

YouTube Annotations: Reflecting Interactive, Web based Hypervideos in Teacher Education

Thomas Winkler¹, Martina Ide², Michael Herczeg¹
Institute for Multimedia and Interactive Systems, University of Luebeck, Germany^{1,2}
Institute for Quality Development of Schools in Schleswig-Holstein, Germany¹
winkler@imis.uni-luebeck.de¹, m.ide@gmx.de², herczeg@imis.uni-luebeck.de³

Abstract: The aim of this paper is to discuss the pertinence of experience of hypertext structures regarding the development of perceptual patterns when using contemporary time based interactive media. Therefore we discuss first the idea of time based hypermedia and related works in arts and education. Then we describe the design and implementation of the pilot project *WeHype*¹, where hypervideos were created at school using YouTube Annotations. Then we focus on the transfer of the experiences at school into teacher education. Results of a comparative evaluation verify that, due to the fact that youngsters communicate and interact more in social nets, they are more familiar with hypermedia spaces than their teachers. Furthermore the evaluation shows, that teachers as well as students think that the easy-to-use tool YouTube Annotations is suitable for understanding hyper-structures and leads to sustainable learning. This leads us to reflect about the importance of artistic approaches for understanding non-linear narrative structures using video. Finally we discuss why teaching with interactive time based hypermedia should be transferred into teacher education.

Introduction

Media have always influenced our learning, whether as a linguistic or non-linguistic (i.e. hypervideo) communication media or as media which (pre-)structure the interactions with the objective world. In the beginning of the 21st century, the significance of the mediated reality of human life changed fundamentally. These changes must be incorporated into teacher training programs. Communication and interaction (between people and between people and digital media, i.e. hyperlink structures) have developed a symbiosis with highly complex information processing systems, both in post-geographical environments (Fassler, 2009), as well as in mixed reality. Teacher training has to reflect and honor these social changes in modified teaching scenarios, so that students are prepared adequately for future challenges in the 21st century. One of the main interests of the KiMM initiative² in co-operation with the IQSH³ is the concentration on the design, development, testing and evaluation of digitally enriched interactive learning environments for children and teenagers (K-12). We have worked for more than ten years in this field to bring a constructivist pedagogical methodology into new, contemporary scenarios of teacher education.

It is important that the learning scenarios are created and evaluated with children and youngsters in mind and can easily become part of daily practical teaching at schools. This means appropriate use of available technology; i.e., it must be inexpensive and easy to handle. Additionally, the technology and the associated learning scenarios are meant to be directly incorporated into teacher education. A prerequisite for the design of up-to-date teaching scenarios is a current understanding of structures of learning processes. Therefore, it also allows for the planning of a model project with students at school, considering changing societal, cultural and social needs in terms of a new

[1] *WeHype* is the name of an arts project in the 11th grade of the Carl-Jacob-Burckhardt Secondary School, UNESCO Associated School in Luebeck, Germany.

[2] Kids in Media and Motion, Initiative of the University of Luebeck, Germany. <http://kimm.uni-luebeck.de>

[3] Institute for Quality Development of Schools in Schleswig-Holstein (Teacher Education, Germany). <http://www.iqsh.de>

learning culture in the 21st century. The novel, unprecedented forms of perceiving and interacting (i.e. hypervideo), have mediated entirely new modes of construction of knowledge (in other words, of learning processes).

