



How pacing of multimedia instructions can influence modality effects: A case of superiority of visual texts

Klaus D. Stiller, Annika Freitag, Peter Zinnbauer and Christian Freitag
Technische Universität Dortmund

“Present text accompanying pictures aurally to promote learning” is a well established principle of instructional design. But recently, it was shown that under certain conditions visual texts can be preferable. Instructional pacing seems to be one of these conditions that mediate effects. Especially, enabling learners to pace an instruction has proven to enhance learning. In an online experiment, the effects of pacing and text modality on cognitive load and performance were investigated. 110 university students studied instructions paced either by the system or by the learner and using either visual or auditory texts accompanying static pictures. Learner pacing reduced mental effort and promoted performance for instructions using visual texts, whilst with auditory texts, only mental effort was positively influenced. Modality effects on performance were found in favour of auditory texts with system paced instructions, but effects vanished or reversed with learner pacing. With respect to study time, the learner paced, visual text instruction was the most efficient.

Introduction

The modality principle of multimedia design recommends using spoken rather than written texts in conjunction with dynamic or static pictures, because learners learn more effectively in computer based and book based environments when spoken texts are used (Low & Sweller, 2005; Mayer, 2005). The superiority of spoken texts was demonstrated in multiple studies (Ginns, 2005; Low & Sweller, 2005) and observed in various measures: less mental effort with spoken words while learning (Tabbers, 2002; Tindall-Ford, Chandler & Sweller, 1997; van Gerven, 2002), less time on solving or answering performance tasks (Jeung, Chandler & Sweller, 1997; Mousavi, Low & Sweller, 1995; Rinck & Glowalla, 1996), higher scores on various tasks of retention, matching verbal and pictorial information, labeling pictures, transfer and practice (Kalyuga, Chandler & Sweller, 1999, 2000; Craig, Gholson & Driscoll, 2002; Mayer & Moreno, 1998; Moreno & Mayer, 1999; Schmidt-Weigand, 2006; Stiller, 2007), and shorter reaction time in a secondary task (Brünken, Plass & Leutner, 2004; Brünken, Steinbacher, Plass & Leutner, 2002).

Older studies, which support modality effects, are limited concerning subject matter, length of study time, settings and pacing of presentation. The implemented instructions contained contents taken from the exact sciences such as geometry, mathematics and physics, and they mostly did not last longer than 5 minutes; they were conducted in laboratory settings and the pacing of presentations was determined by the computer system or the procedure in non-computer settings (e.g. Mayer, 2001; Mousavi et al., 1995). Since then, however, empirical evidence of modality effect has also been found with learner paced or learner sequenced instructions consisting of up

to 90 minutes learning time in classroom and laboratory settings, sometimes using non-scientific contents (e.g. Brünken, Seufert & Zander, 2005; Moreno & Mayer, 2002; Mayer, Dow & Mayer, 2003; Moreno, Mayer, Spires & Lester, 2001; O'Neil et al., 2000; Stiller, 2007; Tabbers, 2002). Thus, evidence of modality effects is founded on a wide range of learning situations.

Explaining modality effects

One of the prominent theories which describe and explain meaningful learning with texts and pictures is the cognitive theory of multimedia learning (CTML; Mayer, 2001). It is based on the dual channel assumption, the limited capacity assumption and the active learning assumption. The dual channel assumption proposes that pictorial and verbal information is processed in two separate systems, a visual/pictorial and an auditory/verbal channel (see Figure 1), respectively. They are functionally independent of each other, but may interact. Written words are initially processed in the visual/pictorial channel before they move to the auditory/verbal channel in the working memory.

The limited capacity assumption suggests that the information processing capacity in each of the two channels is limited. Within capacity limits the learner has to accomplish all the processing necessary for learning. Especially the cognitive load theory (CLT; Sweller, 1999) emphasises the role of the working memory in learning and the problem of occupying working memory by useful and unnecessary processes. The components of cognitive load are intrinsic, germane and extraneous load. Intrinsic load is created by task complexity and is mainly needed for understanding. Germane load goes beyond understanding and is bound to the learning processes in a real sense – processing information in order to build mental representations and storing them. Extraneous load results from the presentation manner of the material and is often the main source that is detrimental to learning. As long as extraneous, intrinsic and germane load do not occupy the entire working memory capacity, some memory capacity is still available for processing. Mayer and Moreno (2002) regard the CLT as an important component of the CTML.

