ELSEVIER

Learning and Instruction 14 (2004) 293–305

Learning and Instruction

www.elsevier.com/locate/learninstruc

The cognitive benefits of interactive videos: learning to tie nautical knots

Stephan Schwan [a,*], Roland Riempp [b]

[a] *Institute for Education and Psychology, Johannes Kepler University, Altenbergerstr 69, A-4040 Linz, Austria*

[b] *Department of Media and Information, Fachhochschule Offenburg, University of Applied Sciences, Offenburg, Germany*

Abstract

In contrast to their traditional, non-interactive counterparts, interactive dynamic visualisations allow users to adapt their form and content to their individual cognitive skills and needs. Provided that the interactive features allow for intuitive use without increasing cognitive load, interactive videos should therefore lead to more efficient forms of learning. This notion was tested in an experimental study, where participants learned to tie four nautical knots of different complexity by watching either non-interactive or interactive videos. The results show that in the interactive condition, participants used the interactive features like stopping, replaying, reversing or changing speed to adapt the pace of the video demonstration. This led to an uneven distribution of their attention and cognitive resources across the videos, which was more pronounced for the difficult knots. Consequently users of non-interactive video presentations, needed substantially more time than users of the interactive videos to acquire the necessary skills for tying the knots.

Over the course of the last two decades, the art of presenting information has undergone dramatic changes due to the development of computer systems that are capable of processing and displaying huge amounts of information almost instantly. Therefore, dynamic visual media like interactive video, hypervideo, and virtual reality have emerged. These media combine realistic, iconic depictions with the three main characteristics of the so-called "new media":

* Corresponding author. Tel.: +43-732-2468; fax: +43-732-2468-9315.
E-mail address: stephan.schwan@jku.at (S. Schwan).

0959-4752/$ - see front matter
doi:10.1016/j.learninstruc.2004.06.005

- non-linear structure,
- concerted use of a great number of different symbol systems, and
- interactivity, which gives users the opportunity to decide on the "what" and the "how" of the information presentation.

It is this last aspect of the new types of dynamic visual media—interactivity—that will be the focus of the following sections. In particular, it will be discussed if and under which circumstances the enlarged scope of action on the side of the user leads to a better processing and understanding of the presented visual information. From a cognitive standpoint, this problem may be decomposed into two different aspects. Firstly, the difference between picking up information by means of direct, unmediated observation and by means of media-based presentation will be considered. Secondly, against this backdrop of the cognitive purposes of media *in general* the question of the cognitive benefits of *interactive* visual media will be discussed. In particular, we will examine if and under which circumstances interactivity leads to a better processing and understanding of the presented visual information.

1. The cognitive benefits of media-based information over direct experience

From the perspective of human evolution, the human information processing system has for the most part developed without the use of media (Donald, 1991). Thus, the cognitive apparatus is well adapted to conditions of direct, unmediated experience. Nevertheless, a large number of empirical studies in the field of media psychology have shown that recipients can also deal with media based information with apparent ease. This is especially true for iconic depictions—regardless of whether they are in the form of static pictures and photographs or dynamic movies and video clips (Hobbs, Frost, Davies, & Stauffer, 1988; Messaris, 1994). This leads to the conclusion that at least realistic, iconic media make use of general, media-unspecific cognitive skills (Anderson, 1996; Levin & Simons, 2000).

Iconic media, however, should not be thought of as merely a convenient means of storing and transmitting visual information that is otherwise more or less cognitively equivalent to its "natural" unmediated counterparts. Instead, it should also be realized that one of the main benefits of media-based information presentations is that contents can be customized according to the cognitive needs of users. In the case of films or videos, customization takes place by staging events, by controlling the recording process (e.g. choosing different camera lenses, using multiple cameras etc.), and by arranging and assembling the recorded material during post-production (e.g. introducing film cuts). Indeed, some film scholars argue that by means of careful design, the film director is to a large extent, able to facilitate and control the cognitive processes of the film viewers (Bordwell, 1985). For example, Schwan, Garsoffky, and Hesse (2000) demonstrated that film depictions of complex activity sequences were better understood if activity boundaries were made more salient through the placement of film cuts. In other words, the placement of the cuts was found to facilitate the process of cognitively segmenting the stream of activity into comprehensible units. Also, Schwan and Garsoffky (2004) could show that film summaries of complex activity sequences led to mental representations similar to those developed from complete, 'un-shortened' depictions of the event, thereby reducing the viewers' needs to cognitively select and aggregate relevant parts of the activity.