For the education and training of teachers it must be stated that young people especially acquire skills outside the educational system (appropriation (construction out of fragments) of other/new scripts (schemata) of perception), which are nonetheless introduced into the educational situation (school). Hyper-structures of the Internet have a strong effect on patterns of perception in young people. Located in all processes of the Internet (chat, blogs, YouTube etc.) hyper-structures are evidence of the change in the communication culture of young people (JIM Studie 2009). According to Franz Josef Roell (Roell, 2003) this is appropriation of reality in the sense of self made *crossmodal* skills of children and young people. This way of thinking is supported by the culture of media, which de-linearizes *several* levels of perception of information while simultaneously pursuing several sources. This perception does not usually correspond to the scripts still prevalent in mono-causal forms of teaching in school and demonstrates the need to think about new approaches in education to prevent educational problems. The principle of Internet communication leads to changes in organization and modes of thought and time-space perception and demonstrate “... *The relevance of hyper-media experience in the development of perception-script development ...*” (Röll, 2003): Information transfer takes place by fragmentation and connection. That these forms of hyper-structures are not fundamentally new, is shown by Vannemar Bush (Bush, 1945), Douglas Engelbart (Engelbart, 1962) and Theodor Nelson (Nelson, 1965). It is not only essential but possible to navigate with hypertext; indeed, it has become the central principle of appropriating reality of Internet communication, because it breaks the prevailing thinking (sequential thinking) in linear structures.

What is Hypermedia?

Conceptually close to branching-type interactivity (in which elements are connected using a branching tree structure), hypermedia reveal the principle of variability: “*We can think of all possible paths through a hypermedia document as being different versions of it. By following the links, the user retrieves a particular version of the document.*” (Manovich, 2001) Branching-type interactivity means menu-based interactivity. That means programs in which all the possible objects the user can visit are accessible from a branching tree structure (Manovich, 2001). The hyper structure is reticular, comparable to associative thinking (cluster development). Links from and to objects lead to a view of the document which yields meaning according to associative semantic connotation to it. The advantage of a hyper-structure is that no categorical, tree-like planning systems for appropriation of the terms are necessary. The design of a hypertext structure is now much more than another form of display level; it is the ability to see sets of information in their interactions with each other in linked contexts. It also implies a playful mind, does not engender mono-causal solutions regarding the explicability of a material property, but uses the navigation potential of user interaction for a variable perspective. In interactive videos hotspots are an example of enabling a non-conceptual interaction. Today, we are able to change the *nature of a story* by the computer by making it an interactive experience (non logical composition). The recipient becomes the producer; the idea of the absolute text, image, etc. no longer exists. “*An interactive narrative is a narrative in which the audience can effect a significant change on the narrative. Navigation means that the audience is able to direct the story, choose different point-of-views (engagement, empathy, and desire to know what happens next). An open structure leaves the narrative creation up on the user – it’s all interaction*” (Manovich, 2001).

Related Work in Arts and Education

Lew Manovich describes within his work *Soft Cinema: Navigating the Database* how a story develops by selecting scenes from a given collection. In contrast to traditional cinema, which prizes narrative as the key form of cultural expression in the modern age, digital media lack this strong narrative component. Using a database, the user can start or stop at any point. “*As a cultural form, database represents the world as a list of items and it refuses to order this list. In contrast, a narrative creates a cause-and-effect trajectory of seemingly unordered items (events). Therefore, database and narrative are natural enemies. Competing for the same territory of human culture, each*

claims an exclusive right to make meaning out of the world.” (Manovich, 2001) In his work *Soft Cinema* (Figure 1) Manovich shows a possibility for representing subjective experience living in a global information society. In contrast to a linear Narration, there is a random algorithm, which combines images again and again in real time. *Soft Cinema* breaks the screen into a number of frames; possibilities of simultaneous multiplicities of perspectives are reflected. In an analogy with hypertext, Manovich (Manovich, 2001) names an interactive narrative “*hyper-narrative*”. It can be understood as the sum of multiple trajectories through a database. For this reason traditional linear narrative can be seen as a particular case of a hyper-narrative.