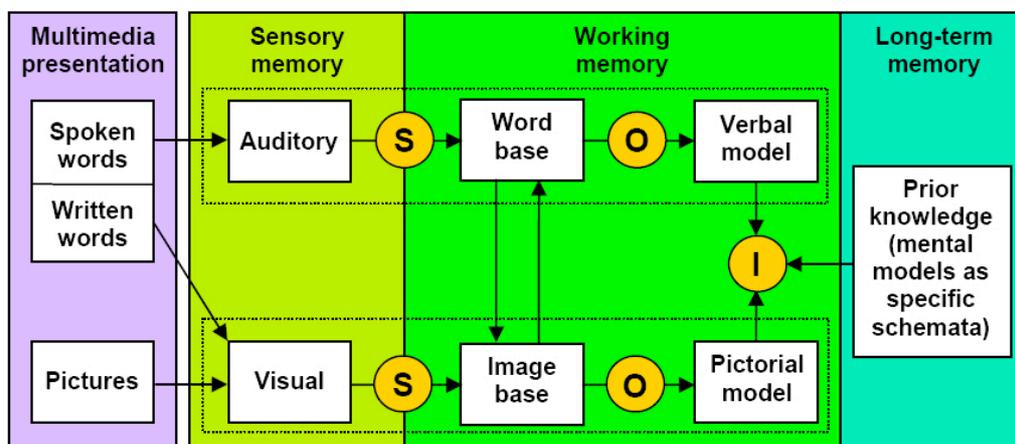


Figure 1: Framework for the CTML (Mayer, 2001)

S = selecting, O = organising, I = integrating

The active learning assumption states that learners have to process information actively in order to achieve meaningful learning. Active cognitive processing includes the selection of relevant information from words and pictures, the mental organisation of this information to form coherent verbal and pictorial representations, and the mental integration of these representations with one another and with prior knowledge from long term memory (see Figure 1).

Modality effect is thought to be composed of visual/pictorial processing channel overload and split attention effect. Firstly, when words and pictures are presented visually, texts exert extraneous load on working memory, because they at least initially use the visual/pictorial channel before the information is transmitted to the auditory/verbal channel. On the other hand, auditory texts enter the auditory/verbal channel directly, thus releasing the visual/pictorial channel from unnecessary load. Secondly, a split attention effect occurs because visual attention has to be spread between visual texts and pictures, which must be integrated mentally to achieve comprehension. This means, for example, that learners must search for corresponding information from texts and pictures, which also often results in less relevant material being selected for further processing (especially with system paced instructions including dynamic presentations), and keep it concurrently in working memory in order to enable integration processes. Splitting attention exerts a heavy extraneous load on working memory. With auditory texts, no splitting of attention is necessary. A simultaneous audio-visual presentation increases the likelihood of learners being able to select relevant information from texts and pictures and to retain corresponding words and pictures in working memory at the same time. Thus it is also more likely that learners engage in active meaningful learning, in particular in integrative processing of verbal and pictorial information.

Learner pacing of instructions and modality effects

In a further step to improve meaningful learning it was proposed to allow learners to pace the instructions. To evidence this, researchers experimentally compared system paced with learner paced instructions. An instruction was system paced when the system determined how long information was presented. Learner pacing was realised by allowing learners to start a part of presentation by their choice after the preceding part has finished, or to choose by themselves when a presentation should be paused and played using interactivity facilities (Hasler, Kersten & Sweller, 2007; Schwan & Riempp, 2004). Typical presentations integrated dynamic visualisations and corresponding texts (Mayer & Chandler, 2001).

As we know from modality effects, audio-visual presentations like narrated animations (Mayer, 2001) are already optimally designed due to split attention effect and the usage of modality specific processing channels, but enabling learners to pace instructions might improve information processing still further. Mayer and Chandler (2001) proposed that learner pacing reduces extraneous cognitive load on working memory and thus enables learners to build a more coherent mental model step by step. Learner pacing allows learners to stay longer on parts of information in order to understand them before the next part is processed; there is more time to get more of the relevant information into working memory as well as to organise and integrate it into a more coherent and complete mental model. This seems to be especially helpful with time dependent presentations, which only enable learners to perceive and process information within a given time period, and with instructions in split attention format (Tabbers, 2002).

The superiority of learner pacing was demonstrated in multiple studies and observed in various measures: less mental effort while learning (Hasler, Kersten & Sweller, 2007; Moreno, 2007), less time on acquiring a skill (Schwan & Riempp, 2004), and higher scores on various tasks of retention and transfer (Hasler, Kersten & Sweller, 2007; Mayer & Chandler, 2001; Moreno, 2007). Empirical evidence has been found with short (less than 10 min) and longer lasting (from 10 to 20 min.) instructions (Hasler, Kersten & Sweller, 2007; Mayer & Chandler, 2001; Schwan & Riempp, 2004, and Moreno, 2007, respectively) using dynamic visualisations, i.e. animations and videos (Hasler, Kersten & Sweller, 2007; Mayer & Chandler, 2001, Moreno, 2007, and Moreno, 2007, Schwan & Riempp, 2004, respectively), with adults and children (Mayer & Chandler, 2001; Moreno, 2007; Schwan & Riempp, 2004, and Hasler, Kersten & Sweller, 2007, respectively), on retention and transfer performance (Mayer & Chandler, 2001; Hasler, Kersten & Sweller, 2007; Moreno, 2007) and on practical skills (Schwan & Riempp, 2004). No effects of pacing were found often for retention or relatively easy performance tasks (Ertelt, Renkl & Spada, 2005; Hasler, Kersten & Sweller, 2007; Mayer & Chandler, 2001), as well as motivation and instructional acceptance (Hasler, Kersten & Sweller, 2007). Studies are limited concerning the experimental setting: they were mostly conducted in the laboratory with participants individually tested and using narrated dynamic visualisations. Only Schwan and Riempp (2004) used pure dynamic visualisations to teach how to tie nautical knots.