In more general terms: media-based information presentations like photographs or movies are not merely valid reproductions of factual information, but are also instruments for information processing. They give authors and producers a great degree of freedom for shaping the presentation of information. These degrees of freedom may even be greater than for a common observer under conditions of natural, everyday experience. For example, a film director can record a given event simultaneously from multiple viewpoints, and can subsequently choose the best, "canonical" view for each part of the event (Blanz, Tarr, & Bülthoff, 1999; Palmer, Rosch, & Chase, 1981). In contrast, everyday observers are typically restricted to their particular viewpoint, which they cannot easily change (Garsoffky, Schwan, & Hesse, 2000). In other words, media authors have the possibility to optimize the experiencing conditions of events those ordinary observers do not possess.

Consequently, these characteristics of visual media lead to a kind of "working division" between author and recipient. The recipients delegate the shaping of their conditions of experience to the author. In turn, the author designs the information presentation in order to facilitate and partly anticipate the cognitive processes of the recipients. The recipient then follows the media presentation and processes it in a mostly predetermined way. This is exemplified by watching a movie or a television report: Here, the sequence and pace of the information presentation is fixed, thereby substantially constraining the mental activities of the viewers.

Keep in mind, that this model is not equivalent with the notion of the recipient as a kind of passive "couch potato". Instead, comprehending the media presentation encompasses a host of information processing activities (Wetzel, Radtke, & Stern, 1994), but these activities are shaped by the characteristics of the media-based information presentation. Additionally, they are restricted to internal, mental processes, whereas overt activities in the sense of modifying the given media presentation are minimized.

2. The benefits of interactive media over non-interactive media

According to the above arguments, one of the main advantages of media over direct experience lies in the possibilities to intentionally shape, arrange, and optimize information with regard to the cognitive apparatus of its recipients. This optimizing process does have limitations, however. Films and videos are mass media presentations in the sense that they do not address a specific viewer, but rather a general audience. Typically, the cognitive characteristics of this audience's members will vary to a greater or lesser extent. Examples of such differences between individual members might include cognitive skills, prior knowledge, current interests, or metacognitive strategies.

For traditional mass media, it is almost impossible to take these individual differences into account. Instead, they have to rely on the notion of an "average user", who is equipped with a set of common cognitive characteristics. Thus, whereas a media presentation may be well suited for the mental apparatus of the user in *general*, its fit with the cognitive needs of a given *individual* viewer may be substantially smaller.

Here, interactive media come into play, because they give up the strict working division between author and user in favor of a more balanced approach, whereby parts of the process of shaping the information presentation are returned to the user. Interactive media enable the user to adapt the presentation to her or his individual cognitive needs by actively deciding about the "what" and the "how" of a given presentation. As a consequence, by introducing interactivity into a dynamic visual presentation, not only the balance between author and user is profoundly changed, but also the interplay of internal (mental) and external (media directed) activities on the side of the user.

Kirsh and Maglio (1994) have termed this user-triggered modification of an external information presentation an "epistemic action", which may serve important cognitive purposes. In particular, by use of epistemic actions, mental processes may be facilitated and simplified, either by reducing the number of elements to be held in memory, by reducing the number of required mental processing steps, or by making the whole process more reliable. Kirsh and Maglio (1994) exemplified this in a detailed analysis of the computer game "Tetris". Here, the player must organize variously shaped tiles as they drop from the top of the computer screen with increasing speed. Kirsh and Maglio (1994) were able to show that skilled players develop a number of epistemic strategies by which they manipulate the appearance of the tiles on the screen. Players use these strategies in order to make the relevant features of the tiles perceptually more salient. This facilitated their cognitive processing and allowed the players to cope with tiles presented at extremely high rates.

Viewers of traditional, non-interactive educational videos typically face problems similar to those of the Tetris players. They must rapidly organize information that is presented at a rate they cannot change. While texts allow their readers to tune their reading behaviour to their cognitive needs, this is not the case for traditional videos or films. Readers can account for the changing complexity of the text passages, e.g. by adapting their reading speed, by skipping unimportant passages, or by re-reading difficult passages (Bazerman, 1985; Guthrie & Mosenthal, 1987). In contrast, viewers follow the film according to its inherent speed of presentation without such adaptive opportunities.

