Advanced Digital Video Technologies to Support Collaborative Learning in School Education and Beyond

Carmen Zahn, Friedrich Hesse, Matthias Finke
Knowledge Media Research Center, Tuebingen,
Computer Graphics Center, Darmstadt,
Germany
c.zahn@iwm-kmrc.de

Roy Pea, Michael Mills, Joseph Rosen
Stanford Center for Innovations in Learning,
Stanford University, CA.
USA
roypea@stanford.edu

Abstract. The aim of the paper is to characterize two new advanced video technology software systems developed for uses in collaborative learning (DIVER[1] and HyperVideo[2]), and how they extend the paradigms of video use in classrooms today. The rationale for and characteristics of these tools are described, and early experiences with their educational uses are characterized.

Keywords: Cognitive tools, advanced video technologies in school education, teacher education, collaborative knowledge construction

INTRODUCTION

In school-based education of today, video is normally utilized as supplement to teacher lectures, in order to enrich regular lessons and to situate or visualize knowledge for a better understanding of a topic at hand. Empirical findings concerning the effectiveness of such video supported learning consistently show that comprehension and transfer of knowledge can be facilitated by dynamic (audio)visual formats in many domains - and that this is especially true for interactive presentations (Cognition and Technology Group at Vanderbilt, 1997; Park & Hopkins, 1993).

Introducing video into school-based education in the future could consist of utilizing *new* advanced digital video technologies as cognitive tools that broaden the spectrum of the existing video use paradigms: Advanced digital video has brought about new conventions of filmic expression in many areas—whether in the arts, at home or in the workplace. In many workplaces, advanced digital video technology is not only a means of communication via video conference, but is also used for collaborative video analyses, e.g. in the area of professional sports, teacher education or in the life sciences. Such advanced technologies may include tools for the selection of single video scenes from existing video information and for the direct integration of video scenes with e-communication facilities. Thus, we are confronted with a situation where we need to establish new components of visual literacy and digital literacy that relate to such work scenarios. Literacy concepts cannot be restricted to static and text-based media anymore, but have to integrate the understanding, analysis and active use of non-linear and audiovisual media as well including the use of digital video technology (Pea, 1991; Pea & Gomez, 1992).

To this end, advanced digital video technologies may become part of our educational systems and the question is then: How can such technologies (and new paradigms, respectively) be implemented in schools and in educational and learning processes broadly? How does the use of digital video technology interact with the cognitive concepts and prerequisites of students and teachers? Our two groups—in Germany and the United States—have been working in parallel on exemplars of a paradigm that is already a part of our everyday lives, but which has been minimally appropriated yet in K-12 education. In the present contribution, we will focus on these two advanced educational digital video systems that were explicitly developed on the basis of cognitive and socio-cognitive psychological and pedagogical considerations.

---

[1] DIVER™, WebDIVER™, Dive™ and “Guided Noticing”™ are trademarks of Stanford University for DIVER software and affiliated services with patents pending. The DIVER project work has been supported by grants from the National Science Foundation (#0216334, #0234456, #0326497) and the Hewlett Foundation. The DIVER team contributing to these efforts includes Roy Pea (Director), Michael Mills, Joe Rosen, Kenneth Dauber, and graduate students Robb Lindgren, Paula Wellings, Sarah Lewis and Lori Takeuchi.

[2] The HyperVideo system was developed at the Computer Graphics Center in Darmstadt, Germany in cooperation with the Knowledge Media Reasearch Center in Tuebingen, Germany.

