Evaluating a framework of theoretical hypotheses for animation learning

S. GUTTORMSEN SCHÄR, H.-J. ZUBERBÜHLER

This paper presents a set of theoretical hypotheses suggesting various relationships between didactical setting and learning effects with animations. Particularly, we investigated whether individual flow-control adequately provides didactical means to reduce the cognitive load imposed by animations. We did not find an effect of individual flow control, probably due to the fact that this learning condition was embedded in a setting where not enough verbal information was offered together with the graphical animation. Overall the multimedia effects found in this study are in line with known principles of didactical multimedia design. Further, this study sheds light on theoretical aspects involved in the complex interaction between learning content, presentation, learning and resulting knowledge.

Keywords: **animations; didactical setting; multimedia effects; theoretical hypotheses**

Untersuchung mehrerer theoretischer Hypothesen zum Lernen mit Animationen.

Dieser Beitrag präsentiert eine Reihe theoretischer Hypothesen, welche auf verschiedene Beziehungen zwischen didaktischer Umgebung, Animation und Lerneffekten hinweisen. Im Besonderen haben die Autoren untersucht, ob eine individuelle Steuerung des Informationsflusses bei Animationen ein wirkungsvolles didaktisches Mittel ist, um die kognitive Überlastung zu reduzieren. Es wurden keine Auswirkungen der individuellen Informationssteuerung gefunden, wahrscheinlich weil diese Lernbedingung zu wenige verbale Informationen zu den Bildern zeigte. Insgesamt stehen die in dieser Studie gefundenen Multimedia-Effekte im Einklang mit schon bekannten Prinzipien des didaktischen Multimedia-Designs. Zusätzlich beleuchtet diese Studie die theoretischen Aspekte, die beteiligt sind an der komplexen Interaktion zwischen Lerninhalt, Präsentation, Lernen und dem daraus resultierenden Wissen.

Schlüsselwörter: **Animationen; didaktische Umgebung; Multimedia-Effekte; theoretische Hypothesen**

1. Introduction

This paper addresses questions related to learning effects of animations. The state of the art of didactical animation research shows that the current question occupying researchers in this field is not anymore *whether* animations support learning or not. Rather, it is more relevant to ask in which didactical setting animations should be embedded in, in order to ensure a learning benefit compared to static media, e.g. picture and text. Most of the media research has been performed with static media (e.g. *Large, 1996, Levin, Anglin, Carney, 1987, Peeck, 1993*). The research on dynamic media has a long tradition in not showing clear learning benefits compared to static media (e.g. *Bètrancourt, Tversky, 2000, Park, Hannafin, 1993, Park, Hokins, 1993*). Only few studies clearly show that dynamic media support learning (*Rieber, 1991 a, Mayer, 2001*). The dilemma of the actual animation research is the insight that animations expectedly should be optimal to represent dynamic information, but the experience after numerous studies is that it is difficult to clearly show that animations support learning better than static media. This imposes a more focused research approach, which aims at defining those factors that restrain learning when animations are provided, while keeping up the assumption that animations should have a positive impact when employed purposefully. We propose the following driving hypothesis in order to investigate cognitive load with animations: Multimedia presentations with animations support learning if they are embedded in a didactical context that a) represent the characteristics of the learning content and b) reduces the cognitive load imposed on learners who have to process a continuous flow of information.

Below, we have formulated the theoretical review related to the issues addressed in this paper as a series of hypotheses, these hypotheses reflect some issues driving the theoretical and empirical research in this field today. The hypotheses were evaluated in an experimental study, which purpose was to refine the theoretical hypotheses rather than classical hypothesis testing.

1.1 The cognitive load hypothesis

An assumption behind this study is that some media or media combinations have a detrimental effect on learning because they induce cognitive load. We have described various sources of cognitive load in more detail in an earlier paper (*Guttormsen Schär, Kaiser, in Print*): *Information redundancy* is caused by the simultaneous presentation of information with more than two media (cf. *Mayer, 2001*). *Split attention*, an effect of information redundancy, induces learners to share their attention between two concurring sources of information (*cf. Sweller, 1999*). Another source of cognitive load is *information channel overload*, which is explained theoretically in the Dual Coding Theory

GUTTORMSEN SCHÄR, Sissel, Prof. Dr., Medical Faculty of the University of Bern, Institute for Medical Education, Inselspital 37a, 3010 Bern, Schweiz (e-mail: sissel.guttormsen@iml.unibe.ch); ZUBERBÜHLER, Hansjörg, Dr., Multimedia- & E-Learning Services, University of Zürich, Winterthurerstrasse 190, 8057 Zürich, Schweiz (e-mail: hansjoerg.zuberbuehler@id.unizh.ch)

