visualization of chain semantics and organization within the archive would have been helpful (for example, a high-level view of the slides as a set of nodes in a graph), but we decided instead that it might be worth exploring a simplified implementation of the model that addressed some of the more basic problems we were encountering with our first prototype.

4.3 Issues to resolve

Before embarking on our third prototype we noted the particular issues that had arisen in development of the initial system. We consolidated the issues mentioned above into several key points to be resolved:

- A better way of organizing slide information for quick access: the context of information on a chained-to track relative to the linked slide was not always clear, and one could only access the chains through certain linked slides.

- A better navigation system and corresponding interface: the multiple buttons for having to enter chains and resynchronize the archive was seen as confusing, especially since it was not always clear if one was on a chain (other than the text changing on the button itself).
- An overall streamlined process for accessing chains that gave appropriate visual context as to the nature and location of the chain in the archive; as mentioned above this was not always clear, and appropriate affordances needed to be used to work with the track and chain metaphor.

From these new requirements we decided that it might be simpler to work with the more basic notion of parallel, independent tracks. We decided to try working with the idea that different tracks, could be used as a way of organizing slides according to different themes. This allowed us to provide a simpler working model for our actual interface and more tightly streamline and organize the slide information for viewers to use, while still maintaining some of the key concepts from our initial vision.

4.4 Prototype 2

As before, this iteration of the prototype was constructed in C#, and continued to use the RealPlayer and SMIL API's for video functionality.

4.4.1 Alternate Design

The basic structure of the interface remained relatively unchanged from the first two iterations. All the major widgets for basic navigation of the archive remain present (see Figure 4.5). Positioning and size were changed to create more real estate for the main slide pane and preview panes to occupy, and the buttons were made slightly more prominent. However, several key changes were made to reflect the new approach towards slides and tracks used in this iteration.

4.4.2 Tabs for Parallel Tracks

As mentioned earlier we chose not to stay with the chain metaphor and instead focused on the simpler concept of parallel tracks, which could be used to present any number of auxiliary uses for the slides independent of the video. While the main timeline track was still present as the default one, synchronized to the video and represented in the slide toolbox, we scrapped

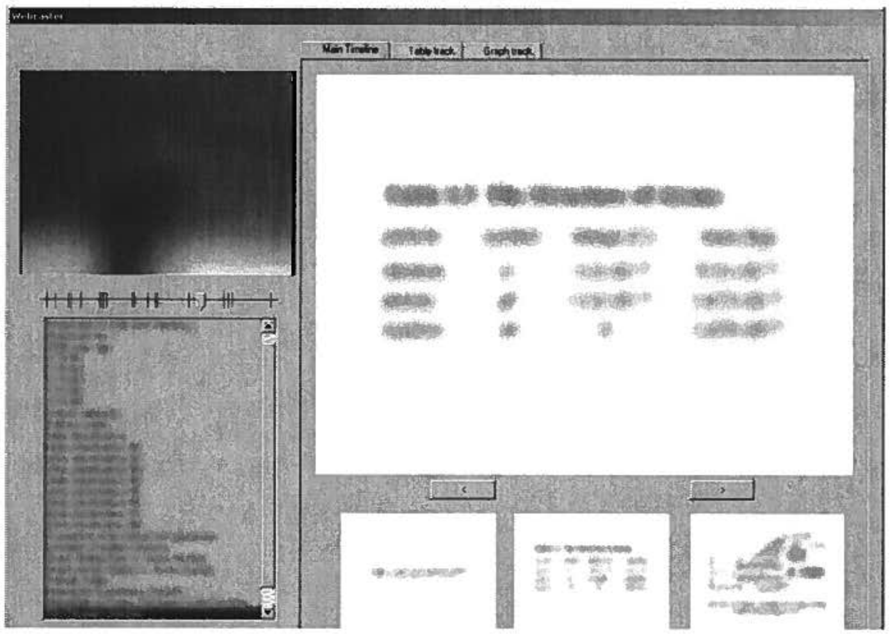


Figure 4.5: The third version of our prototype. Screen real estate has been allotted more efficiently and the interface has been streamlined to support extra tabs for tracks.

the idea of having linked chains and chose to represent contextually grouped material in the form of parallel slide tracks that could be accessed through the tabs at the top of the screen. This is a particular instantiation of the tabbed view model described in Chapter 3. While the tabs described in Chapter 3 let users see any combination of a single video and slide track, in this prototype we chose to restrict the differences between tabs only to slides, and have no switching between videos.

In accordance with our conception of multiple parallel slide tracks, slides included on these tracks are a subset of the entire slide 'library', and can be accessed at any point during the video (see Figure 4.6). Clicking on the tab puts you into the track of choice; the video is automatically assumed to be desynchronized with regards to the main track timeline while browsing the track, as with the chain browsing in the first prototype. However, instead of using a button to navigate on and off the chain in an unintuitive fashion, the tabs more accurately reflect the metaphor of the parallel track and there is much less confusion as to what track one is viewing when using the tabs, which have also been color-coded for redundancy.

Each individual track has its own title so that the context of what is inside is clear. We aimed for these tracks to be used as global mechanisms for organizing slides into meaningful groups, rather than as elements of local context that the chains were used for. The relative un-obtrusiveness of the tabs made it possible for different sets of information to be globally represented (such as a track representing all slides with equations, or all slides with pictures), without occupying large amounts of space or introducing an unwieldy navigation metaphor. We chose not to associate any timings with parallel tracks, as we wanted them to be separate aggregations of data rather than re-interpretations or re-timings of existing timed tracks.

One may click between auxiliary tracks and browse freely without worrying about how to resynchronize; returning to the main track is done by clicking on the main track tab and using the resynchronize button as before. The important difference from the earlier prototypes is that the essential metaphor of the navigation model is much more cleanly captured in the visual form of the tabs lined up at the top edge of the slide viewing pane. In informal followup sessions with users, we found that the concept of parallel tracks was much easier to grasp than chains, as the relative independence of each track kept them as distinct conceptual units.

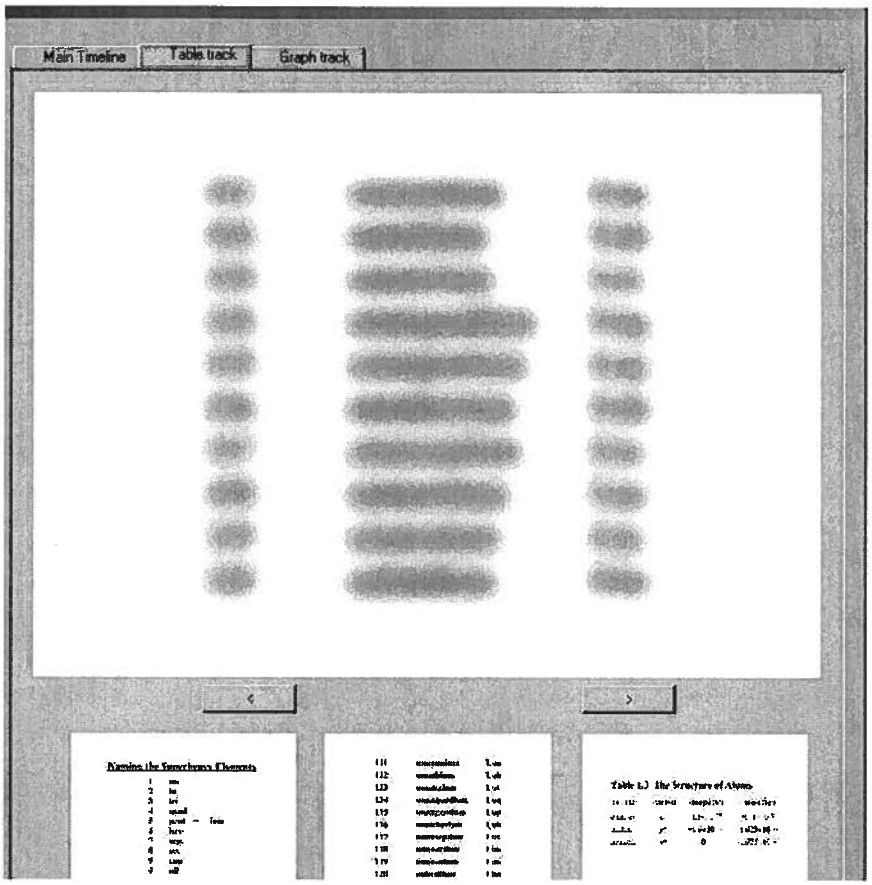


Figure 4.6: The different tabs at the top represent different slide tracks with related material within each one (e.g. tracks of tables, graphs). Redundant color coding reinforces the differences and gives users a sense of contextual awareness when they are browsing slides off the main timeline.

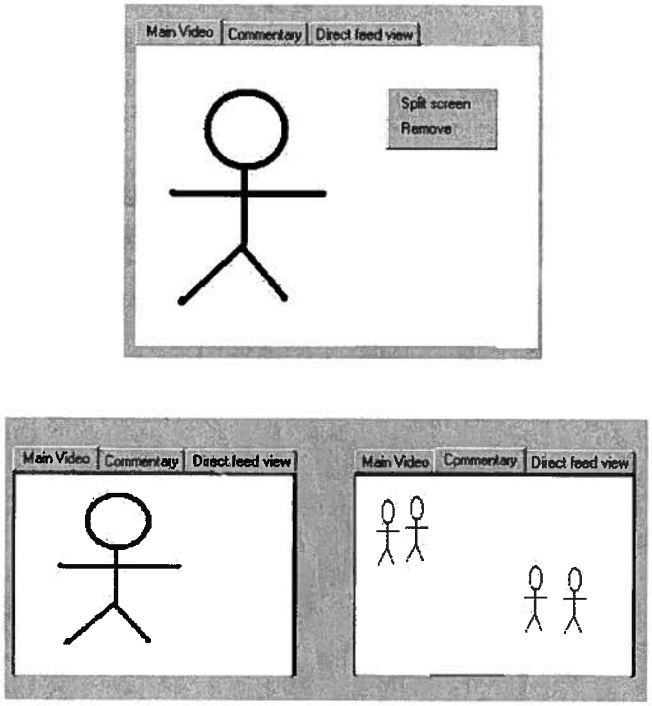


Figure 4.7: Inclusion of multiple video tracks would be one possible future development. We use tabs to switch between video tracks; when the "Split screen" function from the context menu is used, a new set of smaller windows is created, each window with its own video track tab.

4.5 Further possible developments

There are a number of features from the navigational model discussed in Chapter 3 that did not make it into our final prototype. However, our prototype was crafted to be flexible enough that including these features would not conflict with the general design of our system.

4.5.1 Multiple video and timed tracks

Implementing additional video tracks to choose from would be a relatively easy generalization of the current design our prototype uses. As described earlier in Chapter 3's navigational model, the extra video tracks would al-

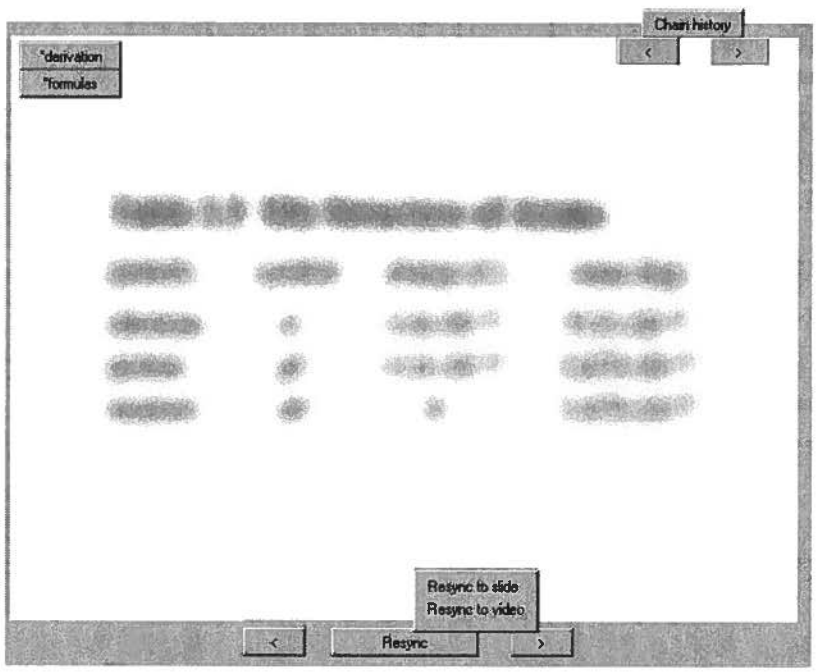


Figure 4.8: A view of the additional slide controls that could be included in future iterations of our prototype. We include the ability to choose which chains to follow from the slides, and the ability to resynchronize the slides to either the video or to the beginning of a particular slide timing.

ways be lined up and synchronized with each other (under the assumption that they have been appropriately padded with blank space to make them equal length), with one track being set as the default video. Figure 4.7 demonstrates how one might go about switching between video tracks in the archive, and also shows how one might use multiple video tracks at once for comparison purposes.

The navigational model also discusses the use of timed alternate slide tracks, a feature which we did not explore in our prototype. This would be as simple as adding extra slide tabs to our existing prototype, and repositioning or re-coloring them as necessary to make a visual distinction between the timed and non-timed groups. We could also generalize our tab concept as described in Chapter 3 and make each tab a single view of a combination of slide and video tracks. This would allow all combinations of video and slides to be represented in the tabs, as described in our model.

The desync button performs the desynchronize function exactly as described in the navigational model, causing all timed tracks to desynchronize themselves from the video(s). The resync button also performs as described in Chapter 3; we can see how this might be implemented in our prototype in Figure 4.8. By default, this button would resynchronize the slides to the current video location, but a context menu activated by right-clicking the resync button would allow users to resync to the beginning time of a particular slide on a particular timed track as well.

4.5.2 Chains

While our first attempt at building chains of slides in the prototype was difficult to understand, a chain system more true to our navigational model might be easier for users to work with.