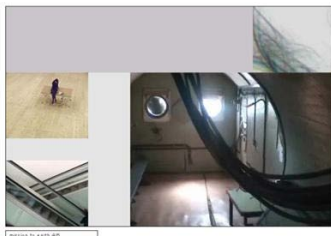


Figure 1: Lev Manovich, *Soft Cinema* (1999-2002), screenshot of the multi video screen

The fictional art work *Portrait One* (1990) by Luc Courchesne focuses interactivity in the context of a hyper-narrative structure. An installation enables a user to encounter a virtual character named Marie. Courchesne reflects upon the meaning of the concept of portrait and uses current media while extending the principle of the portrait of classical painting. The user must “*wait for the magic moment in complicity*” at the chance to communicate interactively with the portrait. “*I use hypermedia to make portraits. ... In my portraits, the entire encounter is recorded, and material is extracted to construct a mechanics of interaction that will allow visitors to conduct their own interviews. As this happens over time, the conversations will evolve toward more intimate considerations.*” (Courchesne, 1994) Similarly to Manovich, later Luc Courchesne puts video in a database which can be accessed via text blocks (Figure 2). Due to the free navigation of the user in a hyperspace, it opens pluralistic methods of communication with the virtual counterpart. The anticipated contents of the dialogue, annoying in its complexity to our consciousness, raise questions about the portrait. Thus, the user is a co-author of work, he can choose the path to follow; he creates the *language* of interaction in an active and vital way. New cognitive demands occur for the viewer.



Figure 2: Luc Courchesne, *Portrait One* (1990)

For several years, the relevance of Hypervideo in education has been discussed regarding learning scenarios. The conception of what characterizes a Hypervideo assumes that a video will be provided with links (hotspots) for additional information to explore other related or complementary media. So Carmen Zahn et al. (Zahn, 2005) summarize that “*in school-based education of today, video is normally utilized as supplement to teacher lectures, in order to enrich regular lessons, ... to visualize knowledge for a better understanding of a topic at hand.*” It is undisputed that in educational and learning processes video technology is mostly used for collaborative video analyses to understand and reflect one’s behavior in relation to the behavior of others, and/or to analyze the lesson’s structure, content or aim of a learning scenario. As a meaningful development, the authors refer to the tool *DIVER*⁴ (Figure 3).

[4] DIVER is a Stanford Center for Innovations in Learning project initiated in 2001 but with its roots in earlier work by Roy Pea, Jeremy Roschelle (Institute for Research and Learning) and Randall Trigg (Xerox PARC) in

It is “based on the notion of a user “diving” into videos, creating new points of view onto a source video and commenting on these by writing short text passages or codes” (Pea, 2004). This digital tool makes it possible to readily create an infinite variety of new digital video clips from any video record (Zahn, 2005). Diving into videos with a virtual camera, it is possible to record a path through the video to create a dive (Pea, 2004), to represent the viewer’s point of view. This tool mainly describes how to “structure” processes of understanding (to trace a path of individual understanding about something), while noticing, selecting or pointing out details.



Figure 3: WebDiver



Figure 4: *Your Food is Your Mood*, 2007: Navigating hypervideos in a walkable cube, using barcodes on objects.

The Hypervideos annotated by WebDiver are not a complex hyper structure; they are not reticularly structured, in fact they remain in a sequential arrested mental model. In our daily life we do not react only to the reality, but interpret *things* with the help of mental models. They are constructed when they are needed to cope with a learning situation (Halford, 1993). While the learners have to form a mental model to understand the idea of “hyperstructures”, they do it *crossmodally*, using *multifunctional intelligence* (Gardener, 1999). The construction of *non-linear, verbal, visual, and acoustic hyper-learning-structures*, as they are implemented in *WeHype*, systematically use the different symbol systems of our different regions of intelligence while constructing dynamic, interactive representations of complex issues.