(DCT) (*Paivio, 1986, Mayer, 2001*). The DCT explains cognitive overload through excess usage of a single information channel (verbal or visual). This theory postulates that language based (voice and text) and pictorial representations are processed differently. Generally, the theory suggests that we have two sub systems for information processing, one system for the processing of non-verbal objects and events (imagery system) and one system specialized for the processing of language (verbal system). In line with the dual coding theory, *split attention* is related to overload on one of the information channels, verbally or visually. This would be the case when e.g. the visual channel must process two information sources simultaneously (e.g. visual text and picture). *Attention* is yet another factor that may be related to cognitive load, particularly when students must spend cognitive capacity to guide their attention to the relevant instead of irrelevant information (*Sweller, 1999*). The presentation of any cues (verbal or visual) that distracts the attention away from the relevant aspects of the learning content may harm learning. The content of a multimedia presentation should overlap with the semantics of the learning content. In case the content of the presentation is irrelevant to the learning target, the attention gaining effect can be distractive. Mayer describes distraction through the presentation of irrelevant. Irrelevant information in form of words, pictures, music or sound hurts learning as long as this kind of information is not in coherence with the learning task (*Mayer, 2001*).

Cognitive load is a more predominant issue with dynamic media than with static media. A general explanation for high cognitive load with animations is that students cannot process the information in the same speed as it is being presented. Rieber's research offers interesting specific suggestions for how to reduce cognitive load. Rieber found that animated presentations better supported learning when the information was presented in smaller logical information chunks (*Rieber, 1991a and 1991b*). Rieber designed the chunks as small information units and introduced a pause between each chunk. The students could initiate the presentation of the next chunk themselves. It is interesting to further investigate other means for reducing cognitive load caused from information overflow of dynamic presentations. Rieber suggested that the chunking of the information reduced the continuous information flow in a way enabling the students to better process the material. We have reported on Riebers experiments on the effects of animations on learning in more detail in Guttormsen Schär, (*in Print*). Our proposed solution to reduce cognitive load builds on Rieber's findings, but applies another approach to information chunking: We assume that multimedia presentations with animations would support learning if students are given full control over the information flow, i.e. that they can control the flow with a "stop and go" interface. This looks more like the current way of presenting animations in a didactical context, as individual control can be regarded as a standard. Hence, in this experiment we investigated whether individual information flow control has the same effect as information chunking as found by Rieber (*1991b*).

1.2 Static – dynamic hypothesis

The conceptual link between animation and learning content has been neglected in many studies. We assume that many studies failing to find the expected benefit of animation failed because the animated presentation was not congruent with the learning material (which may have directed the attention away from the important factors) or the animations were not necessary in order to learn (*Rieber, 1990b, Kinzer et al., 1989, Pane et al., 1996*). The formal analysis of learning content has received little attention in the multimedia literature. There is an inherent problem with this, because many studies suggest that there is a close relationship between characteristics of the learning content and the effect of animated presentations. The analysis of learning content has been given more attention in the computer supported instructional design domain (*Guttormsen Schär, in print*). The work of David Merrill and his taxonomy of learning content into different categories has been widely applied in various e-learning situations (*Merrill, 1983, 1996, 2000, 2001*). Merrill's work builds on the work of Gagné, who was among the first to define instructional events, related mental processes and correlating didactical means (*Gagné, 1965, 1985*).

Some few researchers in the multimedia research domain have addressed the learning content as an important factor for the success of media in a learning setting (e.g. *ISO 14915–3, 2002, Alty, 1993, Rieber, 1990a*). Their argument for learning content analysis holds that good multimedia design depends on descriptions of the nature of the learning content and deciding which medium has characteristics that best transmit such information. Applied to animations, we infer that positive effects of animated presentations depend on whether they reflect dynamic qualities of the learning content. There is a growing body of evidence suggesting that animations are useful to represent dynamic learning content: i.e. temporal characteristics, domain knowledge including movement and action. Furthermore, animations can serve as visual analogy or anchor for understanding abstract and symbolic concepts or processes and for demonstrating sequential actions in a procedural task. Implementation of animation in instructional design is likely to promote learning mainly in three situations: *visualisation*, *motion* and *trajectory* (i.e. the direction of a path of a travelling object), and *temporal aspects* (*Rieber 1990a, Rieber, 1989, Rieber, Kini, 1991, Large, 1996, Park, Hopkins, 1993*):

(1) When conveying a concept or rule over time, or a relation between time and space. When the learning material entails motion, trajectory or change over time, the animation helps to build a mental model of the dynamics (motion can carry relevant implicit information about the learning objective).
(2) For the explicit representation of highly abstract and dynamic concepts and processes in science. When conveying a dynamic process that is difficult for learners to imagine on their own. Animation can provide a convenient means of visually representing imaginable events, actions and ideas, which changes over time.
(3) Animations can be used to simulate functional behaviour of mechanical and electronic systems (especially with dynamic processes that cannot easily be visualized in real, e.g. hydraulics of a pump); to explicitly represent invisible flow of information or current in electronic systems; and as a substitute for real time verbal communication.

In contrary to dynamic learning content, we define static learning content as validated knowledge about objects, matter, or material, states or situations, or descriptions of a unique moment of time. Static learning content remains stable in a relative time span, for instance descriptions, e.g. of a painting, concepts e.g. the definition of "standard-deviation" in statistics is stable and does not change over time (*Guttormsen Schär, Messerli, Schluep, 2004, Merill, 1983, ISO 14915-3, 2002, CISCO, 1999*).