Additionally, Mayer et al. (2003) compared an interactive environment on the functioning of an electric motor to its non-interactive linear version. The interactive and non-interactive instructions were paced and sequenced by learner and system, respectively. The interactive instruction presented a picture of an electric motor; clicking on one of its five parts caused the presentation of five questions each. After clicking on a question, a spoken answer was presented, sometimes accompanied by an animation. The learners could reload any information unit as many times as they wanted to. Students having learned with the interactive instruction proved to be better at problem solving transfer than students having learned with the non-interactive instruction, when transfer was measured immediately after and a week after the learning phase. Unfortunately, study time and usage of the interactive environment were not reported. Moreover, the contributions from pacing and sequencing are not specifiable.

Older research about learner control over pacing in computer programs showed inconsistent results (Dillon & Gabbard, 1998; Niemiec, Sikorski & Walberg, 1996; Reeves, 1993; Williams, 1996), and therefore is not helpful. Two problems were identified: a lack of theoretical foundations for the experiments (Reeves, 1993; Williams, 1996), and a methodological problem due to incomplete work through of learner controlled, compared to system controlled instructions (Lepper, 1985; Williams, 1996).

When learner pacing is suited to reduce split attention effects, it should also influence modality effects (Tabbers, 2002), but only rarely evidence was obtained with experiments that dealt with the pacing of instructions by learner, or system, or that used learner paced instructions alone (Ginns, 2005). Tabbers (2002) and Stiller (2007), for instance, revealed that there are situations in which visual text was better than auditory text. Schmidt-Weigand (2006) showed that modality effects might even vanish when the most demanding presentation modes consisting of commented animations were used.

Schmidt-Weigand (2006) conducted an experiment on modality effect, using learner paced animations on lightning formation. The animation consisted of 16 scenes. After a scene had been played, the learner could decide when to start the next scene. No modality effects were observed as regards time on task ($M = 183.26$ seconds, $SD = 42.68$), cognitive load, verbal retention, transfer and drawing performance. It was concluded that learners chose a presentation pace that matched their individual speed of text comprehension. Learner pacing was suitable for compensating split attention, hence no effects were found. In another experiment, using three speeds for system paced animations, a slower presentation rate led to marginally more success at retention and visual knowledge, but not at transfer. Additionally, only a modality effect on visual knowledge was found and modality effect did not depend on pacing speed.

Tabbers (2002) used web based lessons that lasted from 26 to 70 minutes and dealt with a didactical model of instructional design (static diagrams were explained by auditory or visual texts), focusing especially on learner pacing. Information units were put in linear order and, when instructions were self paced, learners could move forward (and sometimes backward) along this line. In four experiments, two new aspects were found: performance increased and modality effect vanished with more time on task when using system paced instructions, and modality effect vanished or more often reversed for retention and transfer with learner pacing. The vanishing of modality effect was attributed to extended time on task, which is responsible for compensating the split attention effect. The reversing of modality effect was attributed to a higher quality of information processing with on screen texts.

Stiller (2007, Exp. 2) investigated modality effects with hierarchically structured hypertexts on the human eye, which were paced and sequenced by learners. Texts were accessed by clicking on parts of pictures. 10th grade German grammar school pupils learned with visual or auditory texts for 13 minutes. Overall, transfer and drawing performance were not affected by text modality, but a superiority of visual over auditory texts was found as regards verbal retention and labelling pictures. These results contradicted another study by Stiller (2007, Exp. 1), using similar instructions. With university students, a modality effect on structural knowledge measured by completing pictures by drawing and a verbal description task were found in favour of auditory texts; retention, transfer and labelling pictures were not affected.

Research questions and hypotheses

In the following experiment, the effects of instructional pacing as well as text modality on cognitive load and performance were investigated. Therefore, students were asked to work on retention, structural, labelling and transfer tasks and to rate mental effort while learning after studying an instruction.

Firstly, learner pacing was assumed to foster learning, regardless of text modality used in instructions. Learners who can pace the instruction should gain more knowledge and experience less mental effort than learners studying system paced presentations. The pace of a presentation determines the time available for information processing. If learners can affect the pace, they choose one that matches their individual speed of comprehension. This also means that they can spend as much time as they need with the information presented, i.e. learners are more likely to select relevant information, to organise and to integrate it adequately. Hence it is proposed that learner pacing is

suitable for preventing working memory from being overloaded, or at least relieving it from extraneous load (Tabbers, 2002; Schmidt-Weigand, 2006).