Concept and Realisation of the Project *WeHype*

The concept of the project *WeHype* (2009) was carried out in an arts course at 11th grade in a secondary school in Luebeck, Germany. It focused a de-linear structure and no longer represents the traditional function of storytelling of a whole which has a beginning, a middle and an end; exposition, rising action, crisis climax, falling action, denouement, so that the author no longer has the most narrative control (determination). Within new media, the recipient is co-author of the story space; he has the opportunity to join at any node in the hyperspace and to navigate in any direction he chooses. For the conceptual planning of *WeHype* the above named modes of reception and perception scripts of teenagers were central. The design of the hypervideos integrated the form and content of everyday media, aware of the forms of communication of the young people: the cell phone. It is available for everyone, situationally flexible, primarily by the possibility, as an extension of the body, of reflecting each adapted perspective (e.g. persons located in motion) on the world. While the project “*Your Food is Your Mood*” (Winkler, 2007) already enables the experience of four simultaneous perspectives via video on the theme “*Food and digestion*” (by linking the physical space to the digital using physical objects with barcodes within a walkable cube), thus providing facets of the theme in a multiperspective and interactive way (Figure 4). But in *WeHype* web-based videos are linked via “hotspots” (one highlighted, sensitive points in the video).

Since the hyper-videos can and should be found via search engines on the internet, a de-linear cross-linking is embedded and enables an x-anywhere access to the hyper-structure for the recipient. In the thematic direction of the project, all the content, documented by the students in teams, focus on facets of life regarding the city of Luebeck (northern Germany). They focused on places and encounters, to visualize personal, social or educational perspectives. For the aesthetic process it was important to reflect and integrate the quality of cell phone video: in particu-

1990-91 on VideoNoter. DIVER is a project devoted to creating and integrating tools for enhancing the activities of exploring and reflecting on digital video records of learning and teaching.

lar the *inferior* image quality of cell phone video, so conceptually *the blur, the amateur-like, and the sound*. In *WeHype*, a collection of hyper-structures are glued together giving the user options to choose from. This makes the plot flexible enough but still substantial. This visual representation of diverse, interconnected story lines requires the ability of multiple, associative thinking on the part of students. Similar to *Soft Cinema*, in which a random algorithm plays scenes from a database, in *WeHype* it is not possible to see all the possibilities of perspective at any given time. But in contrast to *Soft Cinema*, the user of *WeHype* has to navigate actively to open up the complexity of the content gradually. Introduction to the topic of the project *WeHype* at school was initially to deal with changing forms of reception and principles of composition (departure from cause-and-effect relationships) of various film productions: In *TimeCode*⁵ a screen is divided into four quarters and the four shots are shown simultaneously. The sound mix of the film is designed so that the most significant of the four sequences on screen dominates the soundtrack at any given moment. *Deine Wahrheit* (Melzer 2004) shows a non-linear perspective-movie on DVD. The introduction in the subject matter of *WeHype* was first the analysis of modified forms of reception and principles of composition of different film productions (renunciation of causal mechanism). It follows the principle of a tree structure. The user is able to choose between competing perspectives at various times. *Memento* (Nolan, 2000) visualizes two approaches, marked by color and black and white illustrations indicating chronological and non-chronological elements of the plot. The recipient is situated in the story but without knowing the prehistory. So he feels wrong in a way. With the possibility of navigation through the interactive DVD of the installation *Lorna*, the recipient is able to open up details of past, present and personal conflicts of Lorna dynamically (Hershman, 1984).

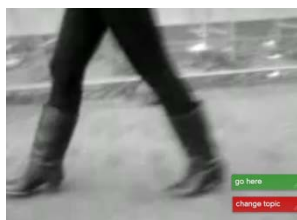


Figure 5: Hotspots in a hypervideo of *WeHype* on YouTube, 2009

In *WeHype*, seven teams (Figure 6) of students each worked on a topic with 9 video clips 20 seconds in length. These clips formed the basis of an initial internal linking, in addition to a superior one linking all existing singular projects into a hyper-structure. The smallest semantic units acting as a liaison between all topics are the visual elements of *shoes* and *stairs*, those provided with the link (hotspot) “*change topic*”, enabling perspective changes regarding the whole subject (Figure 5). The navigation within a topic is enabled by two hotspots on each video, named *go here* and *go there*. The medium of the link is therefore only an option, without evidence of underlying information, focusing the image as a medium of information.⁶ In comparison, the project *art-portrait*, 2010⁷ (also built in the 11th grade arts course) put a clear focus on the aspect of visual communication and language. Depending on the selection of an interaction-text (link option via text) by the recipient, the encounter occurs with a virtual portrait, multiple areas of thought are opened up. The structurally high choice of textual hyperlinks for the user increases the density of navigation and represents forms of complex thought patterns.