DIGITAL VIDEO TECHNOLOGIES AND GROUP KNOWLEDGE PROCESSES

Learning to observe - Learning to analyze: The DIVER system was developed by the Stanford Center for Innovations in Learning. DIVER is based on the notion of a user “diving” into videos, i.e., creating new points of view onto a source video and commenting on these by writing short text passages or codes (Pea, Mills, Rosen, Dauber & Effelsberg, 2004). DIVER makes it possible to readily create an infinite variety of new digital video clips from any video record. A user of DIVER software "dives" into a video record by controlling—with a mouse, joystick, or other input device—a virtual camera that can zoom and pan through space and time within an overview window of the source video. The virtual camera can take a snapshot of a still image clip, or dynamically record a video “path” through the video to create a dive™ (which we also call a DIVER worksheet, see figure 1 below). A dive is made up of a collection of re-orderable “panels”, each of which contains a small key video frame that represents a clip, and a text field that can contain an annotation, code, or other interpretation. Diving on video performs an important action for establishing common ground that is characterized as “guided noticing” (Pea, in press). The use of the virtual camera for the framing of a focus within a complex and dynamic visual array directs the viewer’s attention to notice what it is that is thus circumscribed, and the point-of-view authoring thus guides the viewer to that noticing act. In this way, DIVER can be used as a tool to promote the development of “professional vision” in learning within disciplinary domains (Goodwin, 1994).

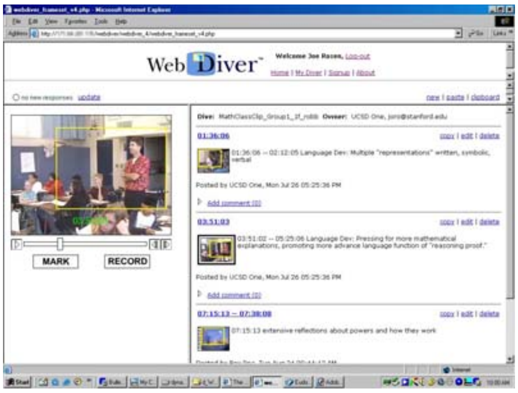

Figure 1: DIVER worksheet

Originally, DIVER’s primary focus was for supporting research activities in the learning sciences (such as interaction analysis: Jordan & Henderson, 1995), and in teacher education, where video analyses play a major role for understanding one’s own behavior and reflecting on it in relation to the behavior of others. DIVER has also been designed to enable the active exploration of panoramic video data—where one or more digital video cameras and associated mirrors are used to capture 360-degree horizontal imagery. In this case as well, the user may select visual information by virtually ‘pointing to it’ in the much larger spatio-temporal data structure of the video, for the purposes of collaborative reflection and analysis. The final product then is a collection of separate short video segments with annotations that represent the user’s point of view on the video.

There are two different ways users work with video using the DIVER approach. In the first, after creating a dive using the desktop DIVER application, the user can upload it onto WebDIVER, a website for interactive browsing, searching, and display of video clips and collaborative commentary on dives. In an alternative version of the WebDIVER system, one can dive on streaming video files that are made accessible through a web server over the Internet, without either requiring the downloading of a DIVER desktop application or the media files upon which the user dives. Using WebDIVER in either of these ways, a dive can be shared over the Internet among teachers, student-to-student, teacher-to-students, or in other scenarios with colleagues and become the focus of knowledge building, argumentative, tutorial, assessment or general communicative exchanges.

On a generic level the system might be described as a cognitive tool that enables “pointing to video” and thus helping to develop skills of observation and noticing details and enhancing the probability that in collaborative processes, the focus of attention and negotiating of meaning between participants in a conversation will build upon a common ground. With DIVER it becomes obvious that advanced technology may not only amplify existing kinds of activities and communication, but that it might augment our spectrum of activities and initiate entirely new forms of learning (Pea, 1985).

The DIVER system distinctively enables what its creators call “point of view” authoring of tours of existing video materials in a way that supports sharing, collaboration, and knowledge building around a common ground of reference (Pea, in press; also see Goldman-Segal, 1998 and Stevens et al., 2002 for related prior work). This form of communication with video is important for tapping the powerful potentials of video-enhanced learning in the classroom.