The full navigational model describes a situation in which following a chain on a slide brings us to a particular slide on a non-timed track; Figure 4.8 shows us how such a conceptual operation might be implemented in our system. Clicking on one of the chain buttons in the top left corner of the slide, when it appears, would automatically switch tabs to the appropriate non-timed track and slide.

This is different from our early prototype where there was no context for where a slide on a chain came from; because there was no way to access a slide on a chain other than through the originating slide, users would feel as if they were "lost" because the slide did not actually exist on any track proper. However, with labels on chains and a clear indication of which track the user was on when following a chain (accomplished by switching the tab view to the appropriate track), these problems may be averted.

In future revisions, the issues that users were having with trying to understand the concept of chains within tracks that were already pointed to by other chains may be mitigated, as all slides on the endpoint of a chain would now belong to an actual non-timed slide track on a physical tab. This might give users a better sense of where they were in the space of slides and slide tracks.

4.6 User study

With our revised prototype, it was possible to conduct a formal experiment to gauge some of the questions we had concerning our webcast archive model. To this end we included experimental logging and timing features to our software, as well as mechanisms to disable and hide various features of the archive viewer. This enabled us to run a comparative study to evaluate the various new features of our prototype against a more traditional kind of webcast viewing system. This is described in more detail in the next chapter.

Chapter 5

User Study

A formal user study was performed to compare two different webcast archive viewing systems. The two systems used in our study were called the Simple viewing system, which was designed to replicate most of the basic archive viewing features available in ePresence, and the Multitrack system, which incorporates some of the enhancements described in our advanced webcast viewing model. A minimal viewing system equivalent to the streaming of a live webcast for the first time (i.e. no rewinding, pausing or the like), called the Linear Viewing System (LVS), was used as a control condition. Our first goal was to evaluate the as-of-yet untested assumption that users could more easily absorb information from an on-demand webcast archive with rewind and playback features than from a live streaming webcast that had no such affordances. The second goal was to determine whether adding multiple tracks of information to an on-demand webcast would be a useful addition to an archive viewing system.

We sought to answer the first question by comparing the Simple Viewing System with the Linear Viewing System. We sought to answer the second question by comparing the Multitrack system to the Simple system. Our hypotheses were as follows:

H1) People will absorb and comprehend information from webcast archives better using the Linear Viewing System to watch them than by using the Multitrack Viewing System. However, people using the Multitrack Viewing System will absorb and comprehend information better than those using the Simple Viewing System.

H2) People will be less satisfied with the Linear Viewing System in terms of system usability and functionality than those using the Simple Viewing System. However, people using the Simple Viewing System will be less satisfied with system usability and functionality than those using the Multitrack Viewing System.

H3) People will not be able to absorb and comprehend as much information per fixed unit of time using the Linear Viewing System when compared to people using the Simple Viewing System. However, people using the Simple Viewing System will absorb and comprehend less information in the same amount of time than those using the Multitrack Viewing System.

We can rephrase these hypotheses more tersely as inequalities:

H1) SVS $<_{comprehension}$ MVS $<_{comprehension}$ LVS

H2) LVS $<_{satisfaction}$ SVS $<_{satisfaction}$ MVS

H3) LVS $<_{comprehension/time}$ SVS $<_{comprehension/time}$ MVS

5.1 Methodology

To test our claims, we conducted a controlled between-subjects experiment in which participants watched an educational webcast in two different sessions, in one of three different experimental conditions. For the first session, participants in each condition watched the webcast with the Linear system. The three conditions differed in the viewing system used for the second session. The systems corresponding to each condition were the Linear, Simple and Multitrack systems respectively.

5.2 Participants

A total of 39 participants were recruited through university campus flyers, electronic mailing lists and an internal user study database (that notifies student registrants of HCI and psychology experiments seeking participants). Participants were all university students, and had a firm grasp of the English language, but had not taken any advanced university courses in the natural sciences.

5.3 Task

Individuals were asked to attend two experiment sessions spread out over the course of one week. In each session, participants were asked to view an archived webcast of an undergraduate classroom lecture on chemistry originally delivered at MIT [3]. They were then asked to complete a series of questions on the lecture material in each session. There were 14 questions worth a total of 24 marks introduced the first session, and 7 new questions worth a total of 9 marks introduced in the second session (see Appendix for the list of questions). Fractional marks were awarded for partially correct answers. Each participant was placed in one of the three experimental conditions that determined the viewing system for the second session.

All participants were informed of the experiment's monetary incentive scheme at the beginning of the first session. Participants in the control condition were told that they would receive an additional $25 honorarium if they achieved the highest score on the second questionnaire within their group of 10 participants. Those in the two experimental groups were told they would receive an additional $25 honorarium if they had the fastest time to finish the questionnaire among those who achieved a score of at least 80%. In the case of ties, the additional honoraria would be divided equally among those eligible within each group.

5.4 Procedure

We describe each step of our experimental procedure below.

5.4.1 First Session

In the first session, all participants were instructed to view the video of the chemistry lecture in its entirety, using the Linear system, which does not allow participants to skip or rewind video sections, or browse the accompanying lecture slides in any order different than the order presented in the webcast. This setup was designed to be similar to the viewing of a live webcast. The participants were given a quiz sheet containing 14 questions about material presented in the lecture. They were instructed to complete as much of the quiz as possible while watching the webcast. No additional time was allotted to answer the questions after the lecture was over. At the end of the first session, participants were reminded to return one week later for the second session of the study.

5.4.2 Second Session

Upon arrival for the second session, participants were given a copy of the original quiz sheet from the last session, along with an extra 7 new questions on the same material. They were asked to view the webcast again, using their assigned archive viewing system. Participants in the navigable conditions were asked to answer the questions as quickly as they could until they

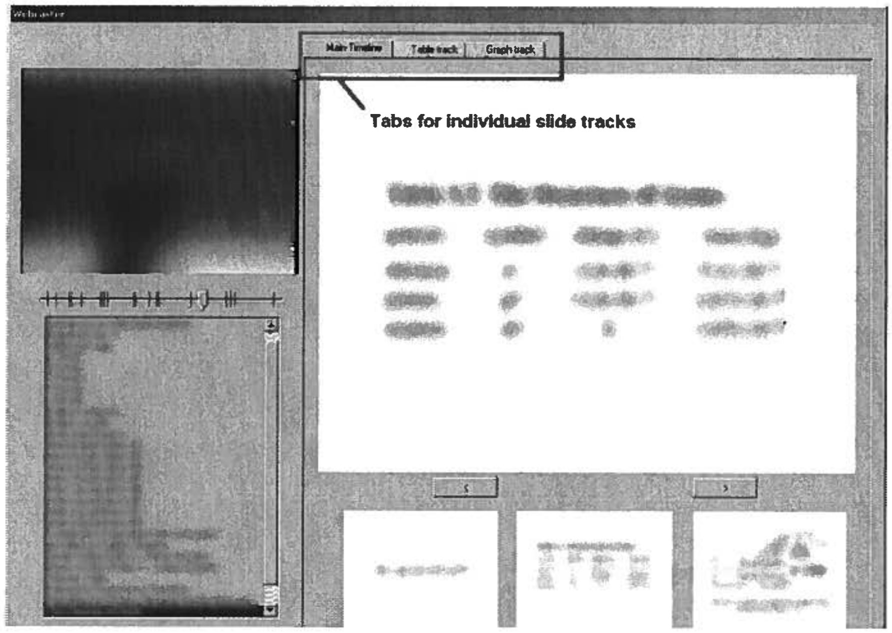


Figure 5.1: Screenshot of the Multitrack Viewing System (MVS) used in our experiment. Tabs for switching between different slide tracks are visible by the top left corner of the slide. The Linear and Simple systems lack this functionality as they contain no extra tracks.

felt they had achieved a score of 80 percent on the quiz (new questions included). Those in the Linear condition had no choice but to sit through the entire 40 minutes of the webcast because they were not allowed to rewind, pause, or shift through slides, as before. They were given a 42 minute time limit to achieve the 80 percent goal. The extra 2 minutes were allotted in that condition so that users could have a chance to finalize their answers.

Upon completion of the second viewing session, participants were given a short questionnaire to gauge their understanding of the software and to gather their opinions on different aspects of the webcast viewing experience.

5.4.3 Experimental conditions

We give some more detail on the particular features of each viewing system used in the second session.

- Linear System: this system was the control condition. The video was played from beginning to end with no pauses. There was no functionality for rewinding or backtracking. The slides advanced according to preset timings in the video. This was identical to the system used in the first session.

- Simple System: The video started at the beginning. Slides advanced according to preset timings in the video, but users were free to navigate the video by adjusting the timeline bar, or by double-clicking on the name of a section to jump to the point in the video where that section's slide first showed. Users could also use the arrow buttons to jump back and forth between adjacent slides as the video continued to play; however doing so de-synchronized the slide viewer, as in the ePresence system. This meant that the currently viewed slide would not update at the preset time markers while the user was viewing slides with the arrows, unless the re-sync button was clicked.

- Multi-Track System: All the features of the Standard system were present. However, as seen in our discussion in Chapter 4, two additional tabs were available which represented separate tracks of information; these tabs are visible in Figure 5.1. By clicking on these tabs, users could navigate an alternate set of slides using the arrow buttons; these alternate sets were effectively filters of the original slides set, displaying all the slides with tables, on all the slides with formulas, respectively. No slide timings were associated with these sets. For all intents and purposes the system was assumed to be in desynchronized mode while viewing these tracks. Returning to the main track put the system back in synchronized mode.

5.5 Measures

We recorded several measures in our experiment. The two primary ones are:

1. Quiz Score - For each of the two sessions we recorded the quiz scores of all participants.

2. Time Spent - For the second session, we recorded how much time it took users to complete the quiz to the specifications given (a self-estimated score of 80% or better).

We logged all actions performed by users (slide changes, video timeline jumps, track switches), and recorded users' opinions on different aspects of the webcast and of the system itself after the study. This information was elicited through questions answered on a 5-point Likert scale, as well as through questions asking for written comments.

5.6 Hypotheses

We refine our initial hypotheses in terms of our individual measures, by taking quiz performance as a metric for a viewer's absorption and comprehension of the webcast material.

H1: a) Participants in the Linear condition will achieve a better overall quiz score at the end of the second session than those in the Multitrack condition.

b) Participants in the Multitrack condition will achieve a better overall quiz score at the end of the second session than those in the Simple condition.

H2: a) Participants in the Simple condition will rate their system higher in terms of functionality, usability and overall satisfaction than those in the Linear condition.

b) Participants in the Multitrack condition will rate their system higher in terms of functionality, usability and overall satisfaction than those in the Simple condition.

H3: a) Participants in the Simple condition will achieve a higher ratio of quiz score to time spent watching the webcast, than those in the Linear condition.

a) Participants in the Multitrack condition will achieve a higher ratio of quiz score to time spent watching the webcast, than those in the Simple condition.

5.7 Results

We report first on the two main performance measures of quiz scores and completion times from the experiment itself, followed by the data collected from participants in the post-experiment survey.

		Linear	Simple	Multitrack	Max
N		13	13	13	-
Pre	Mean	12.84	12.65	12.46	24
	SD	3.31	3.05	3.69	-
Post	Mean	15.84	15.69	17.23	24
	SD	3.31	3.69	3.37	-
New	Mean	4.92	6.53	6.84	9
	SD	2.37	1.49	1.99	-
Time (min)	Mean	-	31.9	29.7	40
	SD	-	7.6	11.85	-
Ratio	Mean	.495	.722	.940	-
	SD	.116	.269	.562	-

Table 5.1: The five measures that were analyzed in an ANOVA. *Pre*: Scores from first session, on Q1-14 (maximum score: 24). *Post*: Scores from second session on Q1-14 (maximum score: 24). *New*: Scores from second session on Q15-21 (maximum score: 9). *Time*: Minutes spent completing quiz in second session. Note: Only 9 of the 13 participants in each of the navigable conditions were able to complete the quiz under the 42 minute time limit; we only consider the 9 in each condition who finished under the limit in the calculation of *Time*. *Rate*: The ratio of one's final score to time spent watching the webcast in the second session (which was automatically 40 minutes for those in the Linear condition). Formally defined as *Pre* + *Post* / *Time*.

5.7.1 Performance Measures

Overall scores

A one-way analysis of variance was performed on five measures. The first three measures were performance measures common to all conditions: the scores from the 14-question quiz administered in the first session (maximum possible score of 24), the scores on those same questions in the second session, the scores on the 7 new questions (maximum possible score of 9) administered during the second session. The last two measures were ones that pertained only to the navigable conditions: the amount of time spent watching the webcast in the second session, and the ratio of a participant's score to this same amount of time. The data varied across a single condition with three levels; this condition was the viewing system for the second session. Table tbl:Exp1 summarizes this data.

While there were no significant differences across conditions for the scores in the first session, or for those same questions in the second viewing session, a significant difference in scores for the new questions was revealed (F(2,36)= 3.521, $p < 0.04$). Contrast analysis on the ANOVA between the Linear and Simple System revealed a significant difference between the two (F(1,36) = 1.6154, $p < 0.045$); this significant difference holds true as well between the Linear system and the Multitrack system (F(1,36) = 4.9231, $p < 0.01$). There was a significant difference in the *Ratio* metric between the Linear and Multitrack systems as revealed by contrast analysis (F(1,36) = 2.886, $p < 0.012$). Users in the Multitrack condition were able to answer more questions correctly per unit of time than participants in the Linear condition.

It is interesting to note the differences in scores, with regards to the instructions given to each group. While users in the Linear condition were merely informed to do as best as they could in the 42 minutes alotted, those in the Simple and Multitrack system were told to quit as soon as they felt they had achieved a score of 80%. Users in navigable conditions, who had an upper bound dictated to them as part of the instructions, managed to attain an overall higher average score than those in the Linear condition, who were given no such limit but were told to do their best.