Coherence between learning content and media representation may also positively influence the cognitive load on the learner. Applying the Cognitive Load Theory, it is reasonable to assume that animations can help to reduce the complexity of cognitive processing of temporal ideas, because the step of translating static information into dynamic information must not be performed (*Sweller, 1999*). The cognitive load on the learner would expectedly be higher when dynamic information represented in a static manner would have to be mentally transformed into a dynamic flow of events. Hence, when the learning content is dynamic in a way that it contains information, which can be visually represented as changing states, activities or events, the assumption is that a corresponding visual dynamic representation supports the cognitive processing of the learning material.

Based on our reflections on the relationship between dynamic learning content and animations (dynamic presentation) we set up a naïve hypothesis suggesting that there is an optimal fit between media and learning content on the static/dynamic dimension. Static learning content should be presented with static media, and dynamic learning content (as described above) should be presented with dynamic media. Rieber has addressed the effect of media on learning performance in a way similar to our naïve hypothesis (*Rieber, 1990a, 1990b*). Rieber tested the effect of animated presentations on incidental and intentional learning. According to the results, incidental learning refers to that kind of information that implicitly is embedded in dynamic presentations (i.e. implicit dynamic information about the relationship between the dynamic learning content). The intentional learning refers to that knowledge the students were instructed to try to learn. Rieber (*1990a, 1990b*) found a larger difference between the animated graphics group and the static graphics group on the incidental learning questions, than there was on the intentional learning questions. Regarding incidental learning, in both studies the students of the animated conditions groups outperformed those students in the static conditions groups in incidental learning. Hence, Rieber's study showed that there is a relationship between media and the resulting knowledge category, and that dynamic knowledge contents can be better learned with dynamic presentations. We aimed at further investigating those findings as formulated in the naïve hypothesis above.

1.3 Knowledge modality hypothesis

As suggested in the static dynamic hypothesis, there is a relationship between the modality of the presented media and the modality of the acquired knowledge. A former study showed different media (-combinations) induced different kinds of knowledge (*Guttormsen Schär, Kaiser, in print*). We found that the extent to which certain media combinations support learning is a mater of how well the media (combinations) fit with the tested knowledge. Hence, we wanted to continue to investigate the degree to which various media generally influence the acquired knowledge or whether various media invoke particular effects on knowledge. Theories deliver support for the assumption that media do not influence knowledge acquisition generally but specifically, e.g. the Dual Coding Theory (*Paivio, 1986, Mayer, 2001*), the Multimodal Memory Theory (*Engelkamp, 1991*) and the Cognitive Load Theory from Sweller (1999). The theories propose that our experienced information is *contained* in representational structures and processes. Hence, our ability to generate, manipulate and transform images of objects and events depends largely on perceptual-motor experiences with such objects. Hence, our representational structures are modality specific rather than general. This is particularly emphasised by Engelkamp, who describes a direct link between sensory experiences and their corresponding mental representations (*Engelkamp, 1991*).

This issue of coherence between media and learning content in terms of static and dynamic qualities is, however, more complicated because complex dynamic learning content is likely to contain both static and dynamic aspects and, hence, would then demand both dynamic and static presentations. A more precise knowledge modality hypothesis, therefore, holds that animated presentations are beneficial for the acquisition of the dynamic knowledge aspects of a given learning task, while static media would better support the acquisition of static knowledge aspects. Consequently, the learning effects of animations should be related to whether the test of knowledge addresses static or dynamic content characteristics. In some studies with animations a difference between dynamic knowledge (involving procedures or processes) and declarative knowledge (relating to facts) has in fact been measured. In Rieber's studies a distinction between fact knowledge (static) and rule learning (dynamic) was made (e.g. *Rieber, 1989, Rieber et al., 1990*).

1.4 Process hypothesis

Effects of animated presentations have often been tested with algorithmic learning material, i.e. dynamic learning content that demands a high degree of insight learning (inference learning) (e.g. *Rieber, Hanafin, 1988, Rieber, 1989, Rieber, Boyce, Assad, 1990, Rieber 1991a, Rieber, 1991b, Sweller, 1999, Mayer, 2001*). Such learning material requires the visualisation of a concept or rule over time. Processes are also dynamic, but in another way than algorithmic information. The process definition from Cisco is general accepted (*Cisco, 1999*). Accordingly processes are defined as: A description of a flow of events, referring to how something functions, may it be of scientific, organisational or mechanical characteristics. Further, we conjectured that a process is different from a procedure, which describes how a certain task is being executed. Rather a process is a (complex) flow of interdependent causalities. We aimed at testing such processes from various topics. One question is, whether the effect of media is relatively stable for those learning contents using this definition, or whether an eventual effect of animated presentations depends on the topic dependent. We hypothesised that the applied definition of processes is a) generic enough to make possible a selection of various processes representing dynamic learning content possible and b) that various such processes, independent of different topics, interact with various media in the same way.