Learning to integrate text and video - learning to design: The web-based HyperVideo system for collaborative learning was developed at the Computer Graphics Center/Darmstadt in cooperation with the Knowledge Media Research Center/Tübingen. It is based on the idea of “annotating movies,” i.e. selecting video segments from a source video and having spatio-temporal hyperlinks added to video: Users of the HyperVideo system can create dynamic sensitive regions within video materials and add multiple links to these sensitive regions. The links can consist of data files uploaded from a local computer, as well as URLs. The links can then be discussed by means of an integrated e-communication tool. Thus, users can include their own annotations and knowledge in a video and share them with others in a group or community. The overall design approach encompasses several steps: (1) information is mainly presented by video, (2) knowledge can be collaboratively expanded by means of both dynamic links and written e-communication, and (3) the process of knowledge building is reflected in a resulting hypervideo structure we denote as a ‘dynamic information space’ of a collaborating group (*DIS*, Chambel, Zahn & Finke, 2004).The system is based on client/server architecture. The web-based graphical user interface (GUI) is shown in figure 2. A special video player (upper left part, see below) displays the spatio-temporal hyperlinks (white rectangles) within the video frame and offers functionalities in order to create new video annotations. The separated navigation space below the video player helps the users to navigate within the DIS. On the right side of the screen, additional information and the users’ discussions/comments are displayed. Newly created video annotations are immediately transferred from the client to the server in order to be instantly shareable by the community.

On a generic level, the HyperVideo system is as a cognitive tool enabling the linking of video information thus helping users to learn to establish non-linear information structures and to focus their attention and discussion in collaborative learning on associated concepts or related external representations of knowledge (e.g., a visible object and a text, or visible object and a formula). Such uses have been discussed in the context of performing collaborative hypervideo design projects (Chambel, Zahn & Finke, 2004). A respective program and the technology were evaluated and further developed during three psychology courses at the University of Muenster/Germany that were planned according to an instructional program based on courses of hypertext writing, originally developed by Stahl and Bromme (2004).

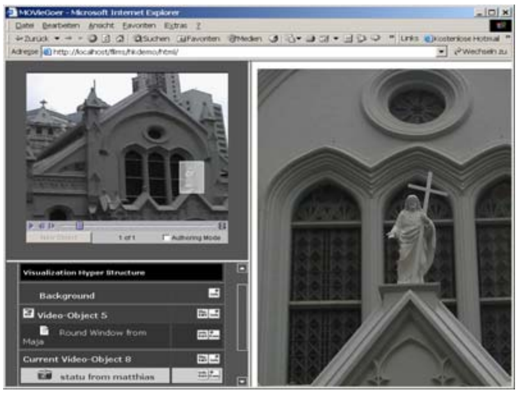

Figure 2: Graphical user interface (GUI) of HyperVideo

ADVANCED DIGITAL VIDEO SYSTEMS AS ‘RHETORICAL PROBLEM SPACES’ IN COLLABORATIVE SITUATIONS

Stahl and Bromme (2004) - applying Bereiter & Scardamalias (1987) model of knowledge transformation to the process of learning by hypertext design - assume that the peculiarities of hypertext may influence the process of learning in very specific ways: 1) Hypertexts are non-linear media, so hypertext design processes do not only include linear writing processes, but also the selection and creation of small “nodes” and the representation of concept relations by links and an overall structure (integration). Also multiple ways of “reading” the hypertext must be considered (e.g., multiple audience perspectives). This should lead learners to a very deep elaboration of content. 2) Hypertext design problems (due to their complex nature) are solved in cooperation and collaboration with others, so the production process has to be coordinated in a group. This should lead to collaborative knowledge building and knowledge exchange. 3) Hypertext design has just begun to emerge, so that even among professionals different ‘metaphors’ (= genre knowledge and mental models of the medium) can be applied. To be able to work and learn, students have to consciously develop and negotiate upon a joint idea of ‘what a hypertext is’ as a first step of their coordinated work. Finding an appropriate metaphor should lead to developing discourse knowledge, on the one hand, and further joint elaborations of the content, on the other hand. These assumptions are also substantiated by empirical results: The reflection of different audience perspectives has been found superior to not doing so. The thorough evaluation of links representing semantic relations between nodes has been found to lead to a deeper elaboration than not using such activities. And finally, a space metaphor showed to guide knowledge transformation processes better than a book metaphor of hypertext (Stahl & Bromme, 2004).