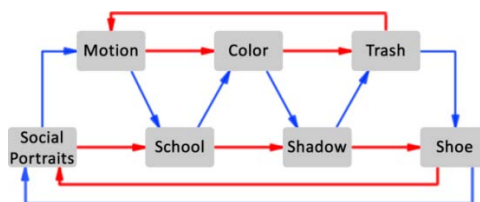


Figure 6: Main structure of *WeHype*. Links between topics/projects: red = shoe, blue = steps

[5] Timecode at Wikipedia: [http://en.wikipedia.org/wiki/Timecode_\(film\)](http://en.wikipedia.org/wiki/Timecode_(film))

[6] *WeHype* on YouTube: <http://www.youtube.com/watch?v=sNKiDutaSng>

[7] *Art-Portrait* on YouTube: http://www.youtube.com/results?search_query=KunstPortr%C3%A4t&aq=f

The potential of hypervideo-based learning in the classroom is to draw the pre-selected non-linear video-scenes, as an endless collection of individual items, into one’s interpretation of these. Developing learning spaces in creating new hypervideo structures influence processes of learning in a specific way: Hyperstructures are non-linear; they include multiple ways of “*decryption*”, because the order of the presentation is not fixed. As an appropriate tool to create a web-based hyper-structure, characterized by “non-hierarchically” linking structure of videos to create stories with multiple possibilities, the free tool *YouTube Annotations* (open source since 2008) was used. This tool challenges the tradition of the screen (frontal viewing) by fragmenting in nodes (scenes) and linking them into a screen of “*non-hierarchical linked videos*”. Thus, while the production processes were realized in groups producing hypervideos (as a collection of fragments), with different perspectives on the general topic among the groups, students develop a deep elaboration of content (visual, acoustic, etc. codes) and reflect on intermediality as well.

Transfer into Teacher Education

A multiplicity of workshops focusing an interdisciplinary approach was accomplished within teacher education and further teacher education regarding all subjects and different kinds of schools.⁸ Therefore a module for teacher education was derived from the results of the evaluation of the pilot project *WeHype*. This teaching module structures workshops regarding the use of non-linear hyper structures in pedagogical processes, concerning the special field “*Learning with Digital Media*” within teacher education. The scaffolding of the teaching module is developed out of an understanding of hypermedia structure in an arts project context. The content includes comprehensive material on innovative pedagogical processes in schools.

Comparative Evaluation of *WeHype* and Teacher Education

The pilot project took place in the 11th grade arts class with 16 (7 female and 9 male) students at a secondary school (average age 16.5 years). The transfer in further teacher education took place with 17 (12 female and 5 male) teachers (average age 42 years). Concerning the answers to the questions two scales were used. One scale with yes and no answers show the results in percentage of yes-answers. The other scale reaches from 1 = *does not apply at all* to 6 = *applies fully and completely*. The mean (arithmetic average), the standard deviation (measure of the statistical dispersion), and the median (relative frequency, not influenced by extremes) were calculated. The whole evaluation questionnaire consists of more than 50 questions. Here we present only the most important findings regarding the research questions of this paper.

The survey of the students shows clearly, that the design and implementation of the project *WeHype* with YouTube Annotations was motivating. All have worked with pleasure on it (Figure 7).

