Multimedia and Human-in-the-Loop: Interaction as content enrichment

Bruno Emond
Institute for Information Technology
National Research Council Canada
1200 Montreal Road, Ottawa, ON. Canada. K1A 0R6
613.991.5471
bruno.emond@nrc-cnrc.gc.ca

ABSTRACT
The current work is part of the broadband visual research program at the Institute for Information Technology (National Research Council Canada). The research program is currently focused on developing human-centered multimedia technology to support large group visual communication and collaboration. This paper outlines some conceptual foundations for the development of a human-centered multimedia research tool to capture interaction data, which could be linked to users cognitive processing. The approach is based on the notion of multimedia interaction as content enrichment and on cognitive modeling methodology.

Categories and Subject Descriptors
H.1.2 [**User/Machine Systems**]: Human information processing.
H.5.3 [**Group and Organization Interfaces**]: Asynchronous interaction, Collaborative computing, Synchronous interaction.

General Terms
Measurement, Human Factors, Theory.

Keywords
Cognitive Modeling; Unified theories of cognition; User, context, and task modeling in multimedia systems; Human interaction modeling from multimedia.

1. INTRODUCTION
At no other time in human history we have had the computing and communication capabilities currently in place to overcome time and space limits to human communication. In addition, the digital nature of all information produced now enables blending of separate media streams into new representations and media augmentations. There are many technological innovations that are still needed to allow seamless interaction with multimedia. However, these innovations have the best chance of finding broad applications and user acceptation if they are developed with cultural awareness, enable integration of sensors and multiple media, and offer access outside the desktop by a wide range of users [1]. Human-centered multimedia development requires a synergy between technological innovations and understanding of human performance and learning capabilities. Throughout the document, video and multimedia will be used interchangeably, video being in itself some form of multimedia presentation [2].

HCM'07, September 28, 2007, Augsburg, Bavaria, Germany.
ACM 978-1-59593-781-0/07/0009.

This paper has three main sections. The first section argues, from an example that video production and analysis fall on a continuum of sense-making activities with the purpose of enhancing human communication through content enrichment. The second section presents a case for a research agenda on video enrichment that requires detailed models of how machine or human annotations determine digital video interaction patterns. Finally, the last section positions human-centered multimedia in the framework of unified theories of cognition [3]. This framework offers an interesting theoretical and methodological foundation for understanding how human process multimedia information. It focuses on one of the development area of human-centered multimedia, which is to understand how humans interact with multimedia from a perceptual and cognitive perspective, and gain a better understanding of how humans perceive and interpret multiple modalities [1]. The paper concludes by situating the current proposal in the larger context of the broadband visual research program at the Institute for Information Technology (National Research Council Canada).

2. MEDIA PRODUCTION AND ANALYSIS: A CONTINUUM OF SENSE-MAKING ACTIVITIES
The main human-centered activities in multimedia are identified as media production, annotation, organization, archival, retrieval, sharing, analysis, and communication [4]. One could look at these as separated but related activities. However, when the lifecycle of a video document is the focus of attention, the basic distinction between media production and media analysis tends to blur, as the media analysis can produce additional media to be blended with the original media, such as index information included as text captions, resulting into a new media composition subjected to further interpretation and analysis. This interpretation process is similar to the semiosis process defined by Peirce, where the interpretant of a sign is another sign for the same object [5, 6]. Looking at the continuum of multimedia production and analysis from this perspective, one is draw to the position that the multimedia lifecycle is a continuum of sense making activities mediated by multimedia interpretation generated by humans and by machines. Human enrichments include content expert analysis and user annotations, and machine enrichments include automatic generation of metadata [7], summaries [8], multimodal annotations [9, 10], and visual

abstractions [11]. These machine enhancements are certainly at the core of a human-centered multimedia strategy in a multimedia rich environment [7].

The Figure 1 presents the different moment in the life cycle of video production where annotations can be added to an audio-visual document. This Figure represents a good abstraction of the type of multimedia application that we have pursued at the National Research Council Canada to support human-to-human visual communication in broadband networks. The figure is divided in three sections, representing the main components of a video, from video capture to multimedia distribution going through video production and video assets management. The left part of the figure represents some key steps involved in a video life cycle. The right part of the figure shows the main video annotation processes and components that can be added to the key steps of video production for enhancing audio-visual documents with annotations. The boxes in purple represent processes performed by humans, while the boxes in green represent processes performed by machines.