Secondly, modality effects in favour of auditory texts were assumed to occur due to split attention effects, which mainly constitute modality effects. But moreover, effects are suggested to vary with pacing mode: modality effects should be observed for system paced instructions and decrease with learner paced instructions, because learner pacing reduces split attention effects. This pattern of effects is also reflected by the effect sizes reported by Ginns (2005). It might even be more likely that modality effect would vanish (Schmidt-Weigand, 2006) or reverse in favour of visual texts (Tabbers, 2002). The latter could be the case when learners are able to regulate their learning more adequately (Tabbers, 2002).

Method

Participants

118 students of German universities participated in the experiment. 7 students were omitted from the data base, because they attained 100% correct answers in the prior knowledge test, hence they were assumed to not improve their knowledge by studying the instructions. One student was eliminated, because he scored zero points in the performance tests. As this is very unlikely, this student was presumed to be a non-serious participant who did not invest any effort in learning. Finally, 110 data sets were analysed. Eighty-four students were female and 26 male, with a mean age of 23.2 years ($SD = 3.5$) and a mean semester of study of 4.4 ($SD = 2.9$). Experimental groups did not differ in aspects of sex (Chi-square test with 4 groups \times 2 sexes; frequencies of male/female = 7/22 and 7/21 for the system paced visual and auditory group, respectively; frequencies of male/female = 6/19 and 6/22 for the system paced visual and auditory group, respectively; Chi-square (3) = .11, ns), age (one way ANOVA; system paced: M-visual = 24.07, $SD = 4.53$, M-auditory = 22.68, $SD = 2.99$; learner-paced: M-visual = 22.24, $SD = 1.72$, M-auditory = 23.75, $SD = 3.68$; $F(3, 106) = 1.72$, ns), and semester of study (one way ANOVA; system paced: M-visual = 4.76, $SD = 3.23$, M-auditory = 3.81, $SD = 2.83$; learner paced: M-visual = 4.68, $SD = 2.73$, M-auditory = 4.48, $SD = 2.67$; $F(3, 106) = .60$, ns).

The announcement about participating in this experiment was published at two German universities (TU Dortmund and University of Regensburg), therefore most students were studying at these universities: Sixty of them studied at TU Dortmund, 35 at University of Regensburg, and 15 at other universities. Most students studied psychology or pedagogical subjects. Experimental groups did not differ in aspects of university membership either (Chi-square test with 4 groups \times 3 universities; frequencies of Dortmund/ Regensburg/ other = 18/7/4 and 14/12/2 for the system paced visual and auditory group, respectively; frequencies of Dortmund/ Regensburg/ other = 12/9/4 and 15/8/5 for the learner paced visual and auditory group, respectively; Chi-square (6) = 3.70, ns).

Multimedia instructions

The learning material covered the structure of the human eye. Four treatment instructions were created by combining two types of text presentation with two types of instructional pacing. The presentations were either paced by system or learner, the text accompanying a schematic picture of a horizontal section through the eye (see

Figures 2 and 3) was either presented visually or auditory. System pacing meant that the instructional units were presented one after the other without a break in between. First, the picture was verbally introduced (see Figure 2). Then twelve parts of the eye were explained in the following order: cornea, pupil, iris, lens, ciliary muscle, vitreous body, sclera, choroid, retina, optic disc, fovea, and optic nerve. While the text belonging to a pictorial part could be read or heard, the pictorial part was highlighted in green until the next text unit was presented. The presentation durations of visual texts were equal to the durations of the corresponding auditory texts. Learner pacing meant that the learner chose when the next information unit should be presented. The presentation speed of auditory text was about 119 words per minute. Texts (812 words) and pictures did not differ between treatment groups.