2. Method

In this experiment we investigated the effect of various static and dynamic media combinations for the presentation of process information. Based on former research we did not include mono media presentations that clearly have been found to be ineffective in former studies (i.e. "only text" "only voice", or "only picture") (e.g. *Mayer, 2001, Guttormsen Schär, Kaiser, in Print*).

2.1 Design

A $5 \times 5 \times (3)$ within group design was applied, with the factors media, learning topic and knowledge category. *The factor media* was implemented with five media / multimedia conditions: In the animation (A) condition the learning topics were presented as pure animations, with legends but without text or voice. In the animation + voice condition (AV), the learning topics were presented as animations with narration. In the animation + control condition (AC) the presentation equalled the animation condition (A) with the additional option to start, stop and browse the animation with a Quick-Time™ control panel. The picture + voice conditions (PV) presented the learning topics as static pictures with narration (visual text was not presented simultaneously). In the condition picture + text (PT) the processes were presented as still pictures with text. This condition was added as a control. It is well established that presentations with PV support learning better than PT, hence similar results in our study would preserve the relation to established research (*Guttormsen Schär, in Print*).

The factor learning-topic was implemented with five different processes. The processes were selected from different topics, being: Heart functions / ECG; The protein synthesis; The Malaria cycle; Off-side rules in soccer; and The Solar eclipse.

The factor knowledge category was implemented with 3 different knowledge categories, which at the same time also were dependent measures: visual knowledge, verbal facts knowledge and process knowledge.

2.2 Subjects

The participants were 27 men and women at the age of 18 to 39 years. They were all recruited from a school for adults preparing for a university entry degree. The subjects were paid a small allowance for the participation. Pre-experimental knowledge in the tested process topics was tested previous to the experiment among the subjects.

2.3 Experimental tasks and measures

Table 1 gives an overview over the experimental variables. The within group character of the study implied that all subjects were exposed to all learning tasks and test conditions.

The general *learning task* was to learn each of five presented processes. We employed the Cisco Systems definition of processes as a guideline for the selection of experimental learning tasks in this study, as explained above.

The *pre-experimental knowledge* in the topics was tested with a simple and a difficult multiple-choice question. The simple question was expectedly answerable for non-experts, the difficult question addressed an expert degree of knowledge.

We checked the *subjective preferences* for the media and learning topics after the learning phase. In the latter case a screen was shown with the text *"Please rate your preference for the processes."* A picture of each topic was shown on the screen displaying a five-point scale with the items *"very bad, bad, neutral, good and very good"*. In order to determine the media preference a screen was shown with the text *"Please rate your preference for the media."* The media conditions were labelled as *"Animation, Animation with Voice, Animation with Control, Picture with Voice, Picture with Text"*, together with the same five point scale as described above.

Table 1. Overview over the variables

Variables		
Independents		**Dependents**
Media:	**Topic**	Visual knowledge*
Animation (A)	Heart functions;	Verbal facts knowledge*
Animation + Voice (AV)	Protein synthesis;	Process knowledge*
Animation + Control (AC)	Malaria cycle;	Preference: media
Picture + Voice (PC)	Off-side rules;	Preference: Topic
Picture + Text (PT)	Solar eclipse.	also independents
Control variables		
Age		
Gender		
Education		
Computer experience		
Presentation order (Task /media)		
Pre-experimental topic knowledge		

The operational definitions of the *knowledge categories* are explained below. There was one test question per knowledge category and topic. We defined *visual knowledge* as the ability to recognise the only correct picture from a presentation of 6 similar pictures. The pictures were presented on the screen and only one could be selected (with radio buttons). The visual test required the exact recognition of features from the task scene that were not obvious without knowledge from the visual presentations. We scored the responses as correct or false.

We defined *verbal facts knowledge* as the ability to reproduce key concepts from the learning tasks and to put them in the correct context. The subjects were given text fragments with blank sections, which they should complete. An example is the test-question for the malaria cycle: *"The* -------- ⟨sporozoites⟩ from the mosquito salivary gland are injected into the human blood stream. Once in the human bloodstream, they arrive in the liver where they multiply. On release, they return to the blood and penetrate red blood cells in which they produce ------- ⟨merozoites⟩." All the tests in this category prompted the subjects to fill in two missing words. A correct word at the correct place made 3 points, a correct word with wrong spelling at the correct place made 2 points, a correct word but at the wrong place made 1 point. The resulting scores thus ranged from 0 to 6 points.

We defined *process knowledge as* the ability to select correct process related statements among a selection of three correct and three wrong statements. The subjects were asked to choose a set of correct statements, and were not told that always three of the six statements were correct. Correct selections were given $+1$ point and wrong selections -1 point, hence the range of scores were between -3 and $+3$ points. An example shows a test from the topic "protein synthesis", which was implemented as a machine analogy:

"Which of the following statements are correct?"