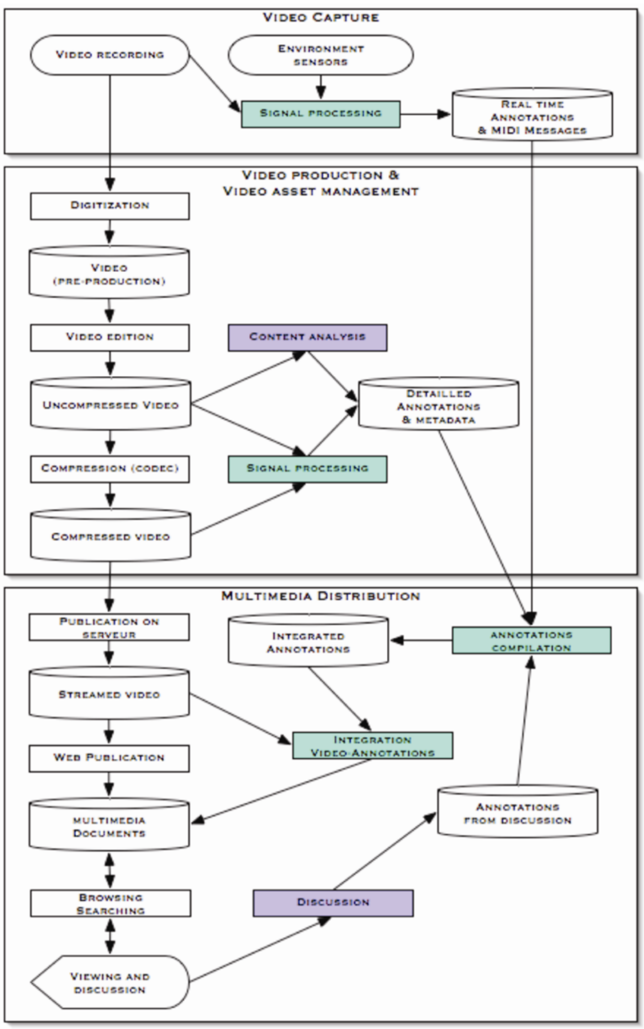

Figure 1. Processes, audio-visual documents, and annotations produced during video capture, video production, media management, and multimedia distribution. Purple indicates human operations while green indicates machine operations.

Adding annotation during video capture is not an easy task for humans. The inherent complexity of learning situations, being a piano lesson or pedagogical situations, often requires the presence of multiple cameras [12, 13], and the rich content of learning situations limit the capability of a human operator to annotate all relevant events that are happening in real time. However, some signal processing algorithms can be applied to extract some relevant features of the situations.

One application for adding annotations in real time is to associate the MIDI messages that are transmitted by a keyboard to a video of the piano playing events that is being recorded. This application could record both the video in MPEG4 [14], as well as MIDI [15, 16] messages in real-time, and automatically associates, without human intervention, every MIDI event to a video frame. Such an application would be useful for annotating or indexing a video from its related MIDI content. The video could then be searched by using MIDI sequences or played with the corresponding MIDI messages sent to a keyboard.

The production of annotations and metadata is an essential element for managing video assets [7]. The intention of metadata schemas such as Dublin Core [17] or IEEE LOM [18] is to provide a common framework for applications to share metadata representations for information reuse and interoperability. Even though metadata schemas are central to asset management applications, they are sometimes too general to capture the specificity of a particular domain like piano pedagogy. At the NRC we have developed an application to capture expert knowledge as part of the video edition process. During the video production phase, an expert can add labels to video segments in order to capture semantic elements that automatic machine techniques cannot yet provide. The results of this analysis are stored in edition lists that are processes in a batch mode to generate multimedia documents encoded in SMIL [19]. The Piano Pedagogy research laboratory at the University of Ottawa in collaboration with the NRC is currently developing a set of annotation tags specific to the piano pedagogy [20]. The application has been also applied to the domain of violin in collaboration with the National Art Center (Canada) with an interface in four languages: English, Chinese, Japanese, and French.

After metadata and different annotation sets have been associated to an audio-visual document, there is still a potential to add annotations at the stage of multimedia distribution. A Web interface to a video server can provide the tools to allow piano students and instructors to discuss content once the material has been published and is being streamed. The Video Annotation Server [21, 22] is such an application and offers to streaming video viewers the possibility of adding free textual messages to a video or a set of videos. This application has been used during two projects involving many school boards across Canada. The first project, LearnCanada (www.learncanada.ca) aimed at supporting teacher professional development, while the focus of the second one, MusicGrid (www.musicgrid.ca), was on supporting music programs using broadband technology [23].