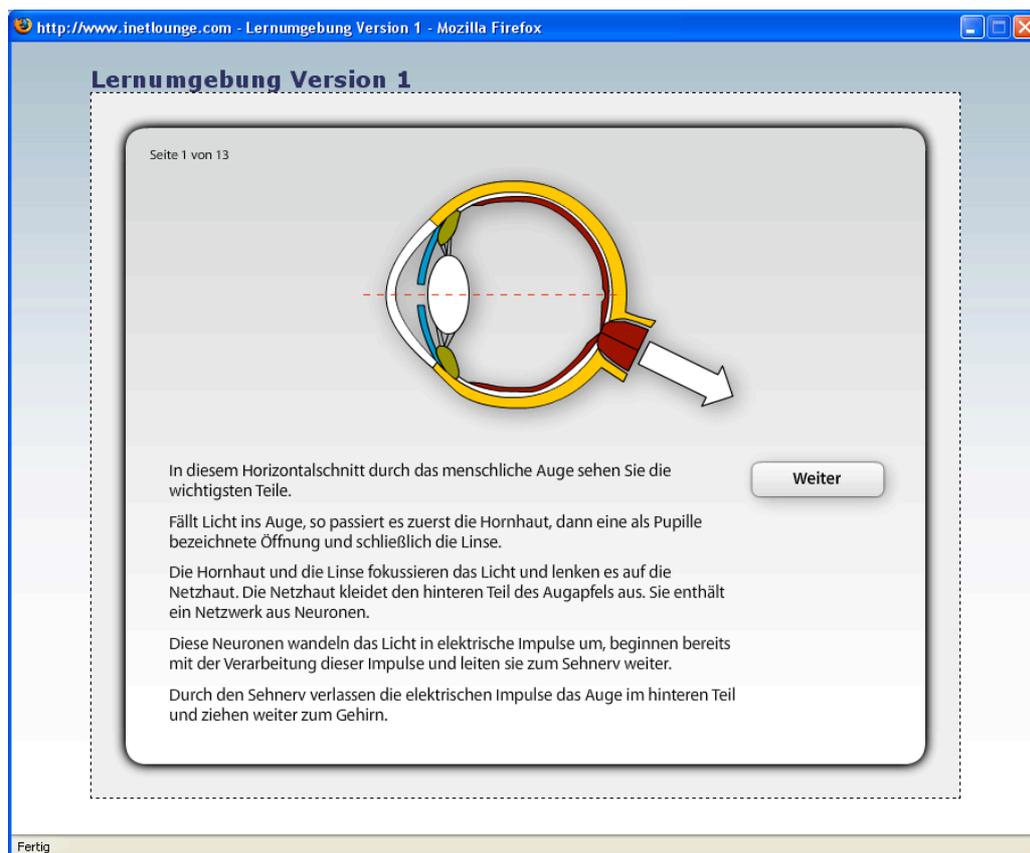


Figure 2: Screenshot of the learner paced, visual text instructions

A couple of numbers that were placed in the top left hand corner of the screen signalled the number of text units in the sequence and at what position of the sequence the learner was. Additionally, in the system paced conditions, the presentation durations of the text units were shown via a count down from total playing time to 0. In the learner paced auditory text condition, the learners could replay the auditory texts.