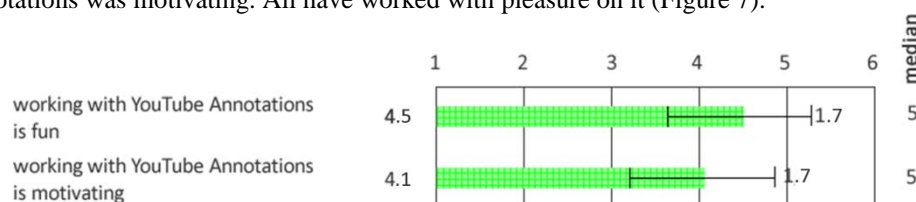


Figure 7: Fun and motivation by students at school using YouTube Annotations

The evaluation clarifies the difference between students and teachers in using social ware and their respective experience using and constructing hyper structures. While the surveyed students communicate mostly through social portals, like *SchülerVZ* (Germany), this behavior is poorly developed among teachers. Only in regard to watching movies on YouTube, the group of teachers shows a distinct usage pattern (3.7 vs. 4.8 of students). In contrast to this result, teachers are less familiar with loading up videos on *YouTube* or *Vimeo* than the students are. But the most indicative fact is that students have previous experience annotating videos (27% of students vs. 6% of

[8] In the state *Schleswig-Holstein*, in northern Germany

teachers). This evidence of the high emersion of young people in processes of the Internet points to the changing structures of appropriating reality. It is obvious that these structures differentiate teachers from students (Figure 8).

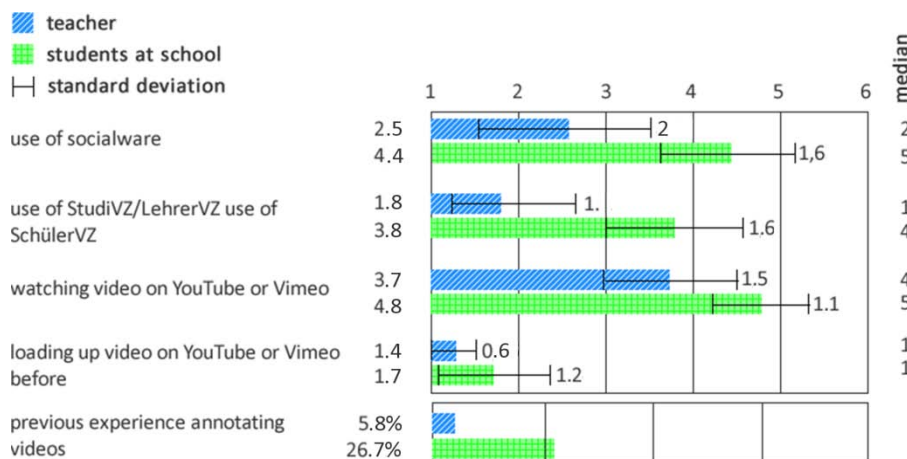


Figure 8: Communication with social ware and experience using and constructing hyper structures

Also the number of students and teachers creating photos and videos using cell phones confirm the divergent behavior of the two groups, although at least 35% of the teachers are creating video using cell phones. More than 53% of them use the photo feature of the cell phone (Figure 9).

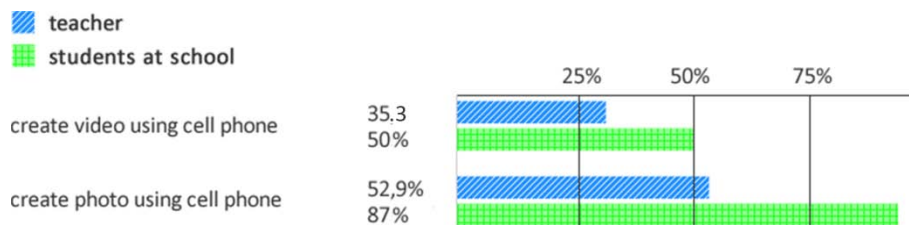


Figure 9: Amount of students and teachers creating photos and videos using cell phones

Almost all of the students as well as the teachers said that *YouTube Annotations* is easy to handle. Both groups pointed out clearly that the tool is suitable to illustrate the complexity of topics. They believe that interactive linking of content promotes the understanding of hyper-structures. Both teachers and students express that visualizing und linking videos leads to sustainable learning processes. Compared to the teachers, it is less important for students to work with *YouTube Annotations* at school, probably because young people already have much more competent behavior in dealing with contemporary media (Figure 10).