3. INSTRUMENTATION FOR MULTIMEDIA INTERACTION RESEARCH

The previous section outlined the life cycle of multimedia production and analysis as a continuous process of content enrichment for the purpose of human-to-human communication. To fulfill one objective of human-centered multimedia development, technological innovations supporting content enrichment have to be in synergy with a proper understanding of human performance and learning capabilities. There is a strong relation between media analysis and interaction, and the ability to interact with multimedia is highly dependent on how a media is indexed [4], but also, adequate indexing requires an understanding of the task context and the human cognitive system. It appears that a human-centered multimedia research agenda requires some form of instrumentation to capture interaction data for the dual purpose of evaluating the adequacy of technological innovations, and modeling human cognitive [24] and social processes [25, 26]. The following paragraphs present some results of an empirical study aimed at evaluating the usefulness of video browsing controls and video browsing patterns using an enhanced media player (ReView)[27].

In addition to media indexing, navigation in time-based media can also be achieved with VCR like controls. Digital video creates the potential for several new browsing controls that could not be implemented on VCRs (which play analogue media), such as random access to video segments, looping, segment preview by thumbnails [28, 29], slide bar, and other indexing mechanisms supporting video browsing. Nonetheless, VCR-like controls that enable quick and user-friendly browsing of multimedia content are desirable in digital video applications [30].

Previous research has indicated that the use of different media player controls (like Fast Forward) depends strongly on the content of the video [31]. In particular audio-centric videos (classroom lectures, conference presentations) tend to reveal uses of media browser controls supporting content indexing. Video-centric content (sports, travel) makes the use of video frame based navigation more relevant, while narrative/entertainment-centric content (shows) does not promote use of browsing functionality because the viewer's enjoinment is bound to the sequential and normal play speed of the media.

We examined the use of media player controls by highly skilled music students viewing one of their own ensemble coaching lessons. Such lessons have both important audio components (music played and verbal instructions), and video components (musician movements and postures). Even if this type of media content is focused on learning, it is certainly different from audio-centric videos studied by Li [31], raising the possibility that users would prefer different media player controls.

ReView was developed for the Windows operating system using Visual Basic. We equipped the player (see Figure 1) with the usual VCR-type features: play, rewind, and fast-forward, and volume and mute features. With a VCR, the different parts of a tape must be accessed in sequence. In contrast, a digital video format provides the potential to access any part of the video in an instant. Like many other standard media players (QuickTime Player, Windows Media, Real Player), ReView has a playback head that can be dragged to move quickly from one frame to any other (Figure 1, number 2). Like a computer window's scroll bar, the playback head provides a spatial indication of the current video frame's position in the video. A small pop-up tool tip dialog box also showed the video's time index when the head was selected.

ReView was developed in consideration of both the video browsing task, and as an aid to use video during instrument practice. Accordingly, to help musicians use auditory signals to recognize sections of the video more easily when browsing, or identifying musical mistakes [32], we added a skip and play function that preserve pitch and tempo (unlike the classic rewind or fast forward). These added play back functions for both the forward and backward buttons (after [31]) and in the playback head. We also included features to create and save segments of the video. Looping these segments continuously was also made possible. Additional information about ReView's features is provided in the Figure 1 caption.