The average score at the end of the second session for participants in the Linear condition was 62%, compared to 67% and 72% in the Simple and Multitrack conditions respectively. While the data was not statistically significant, this phenomenon suggests that users in navigable conditions were able to outperform users in the Linear condition in terms of the efficiency with which they answered questions, since they were able to answer a higher proportion of questions correctly while subject to a time and score limit.

Overall times

We performed a Tukey two-tailed t-test to compare quiz completion times between users in the Multitrack and Simple systems, completion time being defined as the amount of time spent working on the quiz until stopping (participants were instructed to stop upon having achieved what they perceived to be a score of 80 percent or higher). While no significant differences were detected, a straight comparison of mean times indicated that those in the Multi-Track system finished 2.2 minutes earlier than those in the Standard system. We only considered the completion times of those who did not use the full 42 minute time limit to answer the quiz (9 participants in each condition were able to complete the quiz under the time limit).

Viewing habits

To further understand how the navigable systems (specifically, the MVS and SVS conditions) were used to answer questions, we referred to our recordings of which sections of the video participants chose to view in the second session. We compared participant performance on questions to the number of times they viewed the video intervals containing the information necessary to answer those questions. This was done to get a sense of how much relevant information participants could locate and absorb. We also wanted to see the effect that skipping video sections that contained quiz answers or repeatedly viewing the same section containing an answer might have on individual question performance.

For each video-based question in the quiz, we determined what we considered the smallest possible video interval that had to be watched to answer that question; in the case of questions whose answers lay in the slides, we marked down the time interval in which the necessary slide was showing. We then looked at participants' navigation logs to determine how many times participants "viewed" the answer to a question; if a specific segment viewed by a participant completely contained an interval corresponding to a video-based questions, we counted that as a viewing. For video intervals corresponding to slide-based questions, a viewing was counted if a viewed segment overlapped in any manner with (without necessarily containing) such an interval. Any specific navigations to the slide in desynchronized mode were also counted.

		OldSeen		NewSeen		UniqueTime	
	N	Mean	SD	Mean	SD	Mean	SD
Simple	13	67.5%	17.8%	67.03%	18.8%	17.25	7.51
Multitrack	13	66.6%	22.8%	66.6%	12.7%	18.03	7.58

Table 5.2: *OldSeen*: The percentage of answers to questions from the first session, that were seen again by the participant in the second session. *NewSeen*: The percentage of answers to new questions in the second session, that were seen by the participant. *UniqueTime*: The total length, in minutes, of unique footage seen by a participant in the second session. As the term "unique" suggests, any given interval of the video, if watched more than once in the second session, is still only counted once in this total.

Data relevant to this analysis can be seen in Table 5.2. While a Tukey test revealed no statistically significant differences between the Multitrack and Simple conditions, we nonetheless were given interesting insights into the viewing habits of participants.

Participants in the navigable conditions spent around 30 minutes watching the webcast as seen in 5.1, but they only viewed about 17-18 minutes of actual "unique" video. Judging from these numbers, at least a third of most participants' time was invested in re-watching segments of video they had already seen, whether this was intentional or not.

Sections that were frequently rewatched did not vary much between the two conditions, as seen in Figures 5.2 and 5.3. For example, in both conditions, questions 6 through 10 were watched very few times relative to other questions such as 3, 5, and 20.

5.7.2 Subjective measures

The post-experiment survey revealed several interesting trends. Participants were asked questions on a 5-point Likert scale to get a sense of their overall experience with the software. A one-way ANOVA was used to evaluate the questions common to all conditions, while questions that were only given to those in the navigable conditions (i.e. questions concerning usage of navigation controls) were evaluated using a Tukey t-test. While a one-way ANOVA failed to reveal any statistically significant differences, there were interesting observations to be made of the survey data.

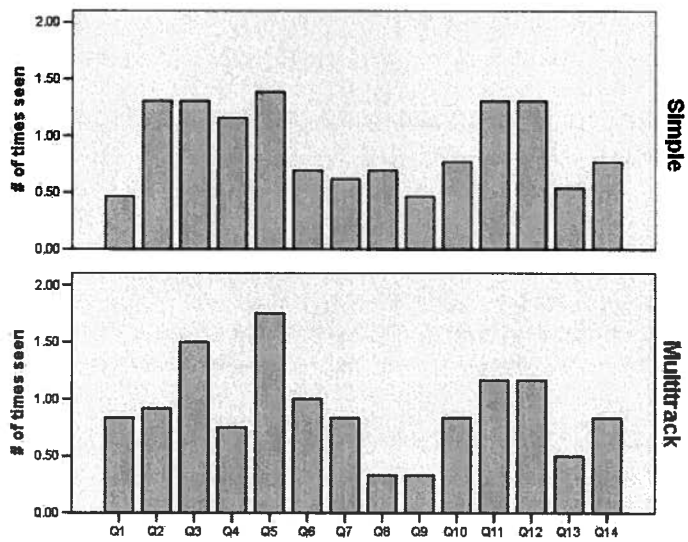


Figure 5.2: The average number of times the answers to questions 1-14 were seen during the second session, for the Standard and Multi-Track conditions.

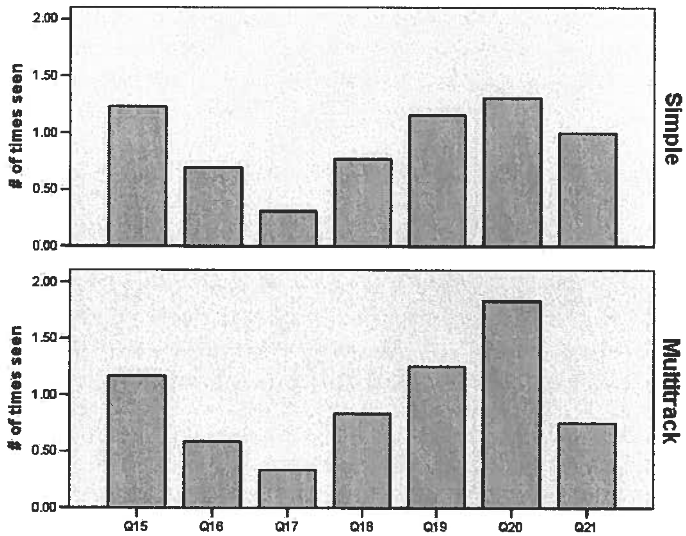


Figure 5.3: The average number of times the answers to questions 15-21 were seen during the second session, for the Standard and Multi-Track conditions.

Ease of use

Ratings were roughly equal in terms of overall ease of system usage between conditions (see Figure 5.4), but there appeared to be a difference between participants in the Multitrack and Simple systems when they were asked to rank the ease of navigation of the video, as well as for the ease of navigation of the slides (1 = very hard, 5 = very easy). A greater proportion of participants in the Multitrack condition ranked their experience as a 4 or higher both for the video and for the slides than those in the Simple system. As seen in Figures 5.5 and 5.6, there were also more participants who gave low ratings (2 or under) to the navigation controls in the Simple system than in the Multitrack system.

Usefulness of slides versus video

When participants were asked how useful the information contained in the video and slides were individually, participants in the MVS and LVS conditions consistently ranked the slides as containing useful information (a score of 4 or more), while those in the SVS condition were divided in their opin-

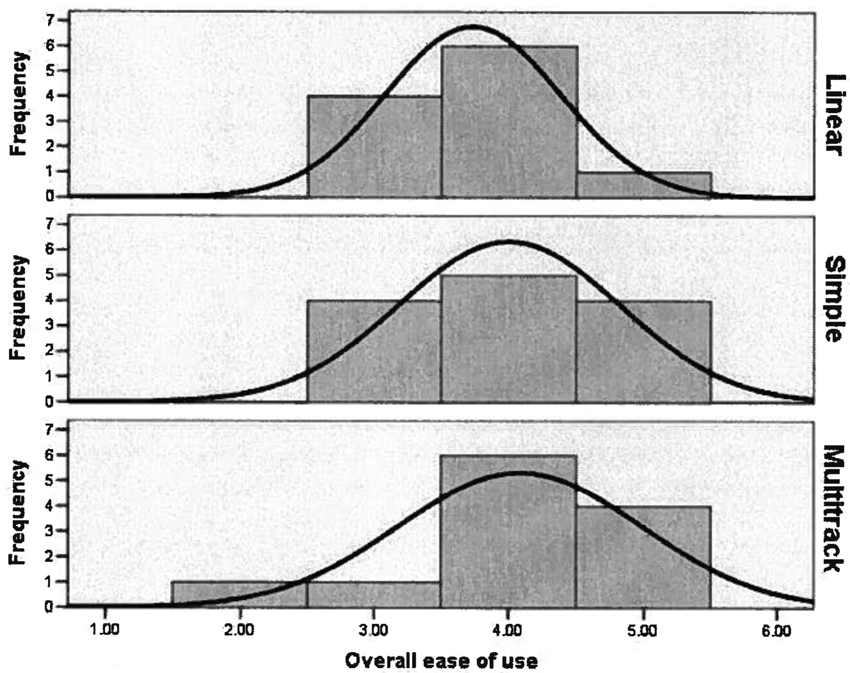


Figure 5.4: Likert scale of participant impressions of the overall ease of use of the software; 1=very difficult to use, 5=very easy to use. A one-way ANOVA showed that differences were not statistically significant ($F(2,36) = .618$, $p = 0.545$).

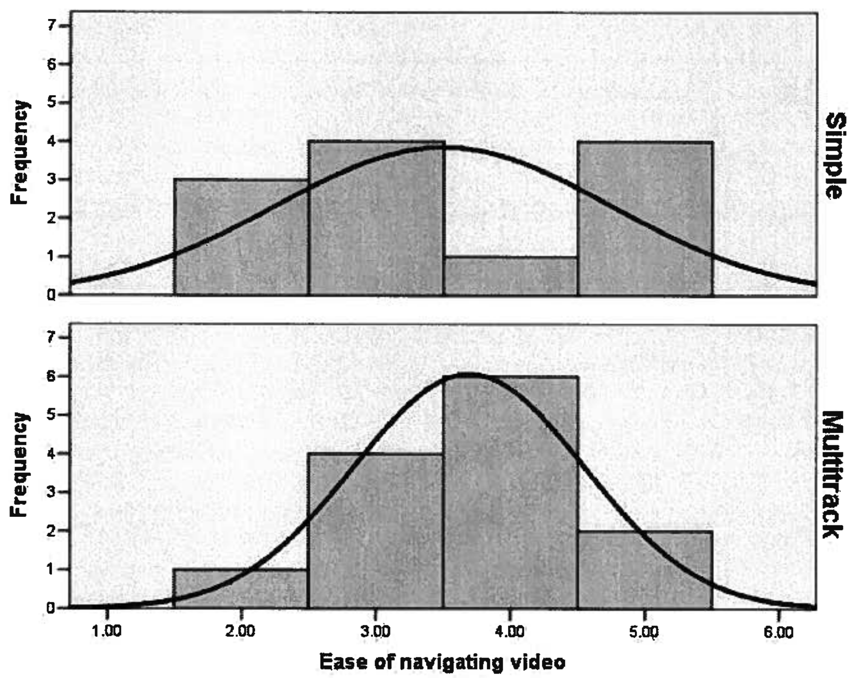


Figure 5.5: Likert scale of participant impressions of the ease of use of the video navigation controls; 1=very difficult to use, 5=very easy to use.

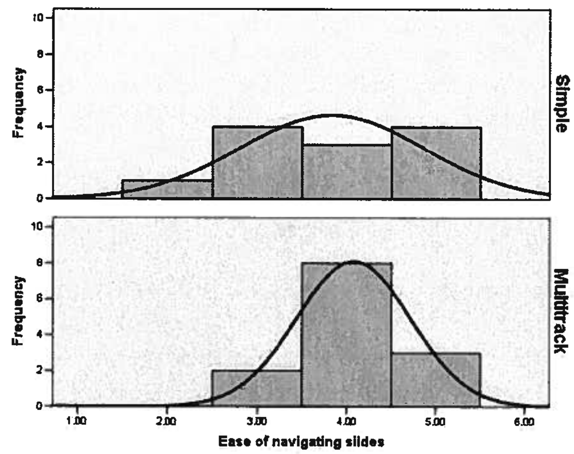


Figure 5.6: Likert scale of participant impressions of the ease of use of the slide navigation controls; 1=very difficult to use, 5=very easy to use. A Tukey t-test showed that differences were not statistically significant ($t(24) = .378$, $p = 0.709$).

ions. Those in the Linear system consistently ranked the video as being only moderately useful from an informational standpoint, relative to the scores for the other systems (Figures 5.7 and 5.8).

Usage of multiple tracks

The majority of participants found it easy and intuitive (score of 4 or higher) to navigate the multiple tracks present in the Multi-Track system (see Figure 5.9). While opinion on the ease of finding information in the different tracks was not as uniform, the majority still indicated a score of 4 or higher (see Figure 5.10). Participants found the resynchronization mechanism easy to use and understand. Only a single participant expressed any kind of difficulty with using the resync controls after having browsed a parallel track.

Difficulty of questions

Generally, the perceived difficulty of the quiz questions was relatively uniform across conditions, although it was in the Linear condition that the most number of participants rated the difficulty of the questions in the first session as tending towards the harder side of things (a rating of 4 or higher,

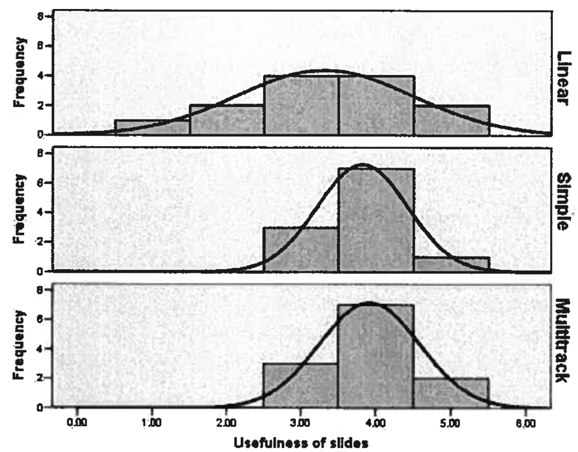


Figure 5.7: Likert scale of participant impressions of how useful the information in the slides was in answering questions; 1=not useful at all, 5=very useful. An ANOVA showed that differences were not statistically significant ($F(1,36) = 1.745$, $p = 0.190$).