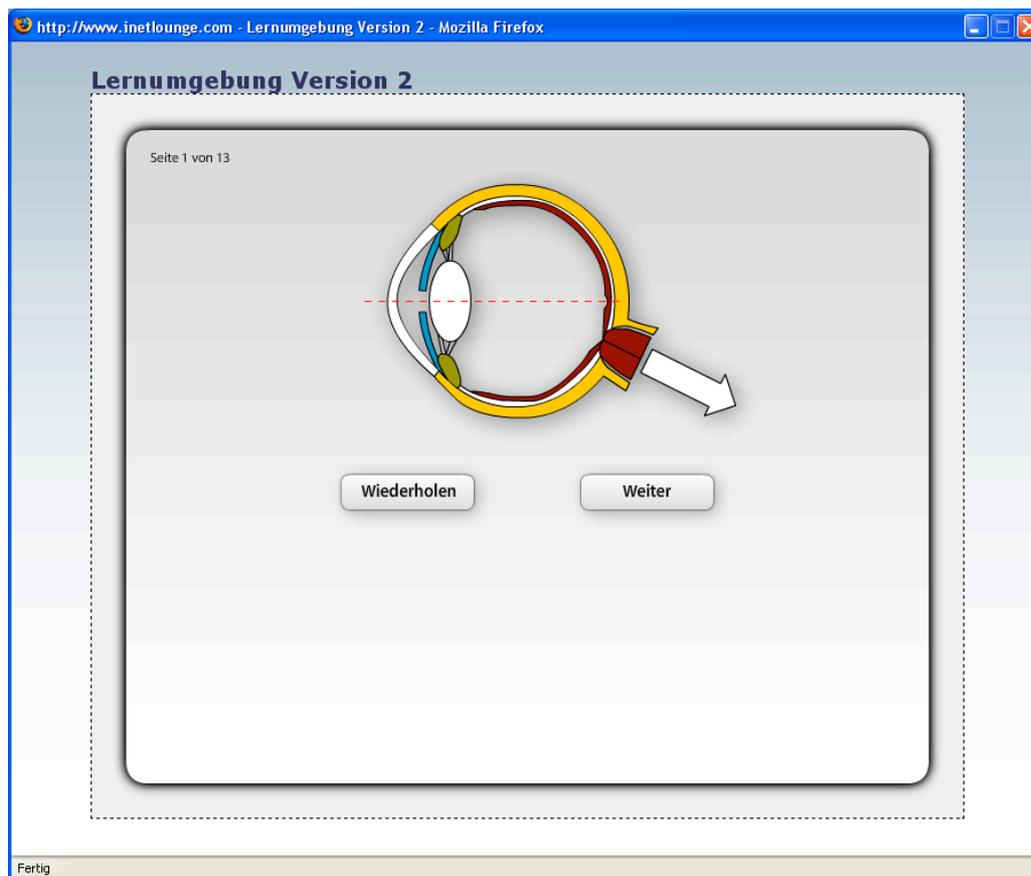


Figure 3: Screenshot of the learner paced, auditory text instructions

Procedure

The experiment was conducted via the WWW. After requesting the URL, the students were introduced to the experiment, its procedure, and the technical affordances. Reading this took about 1 minute. Next, they were asked about their sex, age, university membership, semester of study, and subject and had to commit that their participation was based on seriousness and compliance. Finally, they completed a test of prior knowledge. This took about 2 minutes. After this pre-testing, participants were randomly assigned to one of the four treatment groups: Twenty-nine students served in the system paced, visual text presentation group, 28 in the system paced, auditory text group, 25 in the learner paced, visual text group, and 28 in the learner paced, auditory text group. Unbalanced participants resulted from eliminating students from the sample, due to the criteria mentioned in the first subsection, *Participants*, of the *Methods* section (8 students), technical problems (10 students), students not speaking German as mother language (1 person) and the appearance of some curious, but unserious participating lurkers (32 persons).

Before starting the instruction, the groups were introduced briefly to its handling. Then, the students learned for roughly 7 minutes on average with their treatment

instruction. Subsequently, the students rated the overall difficulty of the presentation and the content, as well as the pleasantness of the kind of presentation. This took about 30 seconds. Last, they took a performance test for 30 minutes maximum. The whole experiment required an average duration of about 21 minutes.

Means of measurement

Prior knowledge, three aspects of cognitive load as well as four kinds of performance were assessed. Table 1 lists each of the questions constituting the measurement scales used.

Table 1: Listing of the questions constituting the measurement scales used

Scales	Questions
Prior knowledge	1 What is regulated by the ciliary muscles?
	2 What happens in the optic chiasm?
	3 What is the function of the choroids?
	4 What kinds of photo receptors exist?
	5 What is the area of the retina called on which the projection of an object falls when it is fixated?
	6 What is the area of the brain called, which is mainly responsible for vision?
	7 What is the blind spot of an eye?
	8 What cells of the retina build up the optic nerve?
Overall cognitive load	1 How easy or difficult was it for you to learn something about the eye from the presentation you just saw?
Intrinsic cognitive load	2 How easy or difficult would you consider the content?
Extraneous cognitive load	3 How pleasant or bothersome would you consider the presentation format?
Verbal structure task	Imagine you are a visual stimulus on your way through the eye.
	1 Please, name all "places" and "streets" you are passing on your way. (Places = structures or similar elements, streets = neuronal paths or similar entities).
	2 Leave out all "places" and "streets", you don't pass by.
	3 If you don't know the exact name of a place or street, simply write down, how we can identify the "place" or "street" (for example, its function, what happens there, a short description).
	Now start listing up the elements according to the sequence of contact with the visual stimulus. Start with the first part of the eye.
Transfer task	Imagine that a friend of yours no longer sees very well with one eye. What possible reasons are there for this? Please list all your ideas about what parts of the eye are malfunctioning!
Labelling task	Here, once again you see the cross section of the eye. Please label the numbered elements.
Verbal retention task	1 Which process takes place in the retina?
	2 What is the function of the choroid?
	3 Which parts of the eye are responsible for a sharp projection of the world on the retina? (coded into two items)
	4 What is the function of the iris?
	5 What elements form the optic nerve?

Prior knowledge was measured by eight multiple choice questions. Four answers were offered for each question along with an option to mark "I do not know". One of the answers was correct. Each answer was scored with a maximum of one point.

The presentations were followed by three questions assessing three kinds of cognitive load. Participants were instructed to place a check mark for each question on a 7-point rating scale from 1-very easy (pleasant), 2-easy (pleasant), 3-rather easy (pleasant), 4-medium, 5-rather difficult (bothersome), 6-difficult (bothersome), to 7-very difficult (bothersome). Question 1 is a standard item for subjective ratings of overall cognitive load (see Table 1; e.g. Paas & van Merriënboer, 1993, 1994). Questions 2 and 3 were introduced to differentiate between intrinsic and extraneous load (see Table 1; Schmidt-Weigand, 2006).

The performance test consisted of a verbal structure task, a transfer task, a pictorial labelling task, and a verbal retention task.