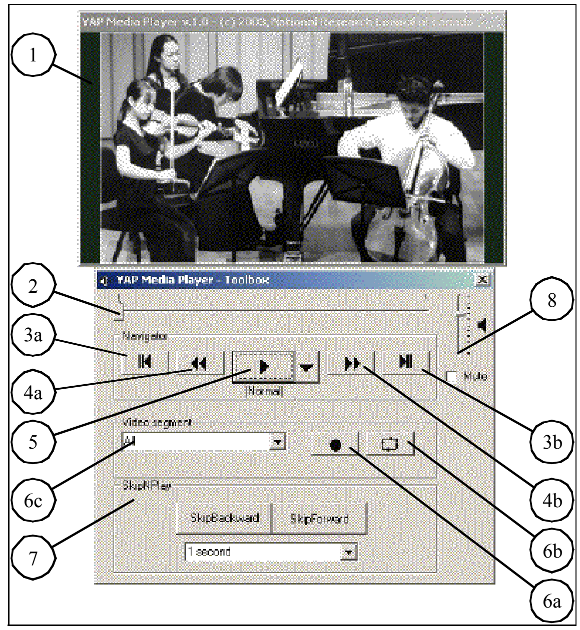

Figure 1. The enhanced media player "ReView". The viewer window displays the video image. 2. The playback head can be clicked and dragged to navigate through the video. 3, 4, and 5. The navigation control panel. 3a. The skip-to-beginning button instantly sets the playback head to the beginning of the video. 3b. The skip-to-end button instantly sets the playback head to the end of the video. 4a. Rewind plays backward through the video, at 4x the normal speed, playing distorted sound. 4b. Fast Forward plays forward through the video, at 4x the normal speed, playing distorted sound. 5. SpeedPlay control drop down menu (options: slowest, slow, normal, fast, and fastest). 6. Video segment control panel. The ClipMark features allow users to save video clips and replay them later. When a video clip is played, the navigation features work in relation to that clip only. 6a. Clicking on the record button starts or stops the recording of a video clip. As the clip plays, it is saved. 6b. The clooping button allows users to loop video clips. 6c. Video segment selection drop down menu, which allows users to choose a named video clip to play. 7. The SkipNPlay features allow the user to quickly skip through the video forward or backward by increments of one, five, and 30 seconds, or, one and five minutes. 8. Volume control and mute.

The content browsing analysis showed that the students spent a disproportionate amount of their time reviewing their performance, as opposed to reviewing video of the coach's instructions. This kind of information is essential in designing an indexing schema over the media, information that would not available if the media player had not been instrumented to collect media browsing behavior.

One goal of this study was to identify which controls should be included in a media player to support video browsing. The preferred control was the slide bar, which supports rapid positioning of the play head as well as variable speed browsing of video content. The data on control use is convergent with this result because the slide bar was the most used media player control. These results suggest that only a simple set of controls is required: pause, play, and the slide bar. However, it is important to note that the students were limited to a single viewing session of 25 minutes. Longer or additional sessions may have revealed different control use patterns.

The empirical study of human-multimedia interaction also revealed a series of possible enhancements such as score annotation, and video zoom-in zoom-out, and remote control by voice.

We intentionally designed ReView as a technological tool to support complex learning in ways that are closely related to the cognitive needs of the learner. As technological tools become more sophisticated, refined, and targeted they can push the boundaries imposed by general applications to more specifically meet the cognitive needs of individual learners in particular domains. There remains great potential to develop such customized technology to more effectively support music learning.

4. INTERACTION FROM THE BIOLOGICAL TO THE SOCIAL

Since the wide spread of Internet applications, such as web browsers, email, and variations of immediate messaging, it has become clear that Internet technologies provided more than an access to a very large repository of information, where behavior can be characterized as information foraging. Internet technologies also support social gathering and relationship building, and define new cultural artifacts shaping cultural and social behavior [33]. Under the information foraging view [34], Internet users are seen as "informavores" maximizing gains of valuable information per unit cost associated to their searching efforts in information ecologies. On the other hand, looking at the Internet as a social phenomenon changes the focus from information processing to social processes such as communication and cooperation in communities of practice [35]. From this perspective, the Internet has created a new psychological space that is a fertile ground for social relationships, roles, and a new sense of self [36].

The task of studying interactive human behavior in information and communication technology environments is a complex one. In particular, it is difficult to create a suitable formal model for cooperation because of the inherent ambiguity and unpredictability of behavior in very opened tasks, and the diversity of actors' point of views [37, 38]. These difficulties are compounded when networked collaborative environments force individuals to deal with user interface problems and time delays affecting remote collaboration. These technical and human factors problems add a layer of complexity to an already overwhelming set of social constructs [37, 38]. The complementary explorative multilevel data analysis [39] and multi-faceted evaluation factors [37] are good examples of the complexity levels embedded in studying human cognition in Internet environments.

To address the complexity problem, computer simulations are increasingly argued to constitute an important method for studying phenomena arising from organized complexity [40]. Simulations have been widely used in contemporary social research [40]. Agent based simulation is ideally suited to exploration of the implications of non-linearity in systems behavior and also lends itself to models that are readily scalable in scope and levels. The approach is useful for examining the relationship between micro-level behavior and macro outcomes. Many social systems may be usefully modeled with simple agents [41, 42]. These do, however, impose limits on the range and scope of social behavior that can be investigated particularly where human social systems are the subject of research [43].

One of the most important aspects of social life and intelligence is the diffusion of mental representations among agents. According to Conte and Paolucci [44], a theory of mental processes is necessary to understand and predict social phenomena that involve mentally complex systems such as humans. According to them, a theory of cultural transmission should account not only for the outcomes of the process (observable behaviors) but also for the process itself, and a model of the process of cultural transmission implies a model of the mental states involved in it.