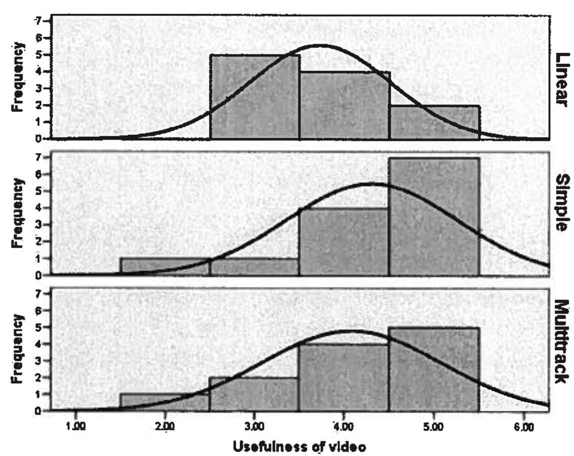


Figure 5.8: Likert scale of participant impressions of how useful the information in the video was in answering questions; 1=not useful at all, 5=very useful. An ANOVA showed that differences were not statistically significant ($F(1,36) = 1.197$, $p = 0.315$).

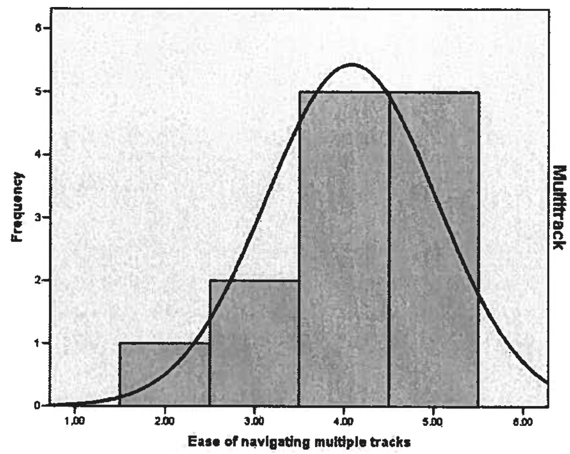


Figure 5.9: Likert scale of participant impressions of the ease of navigating the tabs representing tracks in the Multi-Track system; 1=very difficult, 5=very easy.

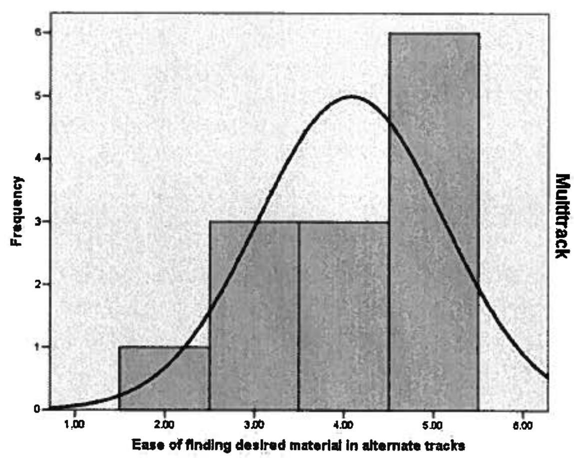


Figure 5.10: Likert scale of participant impressions of the ease of actually finding information in the separate tracks of the Multi-Track system; 1=very difficult, 5=very easy.

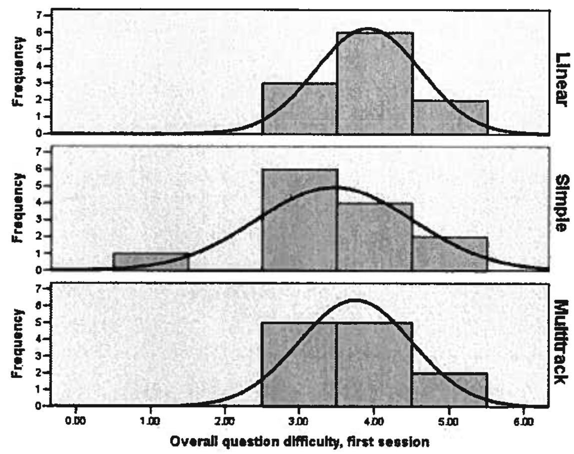


Figure 5.11: Likert scale of participant impressions of the overall difficulty of the questions in the quiz's first session; 1=very easy, 5=very difficult.

5.11). Participants were also asked to gauge their performance on the different sets of questions in each session. They were given 5 intervals of scores, represented by the numbers 1 through 5 on a Likert scale, and asked which interval they thought their final score would belong to (the lowest interval was a final score between 0-20%, and each interval was established in 20% increments, the highest being a final score between 81-100%). Those in the Multitrack system were most confident in their scores, as the average final score estimate of a Multitrack participant, was 4.08 on our Likert scale, compared to 3.58 and 3.73 for the Linear and Simple systems respectively. On the other hand, Multitrack participants were also the furthest off with their estimates; on average, their actual final scores lay 0.75 intervals (i.e. roughly a 15% difference in score) less than what they had estimated, while the Linear and Simple participants' final scores were 0.74 and 0.66 intervals less than their estimates, respectively.

5.8 Observations / Discussion

We now turn to discuss notable trends and other interesting observations of our study results, as well as their possible implications.

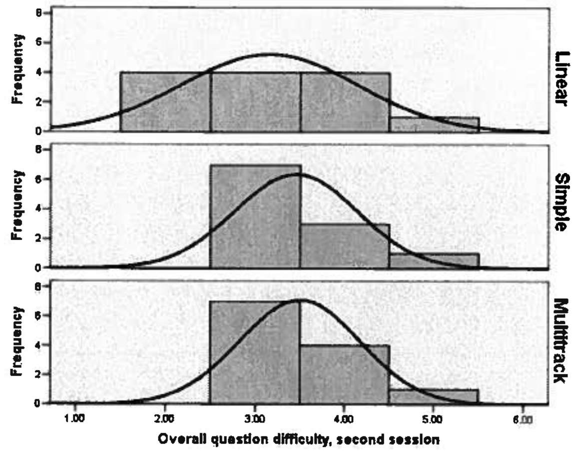


Figure 5.12: Likert scale of participant impressions of the ease of actually finding information in the separate tracks of the Multi-Track system; 1=very difficult, 5=very easy.

5.8.1 Question Order and Performance

Eight of the quiz questions in the first session were numbered "in order" relative to the webcast; that is, the quiz questions were numbered according to the chronological order of their answers in the webcast (video and slides). However, a few quiz questions were intentionally placed "out of order". In general, it appeared that "late" questions (questions that appeared much later in the quiz than when their answers were presented chronologically in the webcast) were more difficult to answer in both the first and second sessions than were "early" questions (those that appeared much earlier in the quiz than when the answers showed up in the webcast). This is supported by the graphs in Figure 5.13 which demonstrate the skew towards low scores across all conditions for "late" questions.

Interestingly, scores for "late" questions appeared to improve to a greater degree for participants in the Multi-Track and Standard systems during the second session, perhaps because it was then possible for them to backtrack to find answers to their questions (as seen in Figures 5.15 through 5.18, the number of participants with non-zero scores on Q11 and Q14 increased much more in the navigable systems relative to the Linear system). There were no major improvements in scores for "early" questions. It is possible that participants performed well on the questions because seeing questions

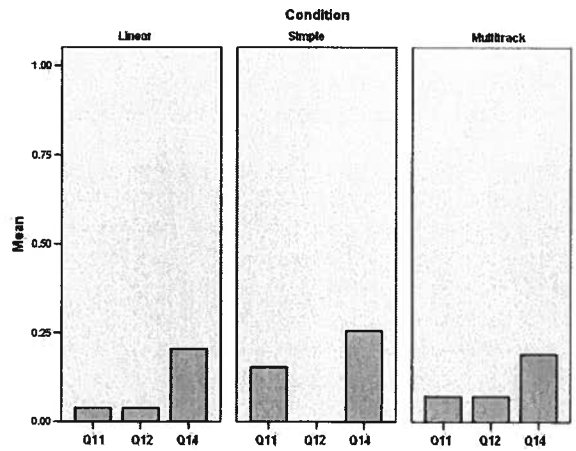


Figure 5.13: Scores on individual "late" questions, as percentages, across each of the three conditions.

first before the answers during the first session was not nearly as hard to deal with as the reverse situation for "late" questions.

5.8.2 Time to completion

Although no significant differences in time to completion between the Simple and Multitrack systems were detected, the absolute best times were all achieved by participants in the Multitrack condition (Figure 5.19); these four participants all achieved completion times under 18 minutes. These users might have been able to complete the second session faster due to having good scores in the first session, nevertheless this hints at the possibility that having quick access to different tracks of information (in this case, tracks containing graphs and tables only) may indeed make information acquisition easier for those who already know what they are looking for. For example, question 14, which was a graph-based question, was answered with better scores in the Multitrack condition than in the other two conditions, which supports the claim that the graphs were easier to access and find in the Multitrack system.

5.8.3 Track usage

We found an interesting phenomenon concerning the usage of the navigation tools in the Simple and Multitrack systems. While participants in the

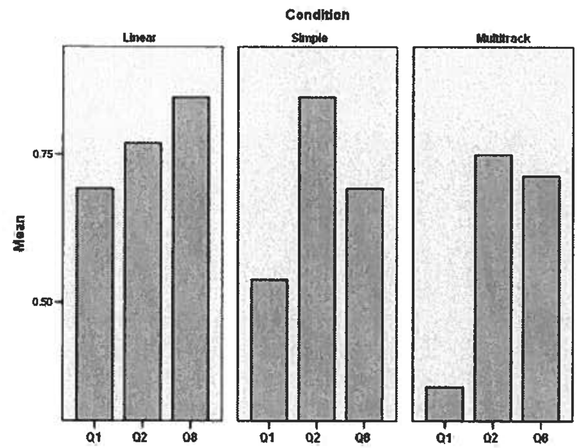


Figure 5.14: Scores on individual "early" questions, as percentages, across each of the three conditions.

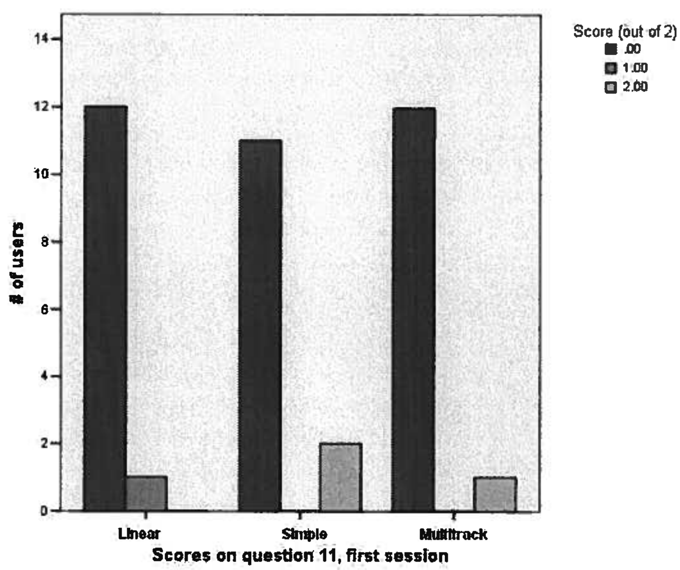


Figure 5.15: Histogram of scores (out of a total of 2) on Question 11 in the first session, for each condition.

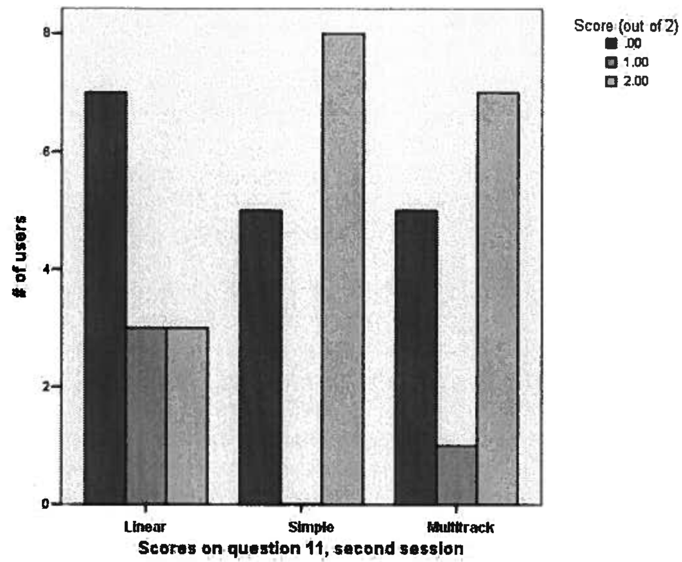


Figure 5.16: Histogram of scores (out of a total of 2) on Question 11 in the second session, for each condition.

Multitrack condition, on average, used the arrows for navigation much more than those in the Simple condition, they also relied less on the toolbox and timeline navigation bars than those in the Simple system. Users spent appreciable amounts of time browsing the other tracks in the Multitrack system (an average of 3.5 minutes per participant); this suggests that users were able to find information relevant to answering the slide-based questions there, which probably reduced the amount of timeline jumping (which those in the Simple system were essentially forced to do to find the same material).