- The structure task required the students to name those structures of the eye passed by a visual stimulus on its way through the eye, and to arrange them according to the sequence of contact (cornea, pupil, lens, vitreous body, retina, optic disk, and optic nerve). For each element listed in the correct order a maximum of one point was scored. The elements that were mentioned without being in the list of correct answers (like ciliary muscle or sclera) were also counted (superfluous details).
- The transfer task demanded that learners use their knowledge in order to find explanations for a given situation. Based on the presented information, 16 reasons were accepted. They were coded into nine items, as multiple causes that referred to one part of the eye (e.g. lens is clouded or inflexible) were put together.
- Then the participants had to name nine parts of the eye (cornea, iris, lens, ciliary muscles, vitreous body, sclera, choroid, retina, and optic nerve). The same picture was used as in the presentations and the relevant parts were marked. For each item 1 point was scored in maximum.
- Retention was assessed by five open questions coded into six items, which were scored with a maximum of one point.

Figures 4 to 7 show what the students saw on their computers when undertaking the various tasks of the performance test.

Table 2 lists the number of items, mean, standard deviation, reliability (Cronbach's alpha, a measure of internal consistency) and objectivity of the measurement scales used. Performance scales were normed to the range of 0 to 100, thus enabling scores to be interpreted as percent correct answers; for superfluous details, only their frequency was counted. Cronbach's alpha of the transfer scale was low, but as all tests are assessed as criteria tests, this is not relevant. The low alpha results from a relatively low production of correct answers (and thus a low standard deviation) facing a relatively high number of accepted answers, or in other words, a single item is not very predictive of the sum score that indicates a lack of item order in aspects of difficulty. Content validity of all performance tests is given; the questions of the performance tests were constructed from the presented content of the instructions. Two raters scored all of the performance tests. The Pearson correlation of sum scores was used as a measure of scale objectivity. Cognitive load scales were normed to their rating range 1 to 7. The computer registered the time needed for presentation.

The study did not provide any feedback from the participants on their experience of going through the experiment. Therefore, no further insights into the study were provided from the participants' point of view on why differences between presentations might have arisen.

Fragebogen II - Nachfragebogen

Stell Dir bitte vor, Du bist ein Sehreiz auf dem Weg durch das Auge

- (1) **Zähle** bitte alle „Orte“ und „Straßen“ **auf**, die Du auf Deinem Weg passierst.
(Orte = Strukturen oder Ähnliches – Straßen = Leitungsbahnen oder Ähnliches).
- (2) **Lasse** alle „Orte“ und „Straßen“ **weg**, die von Dir als Sehreiz nicht besucht werden.
- (3) Wenn Du den genauen Namen eines „Orts“ oder einer „Straße“ nicht weißt, schreib einfach hin, woran wir den „Ort“ oder die „Straße“ erkennen (z.B. Funktion, was passiert da, **irgendeine kurze Umschreibung**).

Beginne nun mit der Aufzählung entsprechend der Reihenfolge des Wegs des Sehreizes durch das Auge. Beginne mit dem ersten Teil des Auges.

Trenne die Orte/Straßen durch **

Seite 2 von 9

Figure 4: Display window of the verbal structure task

Fragebogen II - Nachfragebogen

Stell Dir vor, ein Freund oder eine Freundin kann mit einem Auge nicht mehr gut sehen. Woran könnte das liegen? Zähle bitte alle Gründe auf, die dir einfallen, was im oder am Auge nicht mehr richtig funktionieren könnte.

Trenne die Gründe durch **

Seite 3 von 9

Figure 5: Display window of the transfer task

Table 2: Features of the measurement scales
 (alpha = Cronbach's alpha; r(interrater) = objectivity; n = 110)

Scale		Items	M	SD	alpha	r(interrater)
Control scale	Prior knowledge	8	40.23	27.53	.73	--
Cognitive load	Overall load	1	3.30	1.30	--	--
	Intrinsic load	1	3.75	1.18	--	--
	Extraneous load	1	3.08	1.31	--	--
Performance	Verbal retention	6	46.44	29.47	.74	.99
	Verbal structure: correct details	7	59.74	26.47	.68	.90
	Verbal structure: superfluous details	--	.61	.78	--	.99
	Labelling pictures	9	64.95	26.54	.75	.99
	Transfer	9	20.23	10.47	.30	.94

Fragebogen II - Nachfragebogen

Hier ist noch einmal ein Horizontalschnitt des Auges von oben.

Bitte beschrifte die einzelnen Bestandteile!

(1)

(2)

(3)

(4)

(5)

(6)

(7)

(8)

(9)

Seite 4 von 9

Figure 6: Display window of the labelling task

Fragebogen II - Nachfragebogen

Welcher Prozess findet in der Netzhaut statt?

Bitte hier Antwort eingeben...

WEITER

Seite 5 von 9

Figure 7: Display window of the first question of the verbal retention task. The other four questions were presented in the same manner.