The explicit specification of representations and processes that operate and transform these representations is certainly a strength of cognitive modeling research. Cognitive models are computer programs implemented with the core resources of a cognitive architecture. Cognitive models simulate human performance on cognitive skills. These models have been successfully applied in many domain areas such as perception and attention [45], learning and memory [46-48], problem solving and decision making [49, 50], language processing [51, 52], syllogistic reasoning [53, 54], game playing [55], fMRI [56], communication, negotiation, and group decision-making [57-59].

One of the central arguments of cognitive modeling is that some of the large time scale phenomenon such as learning and social interaction, which can take months and years, can be viewed as emerging properties of cognitive systems at a much smaller time scale (seconds and milliseconds). Anderson [60] has argued that cognitive architectures can bridge 10 milliseconds effects, which reside in the biological band, and significant learning outcomes (100 hours), which reside in the social band. The other cognitive bands are the cognitive and the rational, and fall between the two previous one [3, 60]. The cognitive bands are reproduced in Table 1. The possibility that a unified theory of cognition can cover all of these levels is a significant contribution to the study of complex cognitive systems.

Table 1: Newell's four bands of cognition.

Scale	Time units	System	World
10^7	Months		
10^6	Weeks		Social
10^5	Days		
10^4	Hours	Task	
10^3	10 Minutes	Task	Rational
10^2	1 Minute	Task	
10^1	10 Seconds	Unit Task	
10^0	1 Second	Operations	Cognitive
10^{-1}	100 Milliseconds	Deliberate act	
10^{-2}	10 Milliseconds	Neural circuit	
10^{-3}	1 Millisecond	Neuron	Biological
10^{-4}	100 Microsecond	Organelle	

The bridging argument [60] is based on three theses supported by empirical data. They are the decomposition, the relevance, and the modeling theses. The decomposition thesis states that learning occurring at the social band can be reduced to learning occurring at lower bands. The relevance thesis states that instructional outcomes at the social band can be improved by paying attention to cognition at lower bands. And finally, the modeling thesis states that cognitive modeling can solve the technical barrier of tracing out the consequences of events occurring at a fine-grain temporal structure. According to Anderson, there is good evidence for all the theses. The decomposition thesis has strong support from the social to the neurological band, but the relevance thesis weakens as we move down to the sub-symbolic level, and the modeling thesis needs to overcome some technical barriers below the unit task level.

The decomposition, the relevance, and the modeling theses need additional research to be supported. However, the framework is promising, and relies on some solid cognitive science foundations, and research on cognitive architecture and cognitive modeling. A framework is a general concept for understanding a domain, but it does not have predictive power, whereas a theory is a precise system that embodies specific framework level concepts and can be used to make predictions [61, 62]. For example, the idea that cognition can be understood using production rules (i.e., if/then rules) is a framework level assumption, while specific implementations of production rules to do so constitutes a theory. Cognitive architectures, therefore, can be considered theories capable of explaining complex cognitive behaviors. Models are the result of applying a theory to a specific task or phenomenon to predict measures of performance such as processing time, errors, learning rates and learning patterns.

Cognitive architectures are general in the sense that they claim to be universally capable of modeling cognitive activity. Cognitive architectures are relatively complete proposals about the structure of human cognition. A cognitive architecture provides the resources for developing models. These resources take the form of a set of specifications regarding the functional invariants [63] related to knowledge representation, knowledge processing, memory, perception, and motor actions. Some examples of cognitive architectures are SOAR [3], ACT-R [62], EPIC [64], and CI [65]. Each of these architectures has its strength and was initially developed with some intended modeling purpose. ACT-R is mainly focused on problem solving and memory, SOAR on problem solving and learning, EPIC on multiple task performance, and CI on text comprehension. Bringing cognitive modeling into the realm of human-centered multimedia is not a trivial task but the methodological and computer simulation tools

5. CONCLUSION

The current work is part of the broadband visual research program at the Institute for Information Technology (National Research Council Canada). The research program is currently focused on developing human-centered multimedia technology to support large group visual communication and collaboration. Applications such as virtual citizen town halls and virtual classrooms have been designed with a combined synchronous and asynchronous visual communication environments to support large group collaboration (500-1000 people simultaneously). The research program fosters and combines interdisciplinary research in the areas of asynchronous and asynchronous communication, Web services, augmented reality, and social analysis.