The standard deviation for the number of jumps was much greater for the Simple system than for the Multitrack system. While some participants skipped to places all over the timeline in the Simple system, others did relatively few jumps; by contrast, the number of jumps in the Multitrack system tended to stay in a relatively well-contained range. This seems to suggest that there was little deviation in the strategies used to find answers to questions, as the multiple tracks were essentially invitations to participants to find information pertinent to their task without having to jump to different sections in the video. As many participants had given incorrect answers on graph-related questions in the first session, it is possible that their search

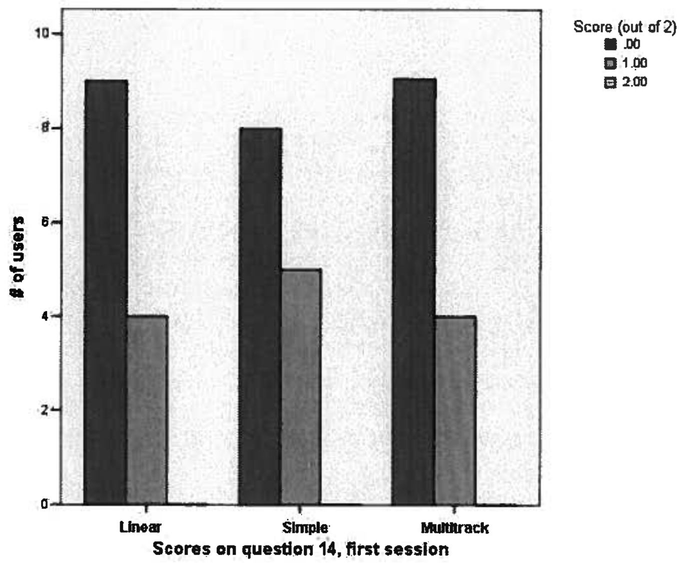


Figure 5.17: Histogram of scores (out of a total of 2) on Question 14 in the first session, for each condition.

		Simple	Multitrack
Forward	Mean	20.07	38.77
	SD	15.45	24.04
Backward	Mean	18.5	22.08
	SD	16.44	11.3
Timeline	Mean	36.75	26.77
	SD	25.81	12.89
Toolbox	Mean	36.17	29.15
	SD	32.31	17.4

Table 5.3: *Forward*: Number of clicks on forward nav button. *Backward*: Number of clicks on backward nav button. *Timeline*: Number of jumps using video timeline. *Toolbox*: Number of jumps using slide toolbox.

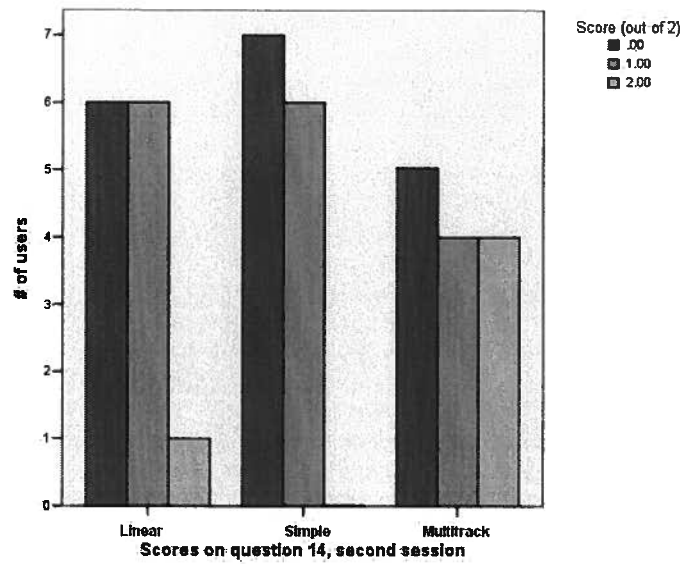


Figure 5.18: Histogram of scores (out of a total of 2) on Question 14 in the second session, for each condition.

strategies were largely expedited since temporal order was no longer the only criterion they had to navigate the webcast.

Most participants tended to use the graph and equation tracks for short bursts of 1-2 minutes before switching back to the standard track. The alternate tracks were used on average about 2.4 times a session to find specific pieces of information. Time spent browsing graph and equation tracks generally demonstrated short intervals of intense slide interaction. This homogeneity in track-browsing strategies suggests that information tracks tailored for a specific purpose promote more active searching and directed data gathering than does the passive activity of watching a single track webcast.

5.8.4 Question Answering Strategies

Questions whose answers could be found in the video tended to be answered equally well across all conditions. The histograms of questions 5, 8 and 10 (Figures 5.20, 5.21, and 5.22) all of which had their answers in the video, indicate that the number of participants who answered each question correctly was roughly the same across all conditions. Although participants in the Simple and Multitrack systems had the capability of navigating back

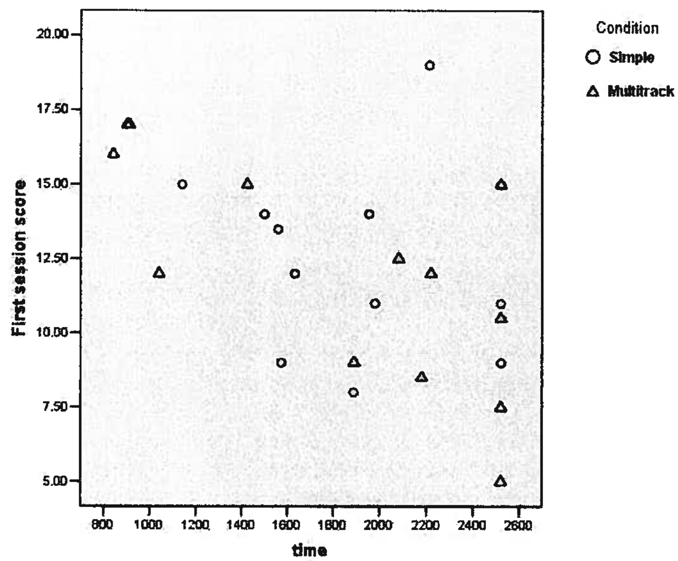


Figure 5.19: Scatter plot of final scores versus completion times for the two navigable conditions.

and forth in the video, the large amounts of video they had to track through meant that their search strategies could not narrow down the probable locations of the relevant information to an appreciably small interval. This is reflected in the fact that logs of timeline traversal for these conditions tended to stay in place, with relatively long intervals of 2-3 minutes between actual jumps. Thus, this strategy turned out to be not much better than watching the entire video through as in the control condition, because relatively large segments of video had to be watched to answer all of the questions.

In contrast, slide-based questions were much more frequently answered correctly in the navigable conditions than in the Linear condition (see Figures 5.23 through 5.26 for notable examples in questions 15, 18, 20, 21). The ability to navigate the slides, which were much more informationally dense than the video itself, made it possible to access a lot of information and collect important numbers / notes from slides in a short period of time. A common strategy was to browse ahead and back in the slides while listening to the main webcast when trying to ascertain whether answers could be found in that section. While there are no immediate benefits to being able to navigate videos for answers (scores were not significantly better for

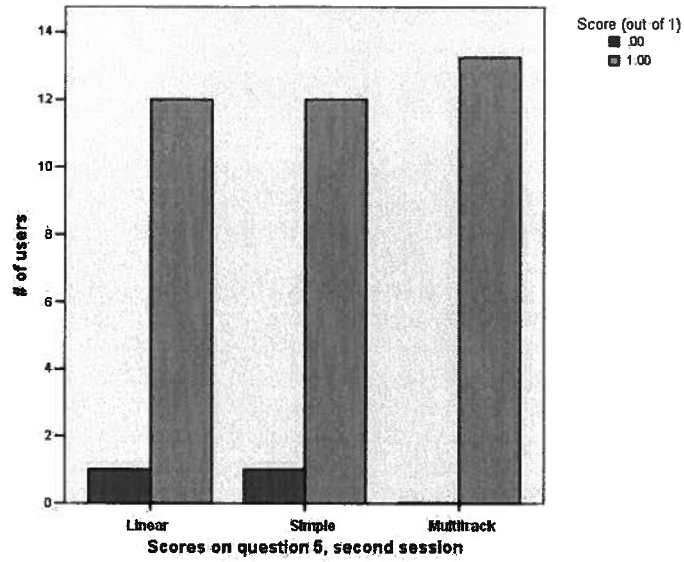


Figure 5.20: Histogram of scores (out of 1 total) on Question 5, second session.

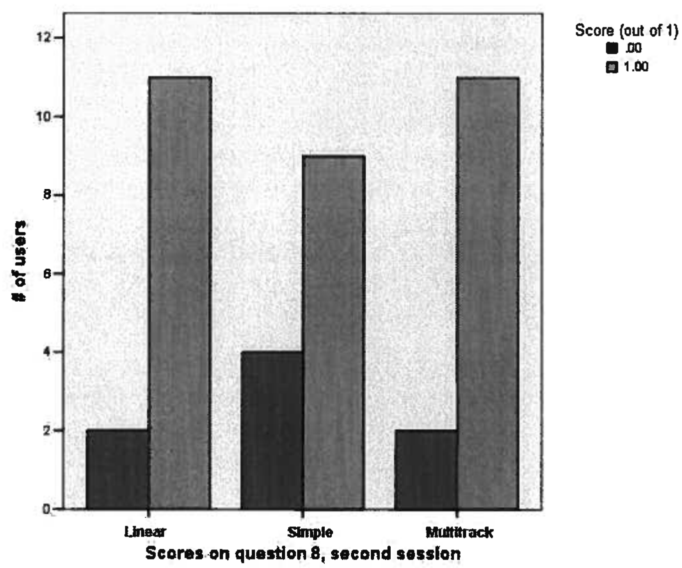


Figure 5.21: Histogram of scores (out of 1 total) on Question 8, second session.

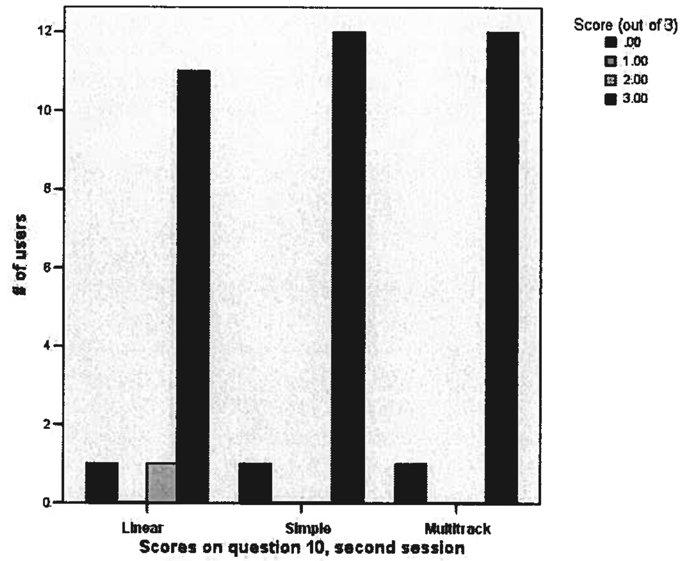


Figure 5.22: Histogram of scores (out of 3 total) on Question 10, second session.

video-specific questions in any of the conditions), there exists the indirect benefit of being able to multi-task and browse slides while watching video in the navigable systems, which is important to quickly understanding the context of a chunk of a webcast.

There were two situations in which performance and improvement were much more noticeable in the Linear system than in the Simple system (and no worse than in the Multi-Track system). More people in the Linear system were able to answer questions 6 and 16 correctly than either of the other two systems. The answer to these questions were found near the beginning and end of the webcast,respectively, and were comprehension questions that required watching of the video to derive the answer. This suggests that a) people in the navigable conditions tended not to view much video at the beginning or the end of the video (this becomes a more plausible situation when we recognize the rapid jump in action past the 4 minute mark, and the bulk of interface activity in the middle of the webcast), and b) that comprehension questions were possibly easier to answer when participants were forced to revisit those concepts by watching through the entirety of the webcast rather than being allowed to visit selective chunks. The first suggestion seems to be strongly reinforced by Figure 5.27, which shows that partici-

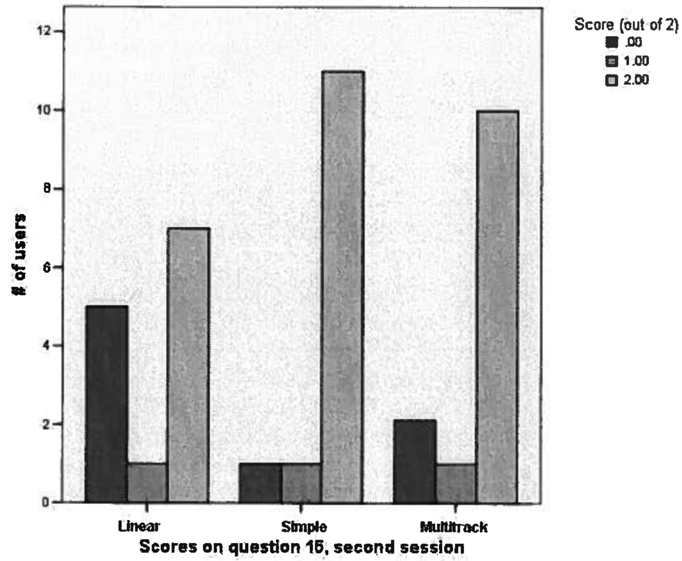


Figure 5.23: Histogram of scores (out of 2 total) on Question 15, second session.

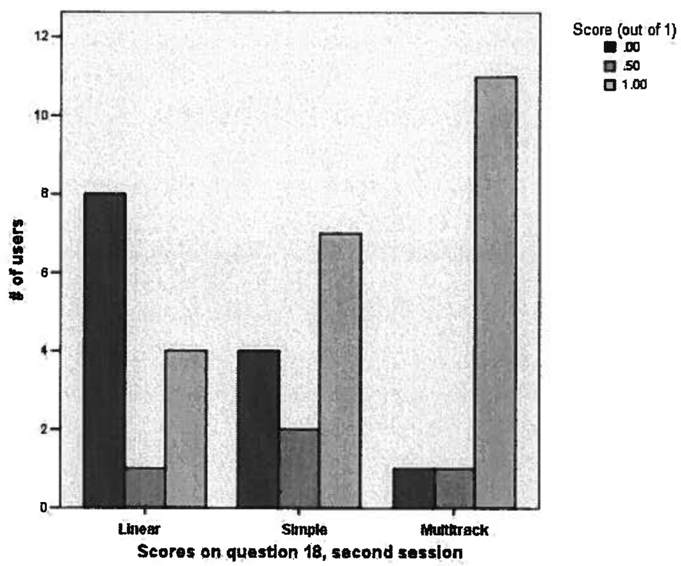


Figure 5.24: Histogram of scores (out of 1 total) on Question 18, second session.