- *Aspecial protein is defined by the order of its elements*
- *Aprotein moves through the cell membrane*
- *The green parts of a cell membrane are movable*
- *The brown parts of the cell membrane are movable*
- *A stretched spring is pressed into the ribosome*
- *The amino acids wander through the cell membrane*

2.4 Production

The five learning topics (processes) were programmed with Macromedia Director and were implemented as Quick time movies in a parent Director movie, which was programmed with the whole experimental procedure. The whole experimental setting was programmed in a parent-movie with 25 child movies (5 learning topics $\times 5$ presentation form). Log-file data with subject data, as well as all the answers to the knowledge questions, subjective evaluations and third variables were recorded in the parent movie. The parent movie also controlled the random procedure for the combinations of media and learning topic, and it controlled the pre-set learning time per learning topic. The learning time varied insignificantly between the processes, and reflected the time it took to run the AV version. Hence, all presentation forms for each process were set to the same duration ($=$AV time). The average learning time per presentation was 10 minutes. Each of the conditions A, AV, PV and PT was played twice (i.e. 2×5 minutes), the condition AC was played once and during the last five minutes, the subjects could control the activities in the last five minutes themselves. A digital clock showed the remaining learning time for each condition. The subjects could not leave the presentation, e.g. start the learning test, before the computer controlled learning time was finished.

The presentation form categories were produced according to the following procedure: A story (content description) was written to each of the process topics. Together with pictorial information it provided the input to a design team formed by design students who specialised in new media design. The design students were instructed to design a short QuickTime animation to each story. We continued to edit the raw animations until they reached our quality requirements: i.e. the presentations should offer a salient view of all the information required for the learning goals. For the topic solar eclipse e.g., this would be a representation of the information enabling naïve learners *to exactly reproduce the setting of the moon and the sun in relationship to their rotation axes'* The final animations were produced from colour drawings with a degree of detail reduced to support the focus on the predefined key factors. All other irrelevant information was not presented. The animations were produced with 20 frames/s. Narration to each topic was recorded with the voice synchronised to the animations. In the AV condition the narration thus followed the flow of the pictorial information. We produced the versions PV and PT from the five animations by capturing four key events as screen shots. In the PV condition the recorded narration was cut to fit the four pictures. In the PT version the corresponding identical information was shown as text on the right side of the picture. Both the conditions A and AC were without text or narration, displaying only conceptual legends.

2.5 Experimental procedure

The experimental situation was fully automated. All tasks were solved using a computer. The data was automatically registered in a log-file. The average total time per subject was 1 hour. The different media conditions were randomly combined with the topics. We made sure that each of the 25 possible combinations of process and media appeared approximately equally in the total subject group. Each subject experienced a random combination of all the topics and media conditions.

All the versions except for the AC version were shown twice. The latter version was shown once, and the subjects could use the QuickTime control-panel to select which of the sequences to view for the rest of the time.

Before learning the experimental tasks, we asked the subjects to answer multiple-choice questions related to the tested topics and randomly mixed with questions referring to processes not addressed in the experiment. The mix of relevant and irrelevant questions should prevent the subjects from focusing the attention to the learning material in front of the experiment. These data served as a reference and would enable us to exclude subjects who obviously had sufficient topic knowledge prior to the experiment.

The experiment proceeded as follows: First we introduced the subjects to the experiment and informed them in general that we were testing the effect of various media on learning. Then they sat down in front of the experimental computer and were asked to type in their personal data into the corresponding fields on the screen. After this, the pre-experimental topic knowledge was tested. The learning session followed in which, the subjects were instructed to learn the information presented to the different topics as good as possible. The subjects initiated the first task by clicking on a start button on the screen. When the learning time was over, a mask was presented for 2 s., in order to delete eventual images on the retina. The knowledge tests to the current topic followed directly after the learning phase. The following four learning tasks proceeded in the same manner. The subjects decided themselves when to continue with the next task. The experiment ended with the preference tests for topic and media presentations.

3. Results

3.1 Main effects of topic

Significant effects of topics were found for visual knowledge and preference. Table 2, and Table 3 show the pairs of compared processes resulting in a significant knowledge differences for those knowledge parameters. Table 4 shows the mean values for the different processes.

Table 2. Effects of topic: visual knowledge

Processes		Processes	Z	p
Heart-functions	>	Protein synthesis	2.5	0.0125
Solar eclipse	>	Protein synthesis	3.1	0.0017
Malaria cycles	>	Protein synthesis	3.7	0.0002
Malaria cycles	>	Offside rules	2.6	0.0108

Table 3. Effects of topic: preference (Fischer's PLSD)

Topics		Topics	Mean diff.	Critical diff.	p
Offside rules	>	Protein synth.	0.963	0.717	0.0089
Malaria cycles	>	Protein synth.	1.11	0.717	0.0027
Heart-functions	>	Protein synth.	1.37	0.717	0.0002
Heart-functions	>	Solar eclipse	0.741	0.717	0.0430

Table 4. Effects of topic: mean values of all the dependent variables

Processes	Preference	Visual knowledge	Process knowledge	Facts knowledge
Offside rules, soccer	2.4	0.48	1.9	3.3
Heart-rhythm / ECG	2.9	0.59	1.5	3.5
Malaria cycles	2.6	0.85	1.9	2.9
Protein synthesis	1.5	0.19	1.4	4.3
Solar eclipse	2.1	0.78	1.3	3.5