This paper outlines some conceptual foundations for the development of a human-centered multimedia research tool to capture interaction data, which could be linked to users cognitive processing. The approach is based on the notion of multimedia interaction as content enrichment and on cognitive modeling methodology. Cognitive modeling, as a methodology, has been used successfully in the area of human computer interaction [24, 66-68], and could be expanded to human-centered multimedia. Cognitive modeling methodology also shares a similar view with human-centered computing that the analysis of interactive behavior has three fundamental units of analysis: machines, humans, and contexts. The latter including the team, organization, and work environment [4]. For cognitive modeling interactive behavior arises from the limits, mutual constraints, and interactions between and among each member of the cognition-artifact-task triad [66, 67]. However, cognitive modeling tends to be focus more on the granular level of task performance and learning rather than on the structural properties or teams and organizations.

6. REFERENCES

[1] A. James and N. Dimitrova, “Human-centered multimedia: culture, deployment, and access,” *IEEE Multimedia Magazine*, vol. 13, pp. 12-19, 2006.

[2] A. James, M. Christel, S. Gilles, R. Sarukkai, and W.-Y. Ma, “Multimedia information retrieval: What is it, and why isn't anyone using it?,” in *Proceedings of the 7th ACM SIGMM international workshop on Multimedia information retrieval*. New York, NY, USA: ACM Press, 2005, pp. 3-8.

[3] A. Newell, *Unified theories of Cognition*. Cambridge, MA: Harvard University Press, 1990.

[4] A. Jaimes, N. Sebe, and D. Gatica-Perez, “Human-centered computing: a multimedia perspective,” in *Proceedings of the 14th annual ACM international conference on Multimedia*, 2006, pp. 855-864.

[5] C. S. Peirce, “On a new list of categories,” *Proceedings of the American Academy of Arts and Sciences*, vol. 7, 1868.

[6] B. Emond, “Pour une grammaire formelle de la classification des signes chez Charles S. Peirce,” *Semiotica*, vol. 72, 1988.

[7] J. Hunter and J. Newmarch, “An indexing, browsing, search and retrieval system for audiovisual libraries,” in *Lecture Notes in Computer Science*, vol. 1696, 1999, pp. 76-91.

[8] D. Zhang and J. F. Nunamaker, “A natural language approach to content-based indexing and retreival for interactive e-learning.,” *IEEE Transactions on Multimedia*, vol. 6, pp. 450-458, 2004.

[9] C. G. M. Snoek and M. Worring, “Multimodal video indexing: a review of the state-of-the-art,” *Multimedia Tools and Applications*, vol. 25, pp. 5-35, 2005.

[10] D. O. Gorodnichy and A. Yogeswaran, “Detection and tracking of pianist hands and fingers,” presented at Proc. of the Canadian conference Computer & Robot Vision (CRV'06), Québec, Canada, 2006.

[11] M. M. Yeung and B. L. Yeo, “Video visualization for compact presentation and fast browsing of picturial content,” *IEEE Transactions on Circuits Systems and Video Technology*, vol. 7, pp. 771-785, 1997.

[12] G. Comeau, M. Brooks, and J. Spence, “Video and broadband video conference in professional development,” presented at Biennial conference of the International Consortium for Educational Development (ICED), University of Ottawa, Ottawa, Ontario., 2004.

[13] J. W. Stigler, P. Gonzales, T. Kawanaka, S. Knoll, and A. Serano, “The TIMSS videotape classroom study: methods and findings from an exploratory research project on eighth-grade mathematics instruction in Germany, Japan, and the United States,” National Center for Education Statistics Report, US Department of Education, Washington 1999.

[14] I. I. JTC1/SC29/WG11, “MPEG-4 Overview - (V.21 – Jeju Version),” 2002.

[15] International_MIDI_Association, *MIDI Musical Instrument Digitial Interface Specification 1.0*. Los Angeles, CA: International MIDI Association, 1983.

[16] International_MIDI_Association, *Standard MIDI Files 1.0*. Los Angeles, CA: International MIDI Association, 1988.

[17] Dublin_Core_Metadata_Initiative, “Dublin Core Metadata Element Set, Version 1.1: Reference Description,” 2004.

[18] C. S. L. T. S. Committee, “IEEE Standard for Learning Object Metadata (1484.12.1-2002),” IEEE Learning Technology Standards Committee, 2002.

[19] World_Wide_Web_Consortium, “Synchronized Multimedia Integration Language (SMIL 2.0),” 2005.

[20] J. Brook, “The Development of an Asynchronous Video Reference Database to Support Suzuki Piano Instruction,” presented at Graduate Music Students' Association (GMSA) Conference, University of Ottawa, Ottawa, 2006.