As argued above, in the case of a thoroughly crafted film presentation, the strict working division between film editor and viewer may well be better suited to the cognitive outfit of an average viewer than a direct unmediated experience. Nevertheless, the lack of adaptability may lead to dangers such as shallow processing or cognitive overload, as has been found in a great number of empirical studies on learning by films or videos (Seels, Berry, Fullerton, & Horn, 1996; Wetzel et al., 1994). Thus, one important purpose of making such videos interactive may be to give the viewers the opportunity to adapt a presentation's pace and sequence to their own cognitive needs and skills.

In contrast with the abovementioned benefits, research on interactive multimedia shows that making use of interactive features increases the number of activities which must be planned and held in memory, and the number of decisions required during the information presentation. Therefore, the benefits of facilitating and simplifying mental processes may be outweighed by the increased cognitive load required to efficiently manage these interactive features (e.g. Conklin, 1987; Schnotz, Boeckheler, & Grzondziel, 1999). To put it in terms of cognitive load theory: interactive features introduces the possible danger of increasing the extraneous cognitive load instead of freeing cognitive resources for further information elaboration (i.e. germane cognitive load, Sweller, Merrienboer, & Paas, 1998). Nevertheless, research on human–computer interaction indicates that interactive functionalities can be designed in a way that keeps increases in extraneous load at a low level. In particular, utilization of natural mapping devices that correspond to everyday perception-action-cycles have been shown to minimize required cognitive effort (Norman, 1988; Zhang & Norman, 1995).

3. Experimental overview

In summary, an advantage of interactive dynamic visualizations over their non-interactive counterparts may be expected in cases where the interactivity allows users to substantially reduce the cognitive processing requirements of the learning task while keeping the cognitive costs of the interaction behaviour at a low level. The following study attempts to show that under these circumstances, learning can indeed profit from the provision of interactive features.

For purposes of empirical investigation, we choose the tying of nautical knots as to-be-learned content. There were several reasons we expected this task to be especially suitable for an interactive video. Firstly, the tying of a nautical knot may be considered as a continuous motor skill, which is typically learned by a combination of observation and practice. According to Park and Hopkins (1993), such skills are best demonstrated through dynamic visualizations, e.g. video clips or animations, instead of static pictures or text descriptions. Secondly, from a practical standpoint, in the case of nautical knots videos may both outperform direct demonstration and interactive animations—the former because of its convenience and its unlimited repeatability, the latter because of its relatively cheap and easy production. Thirdly, different steps of the overall tying procedure may vary greatly in difficulty and complexity (McLeay & Piggins, 1998). Also, due to varying prior knowledge and competencies, a great degree of individual differences during learning may be expected. Therefore, nautical knot tying qualifies as learning content that should benefit from interactively adapting its dynamic visualization to variations in complexity and cognitive processing requirements.

Whereas half of the learners were required to acquire the knot-tying skill by viewing non-interactive videos, the other half could manipulate time-related parameters of the video presentation such as speed and direction. In order to minimize the cognitive effort for handling these interactive functions, an interface similar to Apple Computer's QuickTime was chosen. Here, speed and direction of presentation can by simply manipulated by moving a tile in an analogous manner. Therefore, this kind of interface fulfils the requirements of a natural mapping device, as specified by Norman (1988) or Zhang and Norman (1995).

Within this learning setting, we expected that

- compared to conditions of a non-interactive video demonstration, participants in the interactive video condition should indeed make use of its interactive functionalities, that is, manipulate both speed and direction of the presentation.
- by use of such features, participants should devote different amounts of cognitive effort to different parts of the demonstration (as indicated by different inspection times).
- the uneven distribution of effort should only partly be determined by objective features of the video demonstration, but should also reflect the individual differences of the participants.
- consequently, the uneven distribution of cognitive effort caused by the use of the interactive features should allow the participants to acquire the required motor skill in a more efficient way.
- overall, these effects of interactivity should be more pronounced for complex than for simple knots.

4. Method

4.1. Subjects

Thirty-six participants (eight female and 24 male students at the University of Offenburg, mean age 23) took part in the study. None of them had prior experience with the tying of nautical knots. They were paid for their participation.