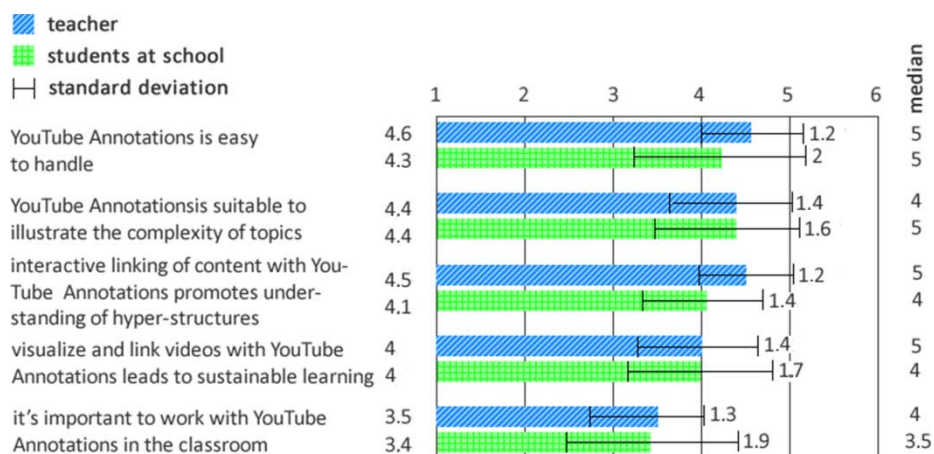


Figure 10: Handling and assumption of emerging value added by learning with *YouTube Annotations*

Additional Evaluation of Pre-Teachers

In a further evaluation with 13 (10 female and three male) pre-teachers (average age 31 years) in relation to the results of the comparative evaluation, the following issues were considered.

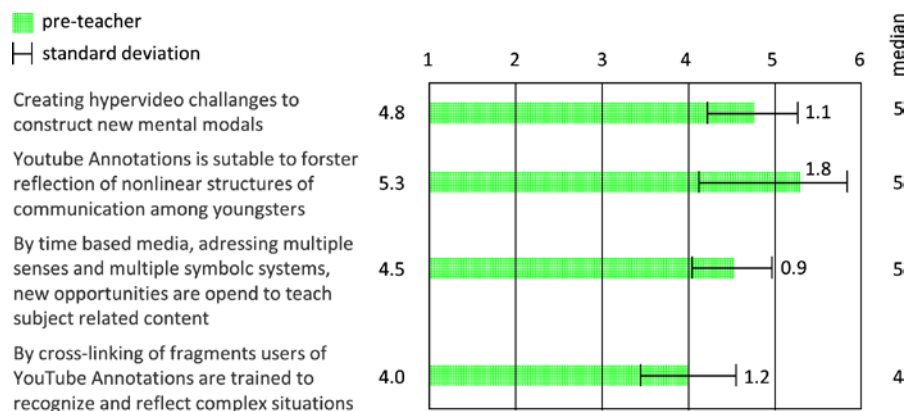


Figure 11: Building up new mental models based on crossmodal perception leads to reflection of complexity