[21] B. Emond, M. Brooks, and A. Smith, “A broadband Web-based Application for Video Sharing and Annotation,” in *Proceedings of Ninth International ACM Conference on Multimedia*, 2001, pp. 603-604.

[22] B. Emond and M. Brooks, “The Private Video Sharing and Annotation Server: A broadband application for teacher training and music education,” presented at International Lisp Conference, New York, NY, 2003.

[23] H. Masum, M. Brooks, and J. Spence, “MusicGrid: A case study in broadband video collaboration,” *First Monday*, vol. 10, 2005.

[24] B. Emond and R. L. West, “Cyberpsychology: A human-interaction perspective based on cognitive modelling,” *Cyberpsychology and Behavior*, vol. 6, pp. 527-536, 2003.

[25] A. Pentland, “Socially aware computation and communication,” *Computer*, vol. 38, pp. 33-40, 2005.

[26] I. McCowan, D. Gatica-Perez, S. Bengio, G. Lathoud, M. Barnard, and D. Zhang, “Automatic analysis of multimodal group actions in meetings,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, pp. 305-317, 2005.

[27] B. Emond, N. G. Vinson, J. A. Singer, M. A. Barfurth, and M. Brooks, “ReView: A Digital Video Player to Support Music Practice and Learning,” *Journal of Technology in Music Learning*, Accepted.

[28] G. Geisler, G. Marchionini, B. M. Wildemuth, A. Hughes, M. Yang, T. Wilkens, and R. Spinks, “Video browsing interfaces for the open video project,” in *CHI '02 extended abstracts on Human factors in computer systems*, 2002, pp. 514-515.

[29] G. Geisler, S. Giersch, D. McArthur, and M. McClelland, “Creating virtual collections in digital libraries: benefits and implementation issues,” in *Proceedings of the second ACM/IEEE-CS joint conference on Digital libraries*, 2002, pp. 210-218.

[30] C.-W. Lin, J. Zhou, J. Youn, and S. M.-T., “MPEG video streaming with vcr functionality,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, pp. 415-425, 2001.

[31] F. C. Li, A. Gupta, A. Sanocki, L. He, and Y. Rui, “Browsing digital video,” presented at Paper presented at the ACM Conference on Human Factors in Computing Systems, 2000.

[32] C. Drake and C. Palmer, “Skill acquisition in music performance: Relations between planning and temporal control,” *Cognition*, vol. 74, pp. 1-32, 2000.

[33] G. Mantovani, “The psychological construction of the internet; from information foraging to social gathering to cultural mediation,” *Cyberpsychology and behavior*, vol. 4, pp. 47-56, 2001.

[34] P. Pirolli and S. K. Card, “Information foraging,” *Psychological Review*, vol. 106, pp. 643-675, 1999.

[35] E. Wenger, *Communities of practices. Learning, meaning, and identity*. Cambridge: Cambridge University Press, 1998.

[36] G. Riva and C. Galimberti, “The mind in the web: Psychology in the internet age,” *Cyberpsychology and behavior*, vol. 4, pp. 1-5, 2001.

[37] M. A. Barfurth, J. Singer, B. Emond, N. Vinson, M. Brooks, and J. Spence, “Evaluation factors for multi-stakeholder broadband visual communication projects,” presented at Eleventh IEEE International Workshops on Enabling Technologies: Infrastrcutures for Collaborative Entreprises – WET-ICE, 2002.

[38] C. Galimberti, S. Ignazi, P. Vercesi, and G. Riva, “Communication and cooperation in networked environments: an experimental analysis,” *Cyberpsychology and behavior*, vol. 4, pp. 131-146, 2001.

[39] G. Riva, “The mind over the web: the quest for the definition of a method for internet research,” *Cyberpsychology and behavior*, vol. 4, pp. 7-16, 2001.

[40] C. Goldspink, “Methodological implications of complex systems approaches to sociality: simulation as a foundation for knowledge,” *Journal of Artificial Societies and Social Simulation*, vol. 5, 2002.

[41] J. M. Epstein and R. Axtell, *Growing Artificial Societies*. Cambridge, MA: MIT Press, 1996.

[42] M. Minsky, *The Society of Mind*. New York: Simon and Schuster, 1986.

[43] C. Goldspink, “Modelling social systems as complex: towards a social simulation meta-model,” *Journal of Artificial Societies and Social Simulation*, vol. 3, 2000.

[44] R. Conte and M. Paolucci, “Intelligent social learning,” *Journal of Artificial Societies and Social Simulation*, vol. 4, 2001.