3.2 Main effects of knowledge category

We also found a main effect of knowledge: Process knowledge generally showed higher scores than both facts knowledge and visual knowledge. The data used to compare facts knowledge and process knowledge was parametrically mapped between 0 and 1. The results revealed that process knowledge was significant higher than facts knowledge (mean = 0.77 and 0.57 respectively; $p < 0.000$, $t = -6.28$). In order to compare visual knowledge (dichotomic data) with the static and process knowledge measures (parametric), the scores for facts and process knowledge were dichotomised. The analysis with the transformed data showed that process knowledge also was better than visual knowledge (mean = 0.90 and = 0.68 respectively, $p < 0.000$, $Z = -6.09$)

3.3 Main effects of media

We calculated the effect of media for each of the three knowledge categories. Table 5 shows the mean values for the dependent variables. The effect of media on process knowledge was calculated with a z-test based on the difference between the correct and wrong answers (possible scores ranged between +3 and −3). Significant effects of media were found for process knowledge and for preference. Tables 6 and 7 show these effects.

Table 5. Effects of media: mean values of all the dependent variables

Media	Visual knowledge	Process knowledge	Facts knowledge	Preference
A	0.63	1.1	3.3	1.6
AC	0.63	1.1	2.8	2.5
AV	0.63	2.1	3.7	3.5
PV	0.52	2.1	4.0	3.0
PT	0.48	1.6	3.6	2.5

Table 6. Effects of media: process knowledge (Fischer's PLSD)

Media		Media	Mean diff.	Critical diff.	p
AV, PV	>	A, AC	0.963	0.607	0.0021

Table 7. Effects of media: preference (Fischer's PLSD)

Media		Media	Mean diff.	Critical diff.	p
AV	>	AC	1.00	0.597	0.0012
AV		A	1.93		< 0.0001
AV		PT	1.04		0.0008
PV		A	1.37		< 0.0001
AC		A	0.93		0.0027
PT		A	0.89		0.0039

3.4 Other effects

There were no interaction effects between media and learning topic. There were also no effects of age, gender, computer-experience, order of the topics or order of the media. A correlation was found between pre-experimental topic knowledge and total knowledge scores. From 15 possible pairs of pre- and post-experimental knowledge, 11 correlated positively ($p < 5\,\%$). On the other hand, a correlation between pre-experimental knowledge in each task and the related post experimental knowledge was not obvious. From 15 possible pairs of pre- and post-experimental knowledge, only 4 showed a positive correlation ($p < 5\,\%$).

4. Discussion

4.1 Main effects of topic

The main effects of topic show some statistical differences between the learning topics (i.e. the different processes). The variation between the topics, however, was mainly related to the topic "protein synthesis", and the results were related to the two measures visual knowledge and preference. The other topics were perceived equally when it comes to learning performance. There were also no effects of topic for the other knowledge categories (facts and visual knowledge). Hence, this result supports that four of the learning topics were regarded as similar, and that they also were implemented and presented with a comparable quality. Hence, only the protein synthesis resulted in lower learning scores. This process was implemented slightly different than the others. The protein synthesis was represented as a machine analogy. Instead of showing how the amino acids moved through a ribosome, the synthesis took place in a mechanical machine. This analogy obviously had a negative effect on visual knowledge. It added another layer of information processing. The students also consciously reported a negative experience for this topic; the machine analogy was the least preferred compared to the other topics. It seems likely that the students still tried to transform the information onto their knowledge about human cells, and this extra loop of processing the material withdrew their attention from visual details. It is therefore not surprising that the interference of the uncommon analogy hindered learning of visual aspects. In general this main effect resembles knowledge associated with multimedia learning formulated in Mayer's Coherence Principle (*Mayer, 2001*). In this case we did not add extra pictures or irrelevant words, but the analogy itself seems to be cognitive distracting. Sweller (*1999*) has called such effects extraneous cognitive load, leading to a cognitive split attention. It seems plausible to assume that the analogy worked like a redundant information source increasing the cognitive load, due to the students having to (or wanting to) compare the visual information with their mental image of cells performing the same process. Our results also show, that failing coherence between the visual presentation and the semantic content did not affect all kinds of learning, e.g. the process knowledge that particularly addresses the deeper understanding of the learning content. This shows that an instructional bias does not necessarily affect learning performance generally, but rather that specific reactions can result. We have described this effect above as the "knowledge modality hypothesis".

In general, the analysis of the learning topic supports the process hypothesis, in that the different processes tested, expect from the protein synthesis, interacted with the media in a similar manner. This also supports that the applied process definition is usable for a stable selection and categorisation of such learning content.

4.2 Effects of knowledge categories

Process knowledge scored higher compared to both facts and visual knowledge. There were no other differences between the other knowledge categories. We do not find it evident that this result was related to a complexity bias between the knowledge test categories. The scoring procedures for the three knowledge categories, as well as the fact that all the knowledge questions were carefully designed *not* to induce a complexity bias between the test categories, incline other interpretations of this result.