4.2. Stimulus material and experimental setting

From a commercially available CD-ROM entitled "Segeln lernen interaktiv" (Learning to sail interactively)[1], four video clips were selected. Each video clip showed—framed in close-up—the hands of an actor demonstrating the tying of a specific nautical knot (see Fig. 1). Each video clip was purely visual, i.e. contained no spoken commentary nor sound and lasted between 14 and 35 s. Two experts rated the knots according to their difficulty. In ascending order of difficulty, the knots were double half hitch, cleat wind, anchor bend, and bowline.

Based on the video clips, two different learning environments were realized. In the non-interactive environment, the video clips were presented in an all-or-nothing fashion. That is, the learner was free to view the respective clip as often as she wished, but she had to view it always completely from beginning to end at normal

[1] Segeln lernen interaktiv (1994), Delius Klasing Verlag Bielefeld

Fig. 1. Video still from one of the experimental videos.

speed without the possibility of interruption. In contrast, in the interactive environment, the learner could interrupt the clip at will, change the speed of presentation from slow motion to normal speed to time-lapse and vice versa, and could also change the direction of presentation from forwards to backwards and vice versa.

4.3. Experimental design and procedure

The experiment was run as a 2×4 factorial design with "video type" (non-interactive vs. interactive) as between-subjects factor and "difficulty of knot" (double half hitch vs. cleat wind vs. anchor bend vs. bowline) as within-subjects factor. The experiment was run in individual sessions, with each session lasting between 80 and 120 min. After a short welcome, the participants were seated in front of a computer and were given a rope. They were told that they had to learn to tie four different nautical knots by watching video clips and practicing the knots. Prior to the four experimental knots, a simple training knot (figure eight knot) had to be learned in order to familiarize the participant with the learning environment. For all five knots, the procedure was identical. In the learning phase, the learner could watch the demonstration video as often as she wished. She could also practice the knot by tying the rope, but only after the video was stopped. In other words, during the learning phase, the learner could alternate between watching the video and practicing the knot at will, but was not allowed to watch and practice simultaneously.

In the non-interactive condition, the video had to be watched completely from beginning to end. In the interactive condition, the video could be stopped at arbitrary points and the speed and direction of the video presentation could be varied. The participants were instructed to proceed through the learning phase until they were able to tie the nautical knot correctly. After finishing the learning phase, the tying of the knot had to be demonstrated to the experimenter. Then, the next knot had to be learned. The knots were presented in fixed order of ascending difficulty (double half hitch, cleat wind, anchor bend, bowline). After having learned all five knots, the learners were interviewed about their learning strategies, and then completed a test on spatial ability (Schlauchfigurentest; Stumpf & Fay, 1983). Also, their experience with the tying of knots, the usage of computers in general, and the usage of interactive learning software in particular was assessed on four-point scales.

4.4. Data analysis

The behaviour of each learner during the learning phase was recorded automatically via log-file protocols. Based on the log-files, the following behavioural indices were determined: Overall duration of the learning phase, which consisted of the overall duration of watching the video during the learning phase, and the overall duration of practicing the knot during the learning phase. To assure comparability across the four different knots, each time measure was divided by the length of the respective demonstrating video. Also, the number of practicing trials during the learning phase was determined.

Additionally, for the interactive condition, it was determined how often the learner stopped the video clip without practising, how often she watched the video in slow motion, how often she watched the video in time-lapse, and how often the direction of presentation was changed. Also, for each frame of the video clip it was determined how long the learner viewed it. From this, both mean and standard deviation of the viewing duration across all frames of the video clip were calculated for each learner.

5. Results

5.1. Individual differences

The participants of the study had no prior expertise in the tying of knots ($x = 0.2$ on a four-point scale), an average experience with computers in general ($x = 2.2$ on a four-point scale), and low expertise with interactive learning software ($x = 1.0$ on a four-point scale). Also, t-tests showed no significant differences between the experimental conditions on these variables, all $p > 0.10$. Similarly, the spatial abilities were also comparable across experimental conditions, $t(34) = 0.29$, $p > 0.10$.

5.2. Learning outcomes

None of participants had to be excluded from the study, because all were able to competently demonstrate the tying of each of the four knots to the experimenter immediately after the respective learning phase.