[45] M. D. Byrne and J. R. Anderson, "Serial modules in parallel: The psychological refractory period and perfect time-sharing," *Psychological Review*, vol. 108, pp. 847-869, 2001.

[46] J. R. Anderson and M. P. Matessa, "A production system theory of serial memory," *Psychological Review*, vol. 104, pp. 728-748, 1997.

[47] J. R. Anderson and L. M. Reder, "The fan effect: New results and new theories," *Journal of Experimental Psychology: General*, vol. 128, pp. 186-197, 1999.

[48] J. R. Anderson, D. Bothell, C. Lebiere, and M. P. Matessa, "An integrated theory of list memory," *Journal of Memory and Language*, pp. 341-380, 1998.

[49] E. M. Altmann and J. G. Trafton, "Memory for goals: An activation-based model," *Cognitive Science*, vol. 26, pp. 39-83, 2002.

[50] B. D. Ehret, W. D. Gray, and S. S. Kirschenbaum, "Contending with complexity: Developing and using a scaled world in applied cognitive research," *Human Factors*, vol. 42, pp. 8-23, 2000.

[51] J. R. Anderson, R. Budiu, and L. M. Reder, "A theory of sentence memory as part of a general theory of memory," *Journal of Memory and Language*, vol. 45, pp. 337-367, 2001.

[52] B. Emond, "Modeling natural language comprehension and anaphora resolution with ACT-R," presented at Fourth Annual ACT-R Workshop, Pittsburgh, PA, 1997.

[53] J. Leitão, "An ACT-R Model of Syllogistic Reasoning," presented at Fourth International Conference on Cognitive Modeling, 2001.

[54] B. Emond, "Cognitive representations and processes in syllogistic reasoning: existential graphs and cognitive modelling," *Psychologica*, vol. 32, pp. 311-340, 2003.

[55] R. L. West and C. Lebiere, "Simple games as dynamic, coupled systems: Randomness and other emergent properties," *Cognitive Systems Research*, vol. 1, pp. 221-239, 2001.

[56] J. R. Anderson, Y. Qin, M.-H. Sohn, V. A. Stenger, and C. S. Carter, "An information-processing model of the BOLD response in symbol manipulation tasks," *Psychonomic Bulletin and Review*, in press.

[57] M. P. Matessa and J. R. Anderson, "Modeling Focused Learning in Role Assignment," *Language and Cognitive Processes*, vol. 15, pp. 263-292, 2000.

[58] D. Fum and F. Del Missier, "Modeling counter offer behavior in dyadic distributive negotiation," in *Proceedings of the Fourth International Conference on Cognitive Modeling*: Lawrence Erlbaum Associates, 2001, pp. 79-84.

[59] C. Lebiere, "Modeling group decision making in the ACT-R cognitive architecture," presented at Computational Social and Organizational Science (CASOS), Pittsburgh, PA, 2002.

[60] J. R. Anderson, "Spanning seven orders of magnitude: a challenge for cognitive modeling," *Cognitive Science*, vol. 26, pp. 85-112, 2002.

[61] J. R. Anderson, *The architecture of cognition*. Cambridge, MA, USA: Harvard University Press, 1983.

[62] J. R. Anderson and C. Lebiere, *The Atomic Components of Thought*. Mahwah, NJ, USA: Lawrence Erlbaum Associates, 1998.

[63] Z. W. Pylyshyn, *Computation and cognition: toward a foundation for cognitive science*. Cambridge, MA, USA: MIT Press, 1984.

[64] D. E. Kieras and D. E. Meyer, "An overview of the EPIC architecture for cognition and performance with application to human-computer interaction," *Human-Computer Interaction*, vol. 12, 1997.

[65] W. Kintsch, *Comprehension*. Cambridge, UK: Cambridge University Press, 1998.

[66] W. D. Gray and E. M. Altmann, "Cognitive modeling and human-computer interaction," in *International encyclopedia of ergonomics and human factors*, vol. 1, W. Karwowski, Ed., 2001, pp. 387-391.

[67] M. D. Byrne, "ACT-R/PM and menu selection: Applying a cognitive architecture to HCI," *International Journal of Human-Computer Studies*, vol. 55, pp. 41-84, 2001.

[68] B. Emond and R. L. West, "Using cognitive modelling simulations for user interface design decisions," in *Innovations in Applied Artificial Intelligence: Proceedings of the 17th International Conference on Industrial & Engineering Applications of Artificial Intelligence & Expert Systems*, B. Orchard, C. Yang, and M. Ali, Eds. Berlin: Springer-Verlag, 2004, pp. 305-314.