Similar assumptions can be made for hypervideo design processes, too. As was described in the previous section, we view advanced digital video technologies as cognitive tools according to a perspective of distributed intelligence (Pea 1993, 2004). Merging this view with the works on hypertext design fostering knowledge transformation processes (Stahl & Bromme, 2004), we generally perceive advanced digital video technologies as *establishing new rhetorical problem spaces with their own rhetorical rules*. These rhetorical problem spaces can well be understood in the sense of Bereiter and Scardamalia (1987) who assumed two problem spaces as important for text writing: the content problem space and the rhetoric problem space.

However, because in the present context we have to deal with digital video, the rhetoric problem spaces in hypervideo design tasks are (audio-)visual ones instead of being merely based on text. Consequently, the rhetorical rules of our new rhetorical problem spaces relate mainly to visual and filmic codes (such as mise en scène and montage), rather than relating solely to text. And hypervideo design as a rhetoric problem, finally, includes relating to different text genres, to the visual codes and styles of pictures/graphical displays *and* to the dynamic visual codes of film and animation. This, in turn, constitutes the educational value of such tasks.

COLLABORATIVE ACTIVITIES INVOLVING ADVANCED VIDEO TECHNOLOGIES

We are now exploring in pilot studies a variety of ways that collaboration among school students can be advanced in learning using advanced video technologies such as the two systems we have described. In WebDIVER, learners can collaboratively analyze video records from archival sources (e.g. science videos), or from video they have themselves collected (e.g. of fieldtrips, art museums, classrooms). In the HyperVideo system, learners and teachers can collaboratively create hypervideo documents (e.g. in university courses, as mentioned above) on the basis of existing or of self-shot videos. In both the Stanford and German software systems, collaborative video work can take place either face to face in a computer-intensive school setting or after-school club, or over computer networks, involving distant locations, either synchronously or asynchronously. These are the main features that distinguish the two systems from related works. Moreover, in both systems, learning scientists can also collaboratively engage with video, to interpret and analyze educational interactions or other behaviors of interest to their studies.

In preliminary work with the WebDIVER collaborative video analysis framework, we have found utility in the following scenarios: (1) pre-service secondary teachers in Stanford’s school of education, creating dives of ten-minute unedited videorecordings of their own teaching, which they analyze with respect to the rubrics which their faculty mentors use to evaluate their work; (2) learning science doctoral students collaboratively analyzing teaching videorecords according to different disciplinary perspectives (anthropology, linguistics, sociology, developmental psychology, educational psychology, cognitive science) and then working to combine them to deepen the quality of interaction analyses; (3) distributed researchers working to analyze video data from user studies, in this case, of preschool children interacting with a touch-screen video-based storytelling system we call KiddieDIVER, and providing a collective set of recommendations via a dive on these data that was shared with the software engineer over the web for review and implementations of software improvements based on insights from the collaborative video analysis activity; (4) faculty use in preparing dives on videos of secondary educational practices that are used in lectures to exemplify and explore theoretical concepts from the research literature used in their courses (e.g., cognitive apprenticeship, scaffolding, academic language); and (5) a film studies professor working with his students to compare several different film versions of the Shakespeare Play Henry V.

In each of these scenarios, we are finding that collaborative diving requires working in new rhetorical spaces, in cooperation and coordination with others. We make several points on the last scenario to exemplify the transformative nature of such activities with respect to common pedagogical methods: Film students spend considerable time studying major filmmakers, film genres, the grammar of cinematography (Metz, 1974), as well as narrative techniques and the animated special effects that have defined recent film developments. DIVER provides a new tool for the faculty member and film student to develop the web of perceptive knowledge that ties together the history of films, filmmakers, film methods and techniques and film criticism. In a film studies course now underway using DIVER at Stanford, graduate students in film are studying the relationship between the actor and the written work. For example, students are looking at two clips, the 1989 film adaptation of Henry V directed and played by Kenneth Branagh, and the 1944 film version of the same