5.3. Learning time

As can be seen in Table 1, across all experimental conditions, the mean overall learning time, relative to length of the respective demonstrating video, varied from 7.6 for the easiest knot (double half hitch) under conditions of interactive presentation to 34.2 for the most difficult knot (bowline) under conditions of non-

Table 1
Means (M) and standard deviations (SD) of overall learning time, viewing time and practicing time in the different experimental conditions

			Double half hitch	Cleat wind	Anchor bend	Bowline
Overall learning time	Non-interactive	M	13.5	31.6	25.7	34.3
		SD	5.7	13.9	11.9	13.6
	Interactive	M	7.6	19.0	13.0	20.5
		SD	4.4	12.0	6.8	9.1
Viewing time	Non-interactive	M	6.6	10.5	10.1	12.3
		SD	2.5	4.2	4.0	4.2
	Interactive	M	4.1	9.7	7.7	9.9
		SD	1.5	4.7	3.3	3.9
Practicing time	Non-interactive	M	6.9	21.1	15.6	22.0
		SD	3.4	11.0	8.9	10.9
	Interactive	M	3.6	9.2	5.2	10.6
		SD	3.3	8.2	4.1	5.6

interactive presentation. In a 2×4 MANOVA, both factors "video type" (non-interactive vs. interactive), $F(1, 34) = 14.6$, $p < 0.01$, and "difficulty of knot" (double half hitch vs. cleat wind vs. anchor bend vs. bowline), $F(3, 102) = 46.8$, $p < 0.001$ were highly significant. The interaction between the two factors was also significant, $F(3, 102) = 2.7$, $p < 0.05$. Closer inspection of the results shows that it took the learners drastically longer to learn the knot under the conditions of non-interactive video presentation than under the conditions of interactive video presentation. Depending on the type of knot, this difference was more or less pronounced, ranging from about 66% for cleat wind to about 95% for anchor bend.

Next, the overall learning time was broken down into the time for viewing the demonstrating video and the time for practicing the knot (see Table 1). With regard to the viewing time, a 2×4 MANOVA revealed a highly significant main effect of "difficulty of knot", $F(3, 102) = 34.6$, $p < 0.001$. Although for every type of knot, the interactive group spend less time viewing the video, neither the factor "video type" ($F(1, 34) = 4.7$, $p > 0.05$) nor the interaction of "video type" and "difficulty of knot" ($F(3, 102) < 1$, $p > 0.10$) reached significance. Thus, whereas the viewing times varied depending on the type of knot, they were roughly comparable with regard to the presence or absence of interactive possibilities. In contrast, the practicing times were heavily influenced by the factor "video type". For all knots, the learners practiced at least as twice as long in the non-interactive condition than in the interactive condition. The Anchor Bend was practiced three times longer under non-interactive condition than under interactive condition. Accordingly, a 2×4 MANOVA revealed significant main effects both for "video type", $F(1, 34) = 18.0$, $p < 0.001$ and for "difficulty of knot", $F(3, 102) = 38.7$, $p < 0.001$, as well as an highly significant interaction of both factors, $F(3, 102) = 6.1$, $p < 0.01$.

5.4. Usage of interactive features

In the interactive condition, the learners made heavy use of the interactive features, especially of slow motion, time-lapse (i.e. playing the video with increased presentation speed), and change of direction (see Table 2). The amount of interactivity was also influenced by the difficulty of the to-be-learned knot, with more difficult knots provoking more interactive behaviour. Accordingly, one-factorial analyses of variance showed significant effects of knot difficulty for the slow motion feature, $F(3,51) = 11.6$, $p < 0.001$, for the time-lapse feature, $F(3,51) = 4.0$, $p < 0.05$, and for changes of presentation direction, $F(3,51) = 8.4$, $p < 0.001$. The strategic uses of the interactive features were also apparent in the post-experimental interviews with the participants. Most of them said that they used the interactive features in order to cope with the more difficult parts of the knot tying process.

By means of this usage pattern, the learners distributed their viewing time devoted to the individual video frames unevenly. For each of the four videos, the mean viewing time of each of its frames was determined across all viewers. Based on these data, the mean standard deviation of the viewing times across all frames was calculated for each video. Again, this index of variability of viewing times was more pronounced for the more difficult knots, as shown by a highly significant effect, $F(3,51) = 13.6$, $p < 0.001$ (see Table 2).