Conclusions and Further Work

As we briefly mentioned at the beginning of the paper, digital media of the 21st century, e.g. *YouTube Annotations*, are more than just practical tools. They have the opportunity to create new learning scenarios. Those encourage key competencies such as the ability to build up new mental models based on crossmodal perception, reflecting reality and new options of structures of behavior (Röll, 2003). For this reason, focal points should be newly shaped and implemented in the education of arts teachers. State of the art teaching scenarios must be forward-thinking and integrate the potential usages of digital media into teaching modules. Finally, we believe that the teaching modules make it possible for teachers to capture the interest of high school students, since they penetrate the world of the Digital Natives and offer them an environment in which they can critically explore their world. The high volume of requests for workshops about *YouTube Annotations* and the positive feedback demonstrate the importance of the interdisciplinary (arts and computer science) offering for the teaching of pre-teacher and teacher in service.

References

- Bush, V. (1945) *As We May Think*. In: *Atlantic Monthly*, July 1945, V. 176, Nr. 1, p. 101–108
- Engelbart, D. (1962) *Augmenting Human Intellect: A Conceptual Framework*. Summary report of the SRI Project No. 3578 at Stanford Research Institute.
http://loan.stanford.edu/mousesite/EngelbartPapers/B5_F18_ConceptFrameworkInd.html
- Fassler, M. (2009) *Nach der Gesellschaft. Infogene Welten – anthropologische Zukünfte*, Wilhelm Fink Verlag, Paderborn, Germany.
- Gardner, H. (1999) *Intelligence Reframed. Multiple intelligences for the 21st century*, New York: Basic Books.
- Halford, G.S. (1993). *Children's understanding. The development of mental modals*. Erlbaum, Hillsdale, NJ, USA
- Luc Courchesne L. (1994) *Family Portrait : The Art of Portraiture*. Luc Courchesne : Interactive Portraits (Ottawa : National Gallery of Canada.
- Pea, R., Mills, M., Rosen, J., Dauber, K., Effelsberg, W., Hoffert, E. (2004). *The DIVER Project: Interactive Digital Video Repurposing*. *IEEE Multimedia*, 11 (1), Jan-March 2004. pp 54-61.
- Manovich, L. (2001) *The Language of New Media*. Cambridge: MIT Press.
- Manovich, L. (2006) *The Poetics of Augmented Space*. *Visual Communication*, Vol.5, No.2, London. pp. 219-240.
- Melzer, A. (2004) *The Interactive and Multi-protagonist Film: A Hypermovie on DVD*. In: *Proceedings of the 3rd International Conference on Entertainment Computing, ICEC 2004, LNCS 3166*, Springer, Berlin Heidelberg New York. pp 193-203.
- Nelson, T. (1965). *Complex information processing: a file structure for the complex, the changing and the indeterminate*. *ACM/CSC-ER Proceedings of the 1965 20th national conference*.
- Winkler, T., Ide-Schoening, M. & Herczeg, M. (2007). *Sustainable Teaching through the use of Media Art Technology - Creating Biological Knowledge by Designing a Multimodal Interactive Tangible Media Installation*. In Carl- sen, R., Mc Ferrin, K., Price, J., Weber, R. & Willis D.A. (Eds.) *Proceedings of the SITE 2007*. Chesapeake, USA: AACE. pp. 2155-2162.
- Zahn, C. at all. (2005) *Advanced digital video technologies to support collaborative learning in school education and beyond*, *Proceedings of the 2005 conference on Computer support for collaborative learning: learning 2005: the next 10 years!* International Society of the Learning Sciences, Taipei, Taiwan, pp. 737 – 742.
-
- JIM Studie 2009 (Basic study in Germany on youngsters age 12 to 19 handling media):
www.mpfs.de/fileadmin/JIM-pdf09/JIM-Studie2009.pdf

Proof of pictures:

Fig. 1: <http://www.medienkunstnetz.de/works/soft-cinema/images/3/>

Fig. 2: <http://www.fondation-langlois.org/html/e/page.php?NumPage=158>

Fig. 3: WebDiver: <http://diver.stanford.edu/webdiver.html>