Shakespeare play directed and played by Laurence Olivier. The same scene and words will be analyzed: Henry V's "Crispin's Day" speech. Previously the film studies professor provided a related assignment to students—describing in an essay what was different about each actor's interpretation—but by having them write about the movie scenes from memory. With WebDIVER, film students are able to point to specific space-time regions of the film in real-time examples from each movie, and to justify their analysis with video-based argumentation using the scenes from the movies being compared. This exercise takes place outside of the classroom, as a homework assignment. Each student is given their own protected workspace, and they access the films and the WebDIVER analysis tool on-line via a web browser. Students will then present their analysis in class, also using WebDIVER. The students will have a chance to comment on each other's work, both orally in class and again later on-line by adding messages and comments to the web-based Dive worksheets. Although this same assignment has been used in film class before, this will be the first time a) students will be able to point directly to the scenes they're analyzing and referencing; and b) an informal learning discussion (via web page collaborative commentary) will continue outside of the classroom presentations. In WebDIVER, students can also literally navigate the movie by way of the actor's/script's utterances (i.e. click on an utterance and go directly to the corresponding scene in the movie). The utterances also scroll along with the movie. The professor anticipates a nuance and depth to analysis that he has not experienced using his previous approach to instruction and assessment.

Prior studies involving the HyperVideo system include an experimental test of how users (N= 74) learn with different design versions of a hypervideo in the domain of biology (Zahn, Barquero & Schwan, 2004) and a comparison of how (and where) authors with different prior knowledge would suggest placing hyperlinks in biology videos (Zahn, Schwan & Barquero, 2002). The results of this latter study revealed that authors of different knowledge backgrounds (content-experts, media-experts and novices) developed similar ideas of a hypervideo structure, which were mainly based on formal features of the source video (such as, for example, terms included in the audio track). Results also showed that the linking decisions of expert-authors were quite congruent with those of novice users, indicating that even users with low prior knowledge were capable to make meaningful linking decisions.

Our prior works provide the basis for applying hypervideo design tasks at school. As a starting point, hypervideo design will be applied in German secondary schools to support media education in German native language lessons (“Deutschunterricht”). The topic will be TV-advertising. We plan to study the collaborative analysis of TV-ads based on the DIVER system and the collaborative hyperlinking of TV-ads based on the HyperVideo system. Altogether, we will conduct two large experiments in a learning lab. Our interest is to investigate the interactions of DIVER and HyperVideo as two generic types of digital video technology with a) individual cognition (i.e. mental models of “hypervideo” in learners) and b) teacher’s instructions (i.e. the support of group discussion by teachers) and the influences of these interactions on group knowledge processes. This future-orientation leads us to the last section of this paper.

CONCLUSION

In writing about modern music, writing, art and science, Umberto Eco (1989) notes that "‘open’ works, insofar as they are in movement, are characterized by the invitation to make the work together with the author and that (2) on a wider level (as a subgenus in the species ‘work in movement’) there exist works, which though organically completed, are ‘open’ to a continuous generation of internal relations that the addressee must uncover and select in his act of perceiving the totality of incoming stimuli." To the extent that DIVER and HyperVideo use can make video and movies and other rich media 'open' to HyperVideo linking and to Diving—interpretation and extensible use with guided noticing, DIVER path movie-making making and annotation—there is without question an active role for the reader, who becomes an author in bringing the work of the video or other medium to a more completed state in his or her interpretations of it. DIVER also provides a tool for evidence-based argumentation, in which one uses what one notices in the medium to make a case around it, and thus extends the work in significant ways with the act of authoring the dive. For the constructivist educator or more generally for those who want a more active voice in media uses for communication and knowledge production, these two systems exemplify a video use paradigm for education that moves away from today’s broadcast-centric and asymmetric uses of video to the communicative empowerment of the video user, who can easily craft point-of-view movies within movies with commentaries and hyperlinks to share with others. We view this fundamental shift from consumption to authorship of video points-of-view as a vital transformation in the use of the video medium for advancing learning and education.