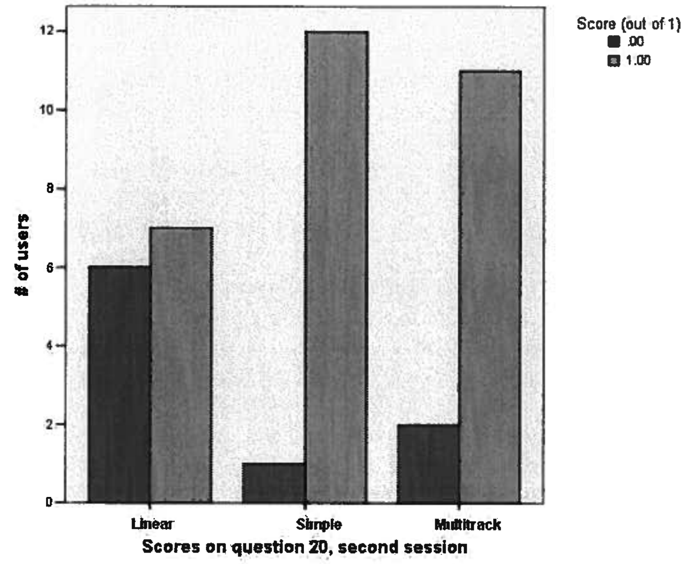


Figure 5.25: Histogram of scores (out of 1 total) on Question 20, second session.

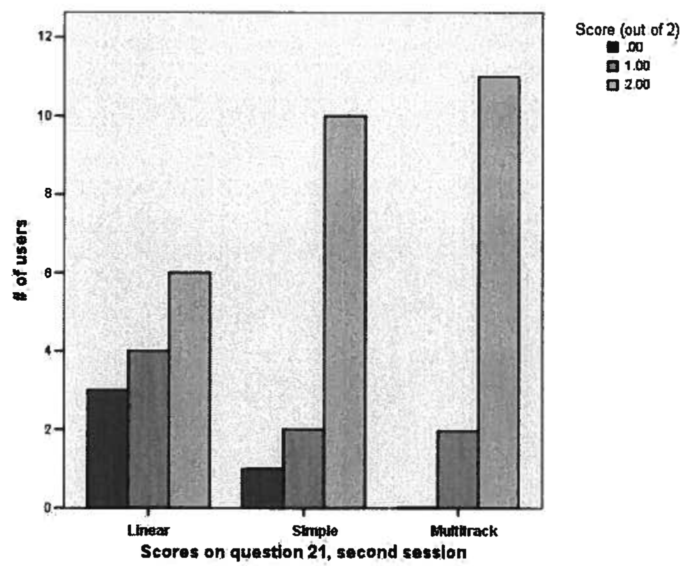


Figure 5.26: Histogram of scores (out of 2 total) on Question 21, second session.

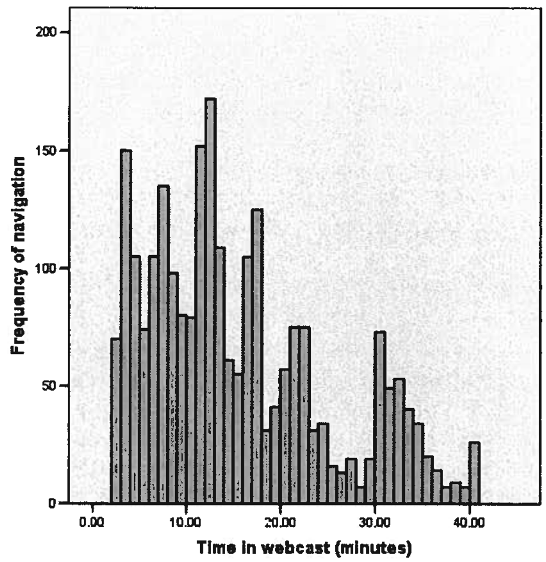


Figure 5.27: The temporal destinations of timeline jumps performed in navigable conditions, represented as a histogram.

pants in the navigable conditions tended not to navigate to the beginning and end of the webcast relative to other sections.

Effects of viewing answers

While the points mentioned above seem to favour the Standard system for certain question answering situations, we make a note of some instances in which the process of actively navigating to find quiz answers may have made the navigable systems a better choice for finding and absorbing the material necessary to answer those specific questions.

It is useful to compare how many times participants viewed the answer to question 9, compared to questions 11 and 12. Question 9 was a comprehension question that required the viewer to draw out two key points in a 2-minute long interlude, while questions 11 and 12 were mathematical calculation questions that required the participant to do the necessary computations using numbers given in the video and slides. Uers in all conditions exhibited relatively poor performance on these questions both in the first and second sessions (Figures 5.15 and 5.16). While participants in the navigable conditions watched the section containing the answer to question 9 an extremely low number of times on average (25% of participants in both navigable conditions skipped over the section and did not watch it at all), yet many participants watched the sections containing answers to questions 11 and 12 at least once (see Figure 5.2).

This suggests that the answers to questions 11 and 12 were easier to find in the webcast than the answer to question 9. This is perhaps not surprising if we consider that the slide titles hint to their topics, and questions 11 and 12 required a combination of slide and video information to be answered; by navigating to the locations of slides that might contain relevant information, one can easily explore the surrounding video as well. Because question 9 was purely video-based and there were very few cues or hints in the question that could point to the possible location of the answer relative to the topics listed in the slides, it was that much more difficult to locate its answer.

However, those participants using the Linear system definitely saw the answer to question 9 by virtue of watching the entire webcast, yet did no better on the question than did participants in the navigable conditions. Of the 7 participants in the Linear system who did not answer question 9 in the first session, 5 were still unable to answer it in the second session despite *seeing the section containing the answer again*, compared to 1 of 7 participants in the Simple system and 2 of 5 in the Multitrack system. It is possible that participants in the Linear condition simply ignored the information or did not recognize its importance to answering question 9.

We theorize that in the navigable conditions, participants may have went with a strategy of actively seeking out answers to questions by navigating to those sections that might contain them, and that their slightly better performance may have been a result of this strategy being more effective than viewing the totality of video information presented in the Linear condition.

Comparing the number of "unique" minutes watched in the video to the total average time spent watching the webcast in the navigable conditions (Tables 5.2 and 5.3), we see that participants' time was used more efficiently than in the navigable conditions than the Linear condition. Participants in the Linear condition watched a total of 40 minutes (excluding the additional 2 minutes given at the end to finalize answers). Participants in the navigable conditions spent about 30 minutes total, watching around 17 minutes of unique video. Only 12 minutes of the webcast contained information relevant to the answers, so it would seem that navigation made the viewing experience more efficient time-wise: more time could be spent rewatching those sections containing the answers once the sections were identified.

We believe that absorbing and reproducing quiz answers depended not only on participants seeing the answers, but also on them being able to

	Simple	Multitrack
Questions 1-14		
Saw/right/pre	105	114
Saw/right/post	21	21
Saw/wrong	35	26
Missed/right	4	7
Missed/wrong	17	14
Questions 15-21		
Saw/right	43	52
Saw/wrong	20	16
Missed/right	5	6
Missed/wrong	23	17

Table 5.4: Sums of questions answered across participants and whether or not their answers were seen, for each navigable condition. *Saw/right/pre*: Participant answered the question for at least partial marks in the first session. *Saw/right/post*: Participant answered the question incorrectly in the first session, saw its answer in the second session and answered it at least partially correct in the second session. *Saw/wrong*: Participant answered the question incorrectly in the first session, saw its answer in the second session and did not get any marks for it the second session. *Missed/right*: Participant answered the question incorrectly in the first session, did not see its answer in the second session and answered it at least partially correct in the second session. *Missed/wrong*: Participant answered the question incorrectly in the first session, did not see its answer in the second session and did not get any marks for it the second session. Note: Questions 15-21 were only shown in the second session; *Saw/right* corresponds to the situation in which the participant saw the answer to the question and got at least partial marks for it.

recognize specific pieces of information as answers through repeated viewing; using active navigation and searching, participants could devote much more attention to finding appropriate information and recognizing it when it appeared.

There were extremely few instances of participants being able to answer a question in the second session without having navigated to the section containing the answer during that same session. The only question in which more than one participant in any navigable condition was able to do this was question 16, a 2-mark question that could only be answered by watching the video. One participant in the Simple condition and 4 in the Multitrack condition correctly answered the question without having viewed the relevant section of the video.

Question 16 asked viewers about the shape and properties of a graph displayed on the board. While question 16 was not among the questions asked in the first session, roughly 3 minutes of time in the webcast were devoted to discussing the theory behind the generation of said graph, with the contents of the graph itself in plain view, near the end of the video. We theorize that these few users were able to answer this question without having to navigate to the section because of the fact that since the graph was in plain view near the end of the video, users still remembered the contents of the graph as the last thing they saw from the first session.

5.8.5 General Impressions

All participants in the navigable conditions unequivocally stated that the second session in which they could navigate was much preferred for answering questions and gathering information. "Being able to [proceed at one's] own time, own pace" was important to one user. Furthermore, the ability to simultaneously process information through different channels was much appreciated, and as one user put it, "[it was beneficial] to be able to listen to the lecture and watch graphs at the same time". On a related note, another user expressed satisfaction at being able to "go back and forth as well as change the slides while [the] video is going". Many reasons were given for preferring navigable archives over non-navigable ones; the ability to re-watch scenes to confirm information was important, as well as the freedom to search for specific answers.

Users in the Linear system condition often expressed dissatisfaction and complaints with having to sit through unnecessary chunks of the webcast and the inability to pause or revisit slides, as mentioned earlier. A number of users also expressed dissatisfaction with having to sit through the same material twice with no way of speeding up the process, and having to revisit questions that had already been answered, using such terms as "boring" and "annoying" to describe the situation.

Another important consideration with navigable systems versus the Linear system was that the lack of a pause button in the Linear system made it hard to write answers down while watching in real time (one user vented frustration over the fact that "[she] could not pause the lecture to read over the questions first before answering them"), especially if they suspected that more answers were coming up quickly in rapid succession.

Many users indicated that the ability to navigate slides was important for note-taking and for answering questions that involved terms or numbers located on the slides. Only two participants in any of the navigable conditions expressed any preference for the LVS; one stated that "the option to skip parts of the lecture made me want to [do so]; thus I could've missed info when I skipped". The other said that there were problems with having to navigate large sections of video just to get one piece of information to answer a question, and that it was potentially more time-consuming to pinpoint the information than to simply just watch a larger interval of video for the information. The implication of these two comments seems to indicate that depending on the location of relevant information in a given webcast, it may be easier to simply watch the webcast straight through to find all of the answers.

Within the navigable conditions, however, users tended to prefer to use the slide toolbox to navigate, and one user said that the video timeline and its markers were "not clearly related [to the] slides in the toolbox". We suspect that it was easier to navigate the material by slide title, and that topical / semantic cues were much preferred for finding answers than blindly searching chronologically. Three participants stated that the resynchronization mechanism was particularly useful for re-establishing video context quickly and easily, with one participant stating that "the button to line up slides with the lecture was helpful". The multiple tracks themselves were also found to be quite useful. As mentioned earlier the ability to search through related slides while watching the video was appreciated, and "finding useful diagrams quickly" and "scan[ning] for answers to related questions" were cited as very important features of the alternate track tabs. It would appear that having multiple tracks of information not only gives the capacity for understanding information in a different context, but also to organize this information for easier search and discovery.

5.8.6 Subjective Ratings

The findings from the Likert scale ratings given by participants indicated that, for the most part, video navigation and slide navigation were deemed easier in the Multitrack system than in the Simple system. We suspect that users felt this way because of the greater number of navigation options that were available in the Multitrack condition; while there were inherently no fundamental differences in the slide-to-slide navigation model, or the video navigation controls, the greater number of options available while browsing the slides in desynchronized mode (whether by using the filter tracks or the standard timeline) perhaps put participants more at ease.

A similar trend is reflected in the ratings of the quality of information in the video and slides. Participants in the Multitrack condition were able to locate the slides with the correct information most easily. The participants in the Multitrack also rated the information on the slides more highly than did participants in the Simple condition, perhaps those using the MVS had both filters and chronological ordering. This probably had a cascading effect. Participants using the MVS believed the quality of information in the video was higher as the number of options for finding information in the slides gave them more time to peruse the video and locate information in it.

In contrast, those participants in the Linear system relied mostly on the slides for information and ranked the video rather lowly, since information in the video passed by them quickly and was not re-viewable. Thus, we speculate that being forced to watch the entire video does not necessarily mean all the material is absorbed since the LVS does not allow for information to be digested and processed by the viewer in a timely manner. The relatively high rankings for the usefulness of slide information in the Linear system was probably due to the fact that the more temporally persistent slides were the only consistent source of information for answering questions.

5.9 Hypothesis Analysis

From our data and analysis we can now assess our hypotheses.

5.9.1 Hypothesis 1a

Participants in the Linear condition will achieve a better overall quiz score at the end of the second session than those in the Multitrack condition.

This hypothesis turned out to be untrue. Participants in the Linear condition did not have a better score at the end of the second session than those in the Multitrack condition. In fact, participants in the Multitrack condition performed significantly better on the new set of questions in the second session than those in the Linear condition. This happened in spite of the fact that users in the Linear system had the chance, in principle, to watch and re-absorb everything, whereas those in the navigable conditions may have skipped important segments in their navigation of the video. There were several instances of this kind of behaviour as discussed earlier.

We believe that there were many mitigating factors that worked against people re-absorbing information in the Linear condition, preventing them from achieving better scores than those in the Multitrack condition. Possible factors include boredom / lack of motivation, and the inability to scroll back to review information in the second session to confirm answers. As mentioned above, users often discussed these issues in relation to their grievances with the non-navigable Linear system.