In addition, an indicator of the homogeneity of the usage of interactive features was calculated. For each of the four videos, and for each participant, the distribution of viewing time across its frames was calculated. Then, the mean correlation of these distributions across the participants was calculated for each video. The mean correlations were $r = 0.19$, $p < 0.05$ for the double half hitch, $r = 0.22$, $p < 0.05$ for the cleat wind, $r = 0.17$, $p < 0.05$ for the anchor bend, and $r = 0.48$, $p < 0.001$ for the Bowline. Thus, for three of the four knots the mean correlations were positive but low, indicating strong inter-individual differences in the distribution of viewing times across the videos. The only exception was the most difficult

Table 2
Means (M) and standard deviations (SD) of use of interactive features for the four different knots

		Double half hitch	Cleat wind	Anchor bend	Bowline
Number of stops (without	M	3.3	2.7	3.5	2.3
practicing)	SD	3.7	1.5	1.5	1.6
Number of slow motions	M	1.5	7.7	6.3	10.1
	SD	2.0	7.0	5.5	8.4
Number of time-lapses	M	14.8	17.1	18.7	23.3
	SD	12.9	14.4	13.8	11.0
Number of direction changes	M	16.0	22.9	24.7	31.7
	SD	13.1	17.6	17.8	15.1
Variability of viewing time	M	2.1	4.0	3.4	5.3
across video	SD	1.2	2.2	1.9	2.2

knot, the bowline. Here, a moderate inter-individual accordance of the viewing time distribution was found.

6. Discussion

The present study addressed the question whether or not dynamic visual media, conceived as cognitive tools, may gain additional effect by making their process of reception interactive. The results show that at least under the specific conditions of the present study, the provision of dynamic visualisations with interactive features may indeed accelerate the process of skill acquisition. When faced with the task of learning nautical knots of varying difficulties, participants made heavy use of the interactive features provided. Similar to well-established reading strategies (Bazerman, 1985; Guthrie & Mosenthal, 1987), viewers of the video clips accelerated, decelerated, stopped, reversed or repeated parts of the videos in a systematic manner, thus distributing their attention and cognitive resources unevenly across the whole video demonstration. Also, this redistribution was even more pronounced for the more difficult knots.

The results fit neatly with the concept of epistemic actions, as introduced by Kirsh and Maglio (1994): the participants obviously possessed a repertoire of strategic interactions, and used it intentionally to adapt the pace and density of the visual information presentation to their cognitive processing needs. Further, these processing needs were to a large degree of an *individual* nature, as indicated by the low correlations of the distribution of viewing times across participants for a given video. Hence, these adaptation processes can only partly be anticipated by a fixed, non-interactive approach to video design.

In effect, despite the substantial redistribution of attention, the average overall viewing times for a given interactive video were largely comparable to its non-interactive counterpart. In other words, the effect of requiring more time to watch the more difficult parts of the video repeatedly, in slow motion, or with stops was more or less compensated by skipping the more easy parts of the videos or watching them in speeded mode. But although the viewing times were roughly the same, the viewers of the interactive videos developed a better understanding of the depicted processes than the viewers of the non-interactive videos. Accordingly, the viewers of the non-interactive videos had to practice the knots at least twice as long as the viewers of their interactive counterparts until they could reproduce it accurately. In other words, learning to tie nautical knots proceeded in a far more efficient manner, if the video demonstration allowed for interactive adaptation to the cognitive needs of the viewers.

Similar gains in learning effectiveness have also been reported for learning origami (paper-folding) by means of interactive video demonstrations (Shyu & Brown, 1995). On the other hand, some other empirical studies of interactive dynamic visualisations have failed to show corresponding increases in learning outcomes (Lowe, 1999; Schnotz et al., 1999). At closer inspection, two major differences seem to contribute to this difference in usefulness of interactive possibilities. Firstly, the studies of Lowe (1999) and of Schnotz et al. (1999) dealt with the acquisition of abstract conceptual knowledge—e.g. meteorological dynamics from weather maps—whereas both the present study and the study by Shyu and Brown (1995) focussed on concrete procedural skills. As Lowe (1999) has convincingly shown, in the case of weather maps, the visible dynamic changes in the video presentation do not necessarily stand in direct connection to the underlying causal principles that have to be learned. In other words, viewers may not possess the appropriate interaction strategies to deal with the video presentation in order to extract its essential conceptual features. Secondly, this lack of appropriate strategies may be accompanied by the problem of finding an adequate user metaphor for the interactive functionalities. The absence of such a metaphor reduces the naturalness of the user interface and may in turn lead to an increased extraneous load, as reported by Schnotz et al. (1999). In other words, balancing the provision of interactivity with the cognitive effort required for strategically handling these functions seems to be a difficult endeavour for abstract conceptual tasks. In contrast, combining a procedural learning task with a highly intuitive interactive interface constitutes an valuable application field for interactive dynamic visualisations, where learners may adapt the presentation to their individual needs at low cognitive costs.