4.3 Effects of media

Due to the main effect of learning topic, we verified the validity of the effects of media by checking if there was an interaction effect between topic and media for performance. No interaction effect of media and topic was found, which shows that the main effects of media for the different knowledge categories can be regarded as media induced and were not a result of eventual systematic differences between the topics.

The effects of media were too unspecific to support the static-dynamic hypothesis or the knowledge modality hypothesis. Due to the fact that there were no interaction effects between knowledge category and media we did not find evident support for the assumption that there is a relationship between presentation form (media) and knowledge category in this study. Our results show that there was no effect for facts or visual knowledge, but only for process knowledge, which was acquired equally well with the multimedia combinations animation + voice and picture + voice. Also for process knowledge, the presentation forms animation and animation + control resulted in significant poorer learning.

We did not find a unique animation effect in this experiment. The picture and voice combinations (regardless whether the picture was presented as an animation or as a still picture) supported learning of process information equally well compared to the only animated presentations. As a result, our assumption that an animation must be embedded in a special didactical setting to effectively support learning has been strengthened. It needs more than a dynamic presentation to support the learning of dynamic learning content. Further, our result is in line with the multimedia principle and the modality principle as formulated by Mayer (*2001*). The former principle postulates a learning benefit when words and pictures are presented together rather than just words alone. The latter principle postulates a learning benefit when pictures are combined with voice instead of text.

We did not find any effect of flow control. Hence, our results did not support that the possibility to control the animation information flow is supports the learning with animations. The missing effect of flow control may, however, be related to the fact that only the mono- animation presentations (no voice) were implemented with flow control. Hence, this result indicates that flow control (and expected reduction of cognitive load) may not have had any effect as long as the additional information through voice was not provided. Hence, to be able to control too little information does not help. Follow up research on this point should therefore be performed in order to investigate whether individual information flow control in the animation + voice condition would have better impact on the learning performance. Hence, our experimental setup did not enable a plausible test of the cognitive load hypothesis.

The effects of media on preference supports the assumption that animated presentations should be supplemented with more information. The animation + voice condition was more preferred than the two animation mono-media presentation forms (AC and A). The AC condition was more preferred than the A condition, which also shows that students prefer to have control over the information flow, even if it does not have immediate effects on their learning performance. Furthermore, multimedia presentations are preferred for single animated presentations, even the condition picture + text was more appreciated than the "only animation" condition.

The analysis of the third variables shows that the effects of media were valid and were not influenced by unintended factors. It is not surprising that the subjects with generally higher pre-experimental knowledge performed better. This effect was, however, not relevant for the performance of the correspondent topics.

5. Conclusions

The research reported in this paper aimed at evaluating some theoretical hypotheses suggesting relationships between the didactical setting and learning effects with animations. In particular we investigated whether individual flow-control adequately provides a didactical means to reduce the cognitive load imposed by animations. Basically this experiment did not support any of the hypotheses statistically, but the experimental setting did not leave clear indications for that the hypotheses should be rejected. The results, as discussed above leave open questions, which should be addressed in further research in which the theoretical hypotheses are operationally tested in new settings. This experiment sheds light on the complex interaction between learning content, presentation, learning and resulting knowledge.

Contrary to the expectations, the major finding in this experiment was the lack of an animation effect. This suggests that we did not offer didactical conditions in which animated information could optimally be learned. The assumption that it would be sufficient to reduce the cognitive load by giving the students full control over the information flow was not confirmed. This study also revealed that the design of the presentation can induce unnecessary cognitive load, e.g. when a visual analogy is used that induces the learners to perform their own logical translation of the relevant factors back to the habitual appearance. Consequently, the cognitive load hypotheses remains plausible, the question remains how to reduce cognitive load when learning with animations.

On a statistical level, we must reject the static-dynamic hypothesis and the knowledge modality hypothesis in this study. On a theoretical level, however, it is too early to reject the hypotheses based on this finding only. The result may have been influenced by the missing animation effect. Hence, as long as an animation effect was not found, the testing of those two hypotheses was not fully possible. Further research should continue to investigate alternative ways of implementing animations in a didactical setting that support the utilisation of the dynamic information for the learning process. Furthermore, continued research applying a differentiated knowledge concept, should be continued in order to shed light on whether particular learning content characteristics should be presented in a particular way.

Although the setting in this study did not deliver any support for the theoretical hypotheses, we are reluctant to discard the theoretical hypotheses based on the results of this study. The relationships addressed in our theoretical hypotheses are very complex, and much research is needed in order to fully understand them.

References

Alty, J. L. (1993): Multimedia: We have the technology but do we have the methodology? Educational Multimedia and Hypermedia Annual, 1993: Proceeding of Ed-Media 93-World conference on Educational Multimedia and Hypermedia, Orlando, Florida, June 1993. Orlando: AACE: 3–10.

Bètrancourt, M., Tversky, B. (2000): Effects of computer animation on users' performance: A review. Le Travail Humain, 63 (4): 311–329.