5.9.2 Hypothesis 1*b*

Participants in the Multitrack condition will achieve a better overall quiz score at the end of the second session than those in the Simple condition.

While participants using the Multitrack system achieved better overall scores than Simple system participants on both the old and new questions in the second session, they did not do so by a statistically significant amount and so our hypothesis turned out to be untrue. However, it is worth noting that participants in the Multitrack condition performed significantly better on the new questions in the second session than users in the Linear condition, whereas participants in the Simple condition were not able to significantly outperform Linear system participants. This could possibly indicate that switching from the LVS to the MVS may be of greater benefit to users than switching to the SVS, since only the MVS can outdo the Linear system in terms of overall quiz performance and information comprehension.

5.9.3 Hypothesis 2*a*

Participants in the Simple condition will rate their system higher in terms of functionality, usability and overall satisfaction than those in the Linear condition.

This hypothesis was untrue. There were no significant differences in Likert scale ratings of system functionality, usability or satisfaction across any of the conditions. However, despite the lack of statistically significant differences, those in the Linear system responded most negatively to the quality of information in the video, and as mentioned earlier, voiced much dissatisfaction with the lack of temporal controls.

Participants felt more efficient using the navigable systems, citing the ability to jump to specific locations or to backtrack over elements they had just missed, that might have been crucial to answering a question. As discussed earlier, this is supported by examples in which active navigation and searching for answers led to better performance on certain questions than the passive way in which users absorbed information in the non-navigable ones.

5.9.4 Hypothesis 2*b*

Participants in the Multitrack condition will rate their system higher in terms of functionality, usability and overall satisfaction than those in the Simple condition.

There was no statistically conclusive evidence supporting this hypothesis, even though participant opinions concerning the ease of use of various parts of each system tended to be in favour of the Multitrack system. A consequence of this trend could possibly have been reflected in the fact that many users in the Multi-Track system were of the opinion the quality of information in the videos and slides was relatively high, compared to the opinions of Standard system users.

5.9.5 Hypothesis 3*a*

Participants in the Simple condition will achieve a higher ratio of quiz score to time spent watching the webcast, than those in the Linear condition.

This hypothesis turned out to be untrue. While participants using the SVS were able to achieve overall scores that were at least as good as the scores of participants using the LVS in a shorter amount of time (SVS participants only spent 32 minutes doing the quiz on average, as opposed to the fixed 40 + 2 minutes in the LVS condition), this level of efficiency was not enough to make a significant difference in overall score-to-time ratios between the two conditions.

However, it is possible that given a full 40 minutes, users of the Simple system could significantly outperform Linear system users in terms of overall quiz score. Users of the SVS were already able to save some browsing time to achieve the same overall score as LVS users; given that participants in the navigable conditions were instructed to quit after achieving what they felt to be score of 80%, it is indeed possible that those users could have potentially done even better if no such restriction were put in place.

5.9.6 Hypothesis 3*b*

Participants in the Multitrack condition will achieve a higher ratio of quiz score to time spent watching the webcast, than those in the Simple condition.

This hypothesis was also untrue. Despite MVS participants achieving better scores in less time than SVS participants, there were no significant differences in ratio between the two conditions. However, the Multitrack condition was significantly better than the Linear condition in terms of ratio, while no such difference was detected between the Simple and Linear systems. If we take the quiz scores to be a valid metric of information absorption and comprehension, we have at least demonstrated that users of the MVS can absorb and comprehend lecture materials at a significantly better rate, than if they simply watched the webcast straight through.

5.9.7 Closing thoughts

The results of our experiment contradicted our base assumption that being made to watch an entire webcast archive would result in a more complete understanding and absorption of the information in it. Those in the Multi-Track condition did significantly better than those in the Linear system in terms of overall score. This result shows that archives, when properly processed in post-production, may be a more useful tool for delivering information than a live webcast. The Linear system emulates the non-navigable nature of a live webcast, and users were not satisfied with the Linear system in terms of usability.

These results suggest that archives can play a specific role in information absorption that live webcasts cannot. The combination of navigability and extra features in augmented archives appears to make them more useful than a live webcast, at least for the process of information absorption and comprehension. This demonstrates that there is definitely potential in researching the enhancement of archives. Exploiting the extra features archives can provide over webcasts may further show how archives are a useful tool in their own right, and that they can be definitely looked to as a way of overcoming the inherent technical shortcomings introduced by webcasts.

Furthermore, because those in the navigable conditions achieved scores comparable to those in the Linear condition without having to watch the entire lecture, we believe that archives can speed up the process of specific information absoprtion and retrieval simply by allowing users to choose what they wish to view, which is an important part of the philosophy behind our archive model.

While we were not able to show that systems augmented with multiple tracks allowed users to absorb significantly more information than a plain archive viewing system, we believe that the satisfaction and high benefit-to-cost ratio of including amenities such as multiple slide tracks, packed densely with related information, are easy for users to grasp and provide alternate sources of information that can prove to be convenient for users. While these measures do not save copious amounts of time, they make the viewing experience more pleasant and the interface easier to navigate, which is an important consideration to many users.

We believe that the approach we have taken in presenting an augmented archive model in a software player shows some promise. By inserting the alternate slide tracks into the archive, we made information available to users in different ways. We feel this kind of variety is important because it does not interfere with the basic understanding of how a webcast archive works, while making parts of the lecture accessible in a much quicker and easier manner. There was no detriment in providing these extra tracks to users, and no one expressed any negative thoughts about the extra interface controls and tabs needed to manage them. If these concepts of track navigation can be fleshed out and expanded upon in future work, then we may be able to more accurately gauge the benefits of these augmentations. We will discuss some possible research directions in the next chapter.

Chapter 6

Conclusion

6.1 Summary of research

We investigated how webcast archives can be re-purposed for different audiences and contexts through the re-organization and re-use of existing archived material. Our research analyzed how existing webcasting systems archived and processed webcasts in post-production to see what kinds of functionality they provided. We constructed a general theoretical model of webcast archiving that could describe existing archive functionalities, as well as describe our ideas for lightweight archive re-purposing in a formal manner. A software prototype for viewing augmented webcast archives was developed. This prototype was based on a simplified version of our model and focused on grouping and organizing presentation slides according to their content on different slide tracks.

A user study was carried out using this prototype to investigate two issues. First, we wanted to see if viewers of webcast archives could absorb and comprehend more information from an archived lecture than those who had viewed non-navigable webcasts. Second, we wanted to determine if the archive augmentations we developed would make it significantly easier to absorb and comprehend information from an archived lecture. We did this by evaluating three different systems for viewing webcast archives: the non-navigable Linear system which was designed to view an archive as if it were a live webcast, the Simple system which featured basic archive navigation functionality present in most modern webcasting systems, and the Multitrack system which was the augmented archive viewing prototype we had developed. We tested users' ability to absorb and comprehend information from an archived lecture viewed using these three systems, by quizzing them on the lecture material.

The results of our study showed that users who used the Multitrack system had better overall quiz scores than those in the Linear condition, suggesting that it was easier to absorb information by watching and navigating augmented lecture archives, than from watching a webcast of the same lecture in its entirety without navigation capabilities. While there were no further statistically significant results, users of the navigable archive systems (the Simple and Multitrack systems) expressed satisfaction with being able to navigate the archives and most opinions on the Linear system were generally quite negative.

Users who watched the lecture using the navigable systems performed at least as well on the quiz as users who watched the lecture using the Linear system, but were able to do so without watching the 40-minute lecture in its entirety. This suggested to us that it was important that users be given the ability to filter out information from an archive as needed, and make it easier for them to identify relevant information.

6.2 Limitations and issues

We would have liked to incorporate more concepts from our model into our final prototype. Some of these have been discussed earlier in Chapter 4. We especially would have liked to incorporate slide chains into our prototype, but the apparent complexity of the chain system was shown to be rather difficult for users to grasp in our informal evaluations. Another feature we would have liked to include was the ability to view multiple slide tracks at once for comparison purposes.

Our study only tested the comprehension skills of users after having watched a webcast archive once, but did not address the question of whether users could *learn* better using more powerful archive viewing systems. Our study did not ask users to answer quiz questions or reproduce content from the lecture without seeing it again, and we did not take advantage of our delayed lecture viewing setup to see if the different systems had any effect on learning. We looked at people's answers on questions that they had not seen in the second session of the experiment but had seen in the first, to see if there was some useful data we could analyze that might hint at learning effects. However, this was not sufficient data to draw any meaningful conclusions from.

In general our results were not significantly strong and did not appear to strongly support or reject our hypotheses. Some of this might be attributed to the complexity of the lecture material, which may have introduced some issues of lecture comprehension across all conditions. If there was indeed a lower overall level of comprehension due to this issue, this may have introduced a potential limit on user performance in general, which would have evened out quiz performance across conditions. This might have made the difference between good scores and bad scores more reliant on chance rather than actual better performance.

6.3 Future work

Further research in this field should build on other aspects of archive augmentation that we did not touch on in our study. For example, it would be useful to analyze the effects of video re-organization in augmented archives, and see if users would respond positively enough to such an augmentation to justify its use. We would also like to implement our idea of chains in webcast archives in another prototype, and further explore the general notion of connecting contextual material via semantic links. Much work has been done in the domain of hypermedia in this area and it would be useful to see if those ideas can help us design a better chain system that might be intuitive and easy to for users to understand.

We would also like to take other directions in future research. In addition to understanding how enhanced archives are navigated and viewed by users, we would also like to research effective ways of creating and editing these archives from the point of view of the archive producer. Since our model is intended to provide a way of doing augmentations in a simple and lightweight manner, we would like to confirm that an appropriately designed archive editing application would an intuitive and effective tool for archive producers to use.

One final possibility worth looking into is the idea of giving webcast archive users themselves the capability to manipulate and create their own derived works from archives. A system supporting such a concept could range from simply supporting user annotations and tags, to giving users the ability to re-cut and re-edit archive video with textual or audio commentary. This would involve research into the rationale and uses of user-created content, and an understanding of the kinds of operations and functions that would be useful to a user creating an edited version of an original archive, be it for personal or public consumption. Taking webcast archive research in this direction would shed new light on the roles of content producers and consumers in the current media-rich Internet landscape.

Bibliography

[1] Breeze resource center. http://www.adobe.com/resources/breeze/. Adobe Systems Inc. Accessed July 20, 2008.

[2] E-learning Solutions - Web conferencing & web based training. http://www.elluminate.com/about.jsp. Elluminate Inc. Accessed July 20, 2008.

[3] MIT OpenCourseWare - introduction to solid state chemistry, fall 2004. http://ocw.mit.edu/OcwWeb/Materials-Science-and-Engineering/3-091Fall-2004/LectureNotes/index.htm. Massachussetts Institute of Technology. Accessed July 20, 2008.

[4] MIT OpenCourseWare Online Course Materials: Audio/Video courses. http://tinyurl.com/5pbbea. Massachussetts Institute of Technology. Accessed July 20, 2008.

[5] Nicovideo. http://nicovideo.jp. Nico Nico Douga. Accessed July 20, 2008.

[6] Youtube company history. http://ca.youtube.com/t/about. Youtube - Broadcast Yourself. Accessed July 20, 2008.

[7] Gregory D. Abowd. Classroom 2000: An experiment with the instrumentation of a living educational environment. *IBM Systems Journal*, 38(4):508–530, 1999.

[8] Gregory D. Abowd, Matthias Gauger, and Andreas Lachenmann. The family video archive: an annotation and browsing environment for home movies. In *ACM SIGMM International Multimedia Conference*, pages 1–8, Berkeley, 2003. ACM.

[9] Ron Baecker. A principled design for scalable internet visual communications with rich media, interactivity, and structured archives. In *Conference of the Centre for Advanced Studies on Collaborative research*, pages 16–29, Toronto, 2003. IBM Press.

[10] Ron Baecker, Melanie Baran, Jeremy Birnholtz, Clarence Chan, Joe Laszlo, Kelly Rankin, Russell Schick, and Peter Wolf. Enhancing interactivity in webcasts using voip. In *Extended abstracts of the SIGCHI conference on Human Factors in computing systems*, pages 235–238, Montreal, 2006. ACM.

[11] Kent Beck and Ward Cunningham. A laboratory for teaching object oriented thinking. In *ACM Conference on Object Oriented Programming Systems Languages and Applications*, pages 1–6, New Orleans, 1989. ACM.

[12] Benjamin Bederson and Lance Good. Zoomable user interfaces as a medium for slide show presentations. *Information Visualization*, 1(1):35, 2002.

[13] Vannevar Bush. As we may think. *Atlantic Monthly*, July 1945. Also available at http://www.theatlantic.com/doc/194507/bush.

[14] Juan Casares, A. Chris Long, Brad Myers, Scott Stevens, and Albert Corbett. Simplifying video editing with silver. In *ACM SIGCHI conference on Human factors in computing systems*, pages 672–673, Minneapolis, 2002. ACM.

[15] Miguel Costa, Nuno Correia, and Nuno Guimares. Annotations as multiple perspectives of video content. In *ACM SIGMM International Multimedia Conference*, pages 283–286, Juan-Les-Pins, 2002. ACM.

[16] Marc Davis. *Media Streams: an iconic visual language for video representation*, pages 854–856. Morgan Kaufmann Publishers Inc., 1995.

[17] David Fono and Ron Baecker. Structuring and supporting persistent chat conversations. In *ACM Conference on Computer-Supported Cooperative Work (CSCW)*, pages 455–458, Banff, 2006. ACM.

[18] Chris Gould. Star wars: The changes. http://tinyurl.com/fp4va, October 2004. DVDActive.com. Accessed July 20, 2008.

[19] Kathy Haramundanis and Larry Rowland. Experience paper: a content reuse documentation design experience. In *ACM International conference on Design of communication*, pages 229–233, El Paso, 2007. ACM.