REFERENCES

Barron, B. J., Schwartz, D. L., Vye, N. J., Moore, A., Petrosino, A., Zech, L., & Bransford, J. D. (1998). Doing with understanding: Lessons from research on problem and project-based learning. *The Journal of the Learning Sciences, 7* (3&4), 271-311.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition.* Hillsdale N.J.: Erlbaum.

Chambel, T., Zahn, C., & Finke, M. (2004). *Hypervideo design and support for contextualised learning.* Paper presented at the *IEEE 2004*, Finland.

Cognition and Technology Group at Vanderbilt. (1997). *The Jasper Project: Lessons in curriculum, instruction, assessment, and professional development.* Mahwah, NJ: Lawrence Erlbaum Associates.

Eco, U. (1989). *The Open Work*. Cambridge, MA: Harvard University Press.

Goldman-Segall, R. (1998). *Points of viewing children's thinking: A digital ethnographer's journey.* Mahwah, New Jersey: LEA.

Goodwin, C. (1994). Professional vision. *American Anthropologist. 96*(3), 606-633.

Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *The Journal of the Learning Sciences, 4,* 1, 39-103.

Metz, C. (1974). *Film Language: a Semiotics of the Cinema*. New York: Oxford University Press.

Park, O., & Hopkins, R. (1993). Instructional conditions for using dynamic visual displays: A review. *Instructional Science, 21*, 427-488.

Pea, R. (1985). Beyond amplification: Using the computer to reorganize mental functioning. *Educational Psychologist*, *20*, 167-182.

Pea, R.D. (1991, July). Learning through multimedia. *IEEE Computer Graphics and Applications, 11*(4), 58-66.

Pea, R. D. (2004). The social and technological dimensions of "scaffolding" and related theoretical concepts for learning, education and human activity. *The Journal of the Learning Sciences, 13*(3), 423-451.

Pea, R. D. (in press). Video-as-data and digital video manipulation techniques for transforming learning sciences research, education and other cultural practices. To appear in J. Weiss, J. Nolan & P. Trifonas (Eds.), *International Handbook of Virtual Learning Environments*. Dordrecht: Kluwer Academic Publishing.

Pea, R. D., & Gomez, L. (1992). Distributed multimedia learning environments: Why and how*? Interactive Learning Environments, 2*(2), 73-109.

Pea, R., Mills, M., Rosen, J., Dauber, K., Effelsberg, W., & Hoffert. E. (2004, Jan-March). The DIVER™ Project: Interactive Digital Video Repurposing. *IEEE Multimedia, 11*(1), 54-61.

Roschelle, J. M., Pea, R. D., Hoadley, C. M., Gordin, D. N., & M., Means B. (2000). Changing how and what children learn in school with computer-based technologies. *The Future of Children - Children and Computer technology, 10*(2), 76-101.

Stahl, E. & Bromme, R. (2004). Learning by writing hypertext: A research based design of university courses in writing hypertext. In H. Van Den Bergh G. Rijlaarsdam, & M. Couzijn, (Ed.), *Effective learning and teaching of writing* (2 ed., Vol. 14, pp. 547-560). Dordrecht: Kluwer Academic Publishers.

Stevens, R., Cherry, G., & Fournier, J. (2002). Video Traces: Rich Media Annotations for Teaching and Learning. *Proceedings of the CSCL 2002 Conference on Computer Supported Collaborative Learning,* Boulder, CO.

Zahn, C., Schwan, S. , & Barquero, B. (2002). Authoring hypervideos: Design for learning and learning by design. In & E. Stahl & R. Bromme (Ed.), *Writing hypertext and learning: Conceptual and empirical approaches* (pp. 153-176). London: Pergamon/Elsevier Science.

Zahn, C., Barquero, B., & Schwan, S. (2004). Learning with hyperlinked videos – design criteria and efficient strategies of using audiovisual hypermedia. *Learning and Instruction*.