Acknowledgements

Thanks to Carmen Zahn and Jahn Hirn for helping conducting the study and analysing the data.

References

Anderson, J. D. (1996). *The reality of illusion. An ecological approach to cognitive film theory*. Carbondale: Southern Illinois University Press.

Bazerman, C. (1985). Physicists reading physics. *Written Communication*, *2*(1), 3–23.

Blanz, V., Tarr, M. J., & Bülthoff, H. H. (1999). What objects attributes determine canonical views? *Perception*, *28*, 575–599.

Bordwell, D. (1985). *Narration in the fiction film*. London: Routledge.

Conklin, J. (1987). Hypertext: an introduction and survey. *IEEE Computer Magazine*, *20*(9), 17–41.

Donald, M. (1991). *Origins of the modern mind*. Cambridge, MA: Harvard University Press.

Garsoffky, B., Schwan, S., & Hesse, F. W. (2002). The viewpoint dependency of recognizing dynamic scenes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(6), 1035–1050.

Guthrie, J., & Mosenthal, P. (1987). Literacy as multidimensional: locating information and reading comprehension. *Educational Psychologist*, *22*(3&4), 279–297.

Hobbs, R., Frost, R., Davis, A., & Stauffer, J. (1988). How first-time viewers comprehend editing conventions. *Journal of Communication*, *38*(4), 50–60.

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, *18*, 513–549.

Levin, D. T., & Simons, D. J. (2000). Perceiving stability in a changing world: combining shots and integrating views in motion pictures and the real world. *Media Psychology*, *2*, 357–380.

Lowe, R. (1999). Extracting information from an animation during complex visual learning. *European Journal of Psychology of Education*, *6*(2), 225–244.

McLeay, H., & Piggins, D. (1998). Mental unknotting. *Perceptual and Motor Skills*, *86*, 707–719.

Messaris, P. (1994). *Visual literacy. Image, mind, & reality*. Boulder: Westview Press.
Norman, D. (1988). *The psychology of everyday things*. New York: Basic Books.
Palmer, S., Rosch, E., & Chase, P. (1981). Canonical perspective and the perception of objects. In J. Long, & A. Baddeley (Eds.), *Attention and performance IX* (pp. 135–151). Hillsdale, NJ: Erlbaum.
Park, O. C., & Hopkins, R. (1993). Instructional conditions for using dynamic visual displays: a review. *Instructional Science*, *21*, 427–449.
Schnotz, W., Boeckheler, J., & Grzondziel, H. (1999). Individual and co-operative learning with interactive animated pictures. *European Journal of Psychology of Education*, *14*(2), 245–265.
Schwan, S., Garsoffky, B., & Hesse, F. W. (2000). Do film cuts facilitate the perceptual and cognitive organization of activity sequences? *Memory and Cognition*, *28*(2), 214–223.
Schwan, S., & Garsoffky, B. (2004). The cognitive representation of filmic event summaries *Applied Cognitive Psychology*, *18*(1), 37–55.
Seels, B., Berry, L. H., Fullerton, K., & Horn, L. J. (1996). Research on learning from television. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 299–377). New York: Macmillan.
Shyu, H., & Brown, S. W. (1995). Learner control: the effects on learning a procedural task during computer-based video instruction. *International Journal of Instructional Media*, *22*(3), 217–231.
Stumpf, H., & Fay, E. (1983). *Schlauchfiguren *Ein Test zur Beurteilung des raeumlichen Vorstellungsvermoegens*. Goettingen: Hogrefe.
Sweller, J., Van Merrienboer, J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*(3), 251–296.
Wetzel, C. D., Radtke, P. H., & Stern, H. (1994). *Instructional effectiveness of video media*. Hillsdale, NJ: Erlbaum.
Zhang, J., & Norman, D. (1995). A representational analysis of numeration systems. *Cognition*, *57*, 271–295.