[20] Beverly Harrison and Ron Baecker. Designing video annotation and analysis systems. In *Graphics Interface*, pages 157–166, Vancouver, 1992. Morgan Kaufmann Publishers Inc.

[21] Manfred Hauswirth and Mehdi Jazayrei. A component and communication model for push systems. In *ACM SIGSOFT international symposium on Foundations of software engineering*, pages 20–38, Toulouse, 1999. ACM.

[22] Liwei He, Jonathan Grudin, and Anoop Gupta. Designing presentations for on-demand viewing. In *ACM Conference on Computer-Supported Cooperative Work (CSCW)*, pages 127–134, Philadelphia, 2000. ACM.

[23] Ellen Isaacs, Trevor Morris, and Thomas K. Rodriguez. A forum for supporting interactive presentations to distributed audiences. In *ACM Conference on Computer-Supported Cooperative Work (CSCW)*, pages 405–416, Chapel Hill, 1994. ACM.

[24] Gavin Jancke, Jonathan Grudin, and Anoop Gupta. Presenting to local and remote audiences: design and use of the telep system. In *ACM SIGCHI conference on Human factors in computing systems*, pages 384–391, The Hague, 2000. ACM.

[25] Sue Kennedy. Using video in the bnr usability lab. *ACM SIGCHI Bulletin*, 21(2):92–95, 1989.

[26] Duy-Dinh Le and Shin'ichi Satoh. National institute of informatics, japan at trecvid 2007: Bbc rushes summarization. In *ACM SIGMM International Multimedia Conference*, pages 70–73, Augsburg, 2007. ACM.

[27] Marcial Losada and Shaun Markovitch. Groupanalyzer: A system for dynamic analysis of group interaction. In *IEEE International Conference on System Sciences*, pages 101–110, Kailua-Kona, 1990. IEEE Publishing.

[28] Wendy E. Mackay and Glorianna Davenport. Virtual video editing in interactive multimedia applications. *Communications of the ACM*, 32(7):802–810, 1989.

[29] Cosmin Munteanu, Ron Baecker, Gerald Penn, Elaine Toms, and David James. The effect of speech recognition accuracy rates on the usefulness and usability of webcast archives. In *ACM SIGCHI conference on Human factors in computing systems*, pages 493–502, Montreal, 2006. ACM.

[30] Theodor Nelson. Complex information processing: a file structure for the complex, the changing and the indeterminate. In *ACM Annual Conference*, pages 84–100, Cleveland, 1965. ACM.

[31] Peter Plessers, Sven Casteleyn, Yeliz Yesilada, Olga De Troyer, Robert Stevens, Simon Harper, and Carole Goble. Accessibility: a web engineering approach. In *ACM International World Wide Web Conference*, pages 353–362, Chiba, 2005. ACM.

[32] Michael J. Potel and Richard E. Sayre. Interacting with the GALATEA film analysis system. In *SIGGRAPH Computer Graphics Conference*, pages 52–59, Philadelphia, 1976. ACM.

[33] Lawrence A. Rowe, Diane Harley, and Peter Pletcher. *BIBS: A Lecture Webcasting System.* Berkeley Multimedia Research Center, Berkeley, June 2001.

[34] Gitta B. Salomon. Designing casual-user hypertext: the CHI'89 InfoBooth. In *ACM SIGCHI conference on Human factors in computing systems*, pages 451–458, Seattle, 1990. ACM.

[35] Dian Tjondronegoro, Phoebe Chen, and Binh Pham. Sports video summarization using highlights and play-breaks. In *ACM SIGMM International Multimedia Conference*, pages 201–208, Berkeley, 2003. ACM.

[36] Elaine Toms, Christine Dufour, Jonathan Lewis, and Ron Baecker. Assessing tools for use with webcasts. In *International Conference on Digital Libraries*, pages 79–88, Denver, 2005. ACM.

[37] Tim Tully. The sounds of evil. *Videography Magazine*, January 1999. Also available at http://filmsound.org/murch/evil/.

[38] Jihua Wang and Tat-Seng Chua. A framework for video scene boundary detection. In *ACM SIGMM International Multimedia Conference*, pages 243–246, Juan-Les-Pins, 2002. ACM.

[39] Stephen White, Anoop Gupta, Jonathan Grudin, Harry Chesley, Greg Kimberly, and Elizabeth Sanocki. Evolving use of a system for education at a distance. In *International Conference on System Sciences*, pages 3047–3057, Hawaii, 2000. IEEE Publishing.

[40] P.T. Zellweger. Scripted documents: a hypermedia path mechanism. In *ACM conference on Hypertext and Hypermedia*, pages 1–14, Pittsburgh, 1989. ACM.

[41] Hong-Jiang Zhang, Chien-Yong Low, Stephen Smoliar, and Jian-Hua Wu. Video parsing, retrieval and browsing: an integrated and content-based solution. In *ACM SIGMM International Multimedia Conference*, pages 15–24, San Francisco, 1995. ACM.

Appendix A

Study Ethics Certificate

The certificate of approval issued by the Behavioural Research Ethics Board at the University of British Columbia, which granted permission to perform the minimal risk study described in Chapter 4 of this thesis.

The University of British Columbia
Office of Research Services
Behavioural Research Ethics Board
Suite 102, 6190 Agronomy Road, Vancouver, B.C. V6T 1Z3

CERTIFICATE OF APPROVAL - MINIMAL RISK AMENDMENT

PRINCIPAL INVESTIGATOR:	DEPARTMENT:	UBC BREB NUMBER:
Kellogg S. Booth	UBC/Science/Computer Science	H07-02306

INSTITUTION(S) WHERE RESEARCH WILL BE CARRIED OUT:

Institution	Site
UBC	Vancouver (excludes UBC Hospital)

Other locations where the research will be conducted:

N/A

CO-INVESTIGATOR(S):

Clarence Chan

SPONSORING AGENCIES:

Natural Sciences and Engineering Research Council of Canada (NSERC)

PROJECT TITLE:

Network for Effective Collaboration Technologies through Advanced Research

Expiry Date - Approval of an amendment does not change the expiry date on the current UBC BREB approval of this study. An application for renewal is required on or before: October 25, 2008

AMENDMENT(S):	AMENDMENT APPROVAL DATE:
	December 6, 2007

Document Name	Version	Date

Consent Forms:

Figure A.1: Certificate of approval from the Behavioural Research Ethics Board of UBC to perform the minimal risk study described in Chapter 4.

Consent form (control group)	v3	November 30, 2007
Consent form (experimental group)	v3	November 30, 2007

The amendment(s) and the document(s) listed above have been reviewed and the procedures were found to be acceptable on ethical grounds for research involving human subjects.

Approval is issued on behalf of the Behavioural Research Ethics Board and signed electronically by one of the following:

Dr. M. Judith Lynam, Chair
Dr. Jim Rupert, Associate Chair
Dr. Laurie Ford, Associate Chair

Figure A.2: Certificate of approval from the Behavioural Research Ethics Board of UBC to perform the minimal risk study described in Chapter 4.

Appendix B

List Of Quiz Questions

Below is the complete list of quiz questions asked in the first and second webcast sessions of our research study.

1. **Name the group of elements that have an outer valence of 0 on the periodic table. (1 mark)**

   noble gases (1)

2. **What was Dalton's preferred method of ordering elements? (2 marks)**

   ascending order (1) of atomic mass (1)

3. **What year was oxygen entered into the periodic table? (1 mark)**

   1774 (1)

4. **What symbols were originally used as the abbreviations for potash and silver? (2 marks)**

    (1)

    (1)

5. **What year did Mendeleev publish his findings in the Russian Chemistry Journal Society? (1 mark)**

   1869 (1)

6. **In what critical way did Mendeleev's work on element classification go beyond that of Julius Meyer's? (2 marks)**

   - left blank spaces in the periodic table (1)
   - made testable predictions based on table entries (1)

Figure B.1: List of quiz questions used in the study. Participants were asked to answer these questions while watching the webcasted lecture.

7. **Complete the following pairings by filling the appropriate selection (a-d) in the adjacent section: (2 marks)**

eka-silicon	______c______	(0.5)
eka-aluminum	______d______	(0.5)
eka-boron	______a______	(0.5)
eka-zirconium	______b______	(0.5)

a) gallium
b) hafnium
c) germanium
d) scandium

8. **What is the unofficial name of the element with atomic number 110? (1 mark)**

darmstadium (1)

9. **Why is there a discrepancy between a) Dalton's assertion that atomic numbers must come in whole-number values and b) the fact that atomic masses of elements are often listed in non-whole number values? (2 marks)**

- the atomic mass is determined by calculating the frequency (1) of isotopes relative each other; isotopes always have whole numbers of neutrons (1)

10. **Label A, X, and Z in this diagram. (3 marks)**

$${}^{A}_{Z}\mathbf{X}$$

A = atomic mass = protons + neutrons (1)

X = atomic symbol (1)

Z = proton number = # of electrons in a neutral atom (1)

Figure B.2: List of quiz questions used in the study. Participants were asked to answer these questions while watching the webcasted lecture.

11. What is the mass in grams of 200 cm3 worth of germanium? (2 marks)

1072 g (2)

12. How far off is the mass for the same volume of germanium according to Mendeleev's calculations? (2 marks)

28 g (2)

13. How many neutrons would a single atom of Carbon-17 contain? (1 mark)

11 neutrons (1)

14. Arrange these elements by first ionization potential, in ASCENDING order. (2 marks)

Neon Titanium Bromine

1. Titanium
2. Bromine
3. Neon

Hint: their atomic numbers are 10, 22 and 35 respectively.

Figure B.3: List of quiz questions used in the study. Participants were asked to answer these questions while watching the webcasted lecture.

15. How are elements grouped according to the triad method (i.e. what relationship exists between these elements?) (2 marks)

- Grouped in threes (1), the middle element = avg atomic number of other two (1)

16. What property of Millikan's velocity function, when graphed out, demonstrated that charge was in fact quantized? (1 mark)

- The velocity function was shown to be discontinuous (1)

17. A lump of 107 g of silver contains the same amount of _____ as 12 g of carbon. (1 mark)

- electric charge (1)

18. What is the atomic radius of element number 9, fluorine? Use the correct unit of measurement. (1 mark)

- 0.5 angstroms (1)

19. Why was hydrogen not entered in the periodic table until the 1700's, despite people knowing of its existence before? (1 mark)

- It was not chemically isolated or characterized until then (1)

20. Which classification method did Newlands present to the Royal Society? (1 mark)

- Octave method (1)

21. Of the elements with atomic #'s 23, 52, 75, which one has the highest boiling point? The lowest? (2 marks)

- 75 high (1), 52 low (1)

Figure B.4: List of quiz questions used in the study. Participants were asked to answer these questions while watching the webcasted lecture.

Appendix C

List Of Post-Study Questions

Below is the complete list of questions users were asked at the end of the study, to get their opinions on different aspects of the webcasting systems.

- How would you rate your overall performauce on the quiz for the *first* sessiou?

0-20% correct	21-40% correct	41-60% correct	61-80% correct	81-100% correct

- How would you rate your overall performance on the *old questions* for the *second session*?

0-20% correct	21-40% correct	41-60% correct	61-80% correct	81-100% correct

- How would you rate your overall performance on the *new questions* for the *second session*?

0-20% correct	21-40% correct	41-60% correct	61-80% correct	81-100% correct

- Can you comment on any *difficulties* the interface presented to you as you attempted to answer the questious?

- Can you comment on any *benefits* the software presented to you as you attempted to answer the questions?

- In which viewing session was it more difficult to do well on the quiz? Can you explain why?

Figure C.1: List of all questions that participants were asked to answer at the conclusion of the second session of the study.

- Rank the overall ease of use of the software

1	2	3	4	5
very hard to use		somewhat easy to use		very easy to use

- Rank the usefulness of the information on the *slides* in answering the questions

1	2	3	4	5
not useful at all				very useful

- Rank the usefulness of the information in the *video* in answering the questions

1	2	3	4	5
not useful at all		somewhat useful		very useful

- Rank the difficulty of the original questions from the *first* session

1	2	3	4	5
too easy given the material presented				too hard given the material presented

- Rank the difficulty of the *new* questions from the *second* session

1	2	3	4	5
too easy given the material presented				too hard given the material presented

- How true is this statement: "The process of navigating through *video* was intuitive and easy to use"

1	2	3	4	5
not true		somewhat true		very true

Figure C.2: List of all questions that participants were asked to answer at the conclusion of the second session of the study.

- How true is this statement: "The process of navigating through *slides* was intuitive and easy to use"

1	2	3	4	5
not true		somewhat true		very true

- How true is this statement: "The *multiple slide tracks* feature was intuitive and easy to use"

1	2	3	4	5
not true		somewhat true		very true

- How true is this statement: "The *multiple slide tracks* feature was helpful for finding information to answer questions"

1	2	3	4	5
not true		somewhat true		very true

- How true is this statement: "Resyncing video to slides was intuitive and easy to use"

1	2	3	4	5
not true		somewhat true		very true

- Rank the following features with regards to how helpful they are in completing the quiz with as HIGH A SCORE as possible (write each of the numbers 1-4 exactly once beside each feature, 1 for most helpful, 4 for least helpful):

Video navigation timeline
Slide navigation arrow buttons
Skipping to a particular section in the video using the slide toolbox on the left
Skipping to a particular slide in the video using the slide toolbox on the left
Availability of multiple slide tracks / filters

- Rank the following features with regards to how helpful they are in completing the quiz as QUICKLY as possible (write each of the numbers 1-4 exactly once beside each feature, 1 for most helpful, 4 for least helpful):

Video navigation timeline
Slide navigation arrow buttons
Skipping to a particular section in the video using the slide toolbox on the left
Skipping to a particular slide in the video using the slide toolbox on the left
Availability of multiple slide tracks / filters

Figure C.3: List of all questions that participants were asked to answer at the conclusion of the second session of the study.

- Do you have any other comments on the software, quiz, or experiment?

Figure C.4: List of all questions that participants were asked to answer at the conclusion of the second session of the study.