CISCO (1999): RIO Strategy v3.0 June 25, 1999, Cisco Systems, Inc. Retrieved January 14, 2003, from the World Wide Web: http://www.cisco.com/warp/public/779/ibs/solutions/learning/white-papers/el cisco rio.pdf

Engelkamp, J. (1991): Das menschliche Gedächtnis (2. Aufl.). Göttingen: Verlag für Psychologie.

Gagné, R. M. (1985): The conditions of learning and theory of instruction (4th ed.). New York: Holt, Rinehart, Winston.

Gagné, R. M. (1965): The conditions of learning. New York: Holt, Rinehart, Wilson.

Guttormsen Schär, S. (in Print): Multimedia Didactics. Aachen: Shaker Verlag.

Guttormsen Schär. S., Kaiser, J. (in Print): Revising (Multi-) Media learning Principles by applying a differentiated knowledge concept. International Journal of Human Computer Studies (in print).

Guttormsen Schär, S., Messerli, S., Schluep, S. (2004): Information categories for learning-object metadata. An empirical study on the practical validity of theoretically founded information categories. Ed-Media 2004 , Lugano, Switzerland.

ISO 14915-3, Norm Schluss-Entwurf prEN; ICS 13.180; 35.200, Juni 2002 (Software-Ergonomie für Multimedia – Informationsinhalte nach ISO Europäische Benutzerschnittstellen – Teil 3: Auswahl und Kombination von Medien (ISO/FDIS 14915-3: 2002).

Kinzer, C. K., Sherwood, R. D., Loofbourrow, M. C. (1989): Simulation Software versus expository text: A comparison of retention across two instructional tools. Reading Research and Instruction, 28: 41–49.

Large, A. (1996): Computer animation in an instructional environment. LISR, 18: 3–23.

Levin, J. R., Anglin, G. J., Carney, R. N. (1987): On empirically validating functions of pictures in prose. In: D. W. H. Willows, H. A. (Ed.): *The psychology of illustration,* (Vol. I, Basic research: 51–85). New York: Springer.

Mayer, R. E. (2001): Multimedia learning. Cambridge: Cambridge University Press.

Merrill, M. D. (1983): Component display theory. In: Instructional-design theories and models: An overview of their current status. Hillsdale: Lawrence Erlbaum associates, Publishers.

Merrill, M. D. (1996): Instructional transaction theory: An instructional design model based on knowledge objects. Educational Technology, 36 (3): 30–37.

Merrill, M. D. (2000): Knowledge objects and mental models. (Vol. 2003).

Merrill, M. D. (2001): Components of Instruction. Toward a theoretical tool for instructional design. Instructional Science, 29 (4): 291–310.

Paivio, A. (1986): Dual coding theory. In: A. Paivio (Ed.): Mental representations. A dual coding approach. Oxford: Oxford University Press: 53–83.

Pane, J. F., Corbett, A. T., John, B. E. (1996): Assessing dynamics in computer-based instructions. In: Proceedings of the International conference on Computer Human Interaction CHI'96: 197–204. New York: ACM Press.

Park, I., Hannafin, M. J. (1993): Empirically-based guidelines for the design of interactive multimedia. Educational Technology Research and Development, 41 (3): 63–85.

Park, O.-C., Hopkins, R. (1993): Instructional conditions for using dynamic visual displays: a review. Instructional Science, 21: 427–449.

Peeck, J. (1993): Wissenserwerb mit darstellenden Bildern. In: B. Weidenmann (Ed.), Wissenserwerb mit Bildern, Instruktionale Bilder in

Printmedien, Film/Video und Computerprogrammen. Verlag Hans Huber: 59–94.

Rieber, L. P. (1989): The effects of computer animated elaboration strategies and practice on factual and application learning in an elementary science lesson. Journal of Educational Computing Research, 5 (4): 431–444.

Rieber, L. P. (1990 a): Animation in computer-based instruction. Educational Technology Research & Development, 38 (1): 77–86.

Rieber, L. P. (1990 b): Using computer animated graphics in science instruction with children. Journal of Educational Psychology, 82: 135–140.

Rieber, L. P. (1991 a): Animation, incidental learning, and continuing motivation. Journal of Educational Psychology, 83 (3): 318–328.

Rieber, L. P. (1991 b): Effects of visual grouping strategies of computer-animated presentations on selective attention in science. Educational Technology Research and Development, 39 (4): 5–15.

Rieber, L. P., Boyce, M. J., Assad, C. (1990): The effects of computer animation on adult learning and retrieval tasks. Journal of Computer Based Instruction, 17 (2): 46–52.

Rieber, L. P., Hannafin, M. J. (1988): Effects of textual and animated orienting activities and practice on learning from computer-based instruction. Computers in the Schools, 5 (1/2): 77–88.

Rieber, L. P., Kini, A. S. (1991): Theoretical foundations of instructional applications of computer-generated animated visuals. Journal of Computer Based Instruction, 18 (3): 83–88.

Sweller, J. (1999): Instructional design in technical areas (Vol. 43): Camberwell, Victoria, Acer Press. ■