Results

Originally, prior knowledge, which is known to influence learning outcomes, was planned to be used as covariate in a series of 2x2 between subjects factorial analyses of covariance, but slopes of the regression lines representing the relationship between the dependent variables and the covariate in the four groups could not be considered as equal. Hence, the treatment groups were tested for differences in prior knowledge and, as no differences were found (system paced, visual text group: $M = 34.05$, $SD = 27.53$, system paced, auditory text group: $M = 41.52$, $SD = 27.53$; learner paced, visual text group: $M = 46.00$, $SD = 24.40$; learner paced, auditory text group: $M = 40.18$, $SD = 27.92$; one way ANOVA, $F(3,106) = .87$, ns), a series of 2x2 between subjects factorial analyses of variance (ANOVA) was conducted with text modality (visual vs. auditory) and presentational pacing (system vs. learner) serving as between subject factors.

As dependent variables, cognitive load and performance were assessed. Assumptions of normality and homogeneity of variance could mostly be retained. Normality of distributions was mainly violated concerning data reporting superfluous details (right skewed data distributions; many students did not report any superfluous details) and study time (right skewed data distributions, due to presentational conditions). As an ANOVA is regarded to be very robust to most plausible violations of the usual assumptions (Bortz, 2005), the strategy of analyses was not changed. Partial eta squared effect sizes (h_p^2) were computed; h_p^2 of .01 are defined as small, of .06 as medium, and of .14 as large (Cohen, 1988).

Results for cognitive load ratings only revealed a significant main effect of pacing on extraneous load (see Table 3). With learner pacing, students experienced lower extraneous load than with system pacing (see Table 4), for overall load a corresponding marginal effect could be observed ($p < .10$). No significant main effects of text modality or pacing by modality interaction effects were found (see Table 3).

Results for performance showed significant main effects of pacing, non-significant main effects of text modality as well as significant pacing by modality interaction effects (see Table 3). Effects were found for all performance tasks except transfer.

Table 3: ANOVAs of cognitive load and performance (* $p \leq .05$, ** $p \leq .01$)

	Pacing		Text modality		Pacing x modality	
	F(1,106)	h_p^2	F(1,106)	h_p^2	F(1,106)	h_p^2
Overall load	3.09	.03	.01	.00	.03	.00
Intrinsic load	.13	.00	.29	.00	.54	.01
Extraneous load	6.43**	.06	.53	.00	.01	.00
Verbal retention	5.28*	.05	.33	.00	1.90	.02
Verbal structure: correct details	.24	.00	.16	.00	5.93*	.05
Verbal structure: superfluous details	6.42**	.06	.47	.00	.02	.00
Labelling pictures	5.14*	.05	1.00	.01	5.71*	.05
Transfer	.87	.01	1.10	.01	.36	.00
Study time	.46	.00	40.27**	.28	30.20**	.22

Pacing had a significant effect on verbal retention. With learner paced presentations, students could recall more facts in the retention task than with system paced instructions (see Table 4). This effect is mainly due to the visual text groups, the auditory groups do not differ as much in performance, but no pacing by modality effect was found. A modality effect on retention did not occur either.

Table 4: Mean scores and standard deviations of cognitive load and performance measures

	System paced presentations				Learner paced presentations			
	Visual text		Auditory text		Visual text		Auditory text	
	M	SD	M	SD	M	SD	M	SD
Overall load	3.52	1.45	3.50	1.40	3.04	1.02	3.11	1.23
Intrinsic load	3.72	1.19	3.68	1.22	3.64	1.22	3.93	1.12
Extraneous load	3.48	1.38	3.29	1.41	2.84	1.14	2.68	1.19
Verbal retention	35.06	26.48	45.83	29.87	55.33	27.52	50.89	31.29
Verbal structure: correct details	51.72	27.11	65.82	22.11	66.29	18.84	56.12	33.20
Verbal structure: superfluous details	.83	.80	.75	.89	.48	.51	.36	.78
Labelling pictures	51.53	28.40	68.06	25.46	74.22	20.72	67.46	26.22
Transfer	17.67	10.31	20.98	10.65	20.75	10.00	21.65	10.96
Study time (min:sec)	7:09	0:13	7:20	0:13	5:49	1:34	8:22	1:38

For structural knowledge as measured by listing correct elements, a significant interaction effect was found. Thereby, a modality effect occurred with system paced presentations (post-hoc Welch test; $t(53.53) = 2.15$, $p \leq .02$, one sided). Learners using auditory texts could report more relevant structural items than learners using visual texts. In the learner paced conditions, the effect seems to invert (post-hoc Welch test; $t(43.61) = -1.39$, $p \leq .09$, one sided). The group with visual texts reported more relevant items than the auditory text group (see Table 4). Alternatively, one can state that learner pacing fostered performance when visual texts were used in instructions (post-hoc Welch test; $t(49.40) = 2.32$, $p \leq .01$, one sided), for auditory texts, pacing had a reversing effect (post-hoc Welch test; $t(47.01) = -1.29$, $p \leq .10$, one sided). This

interaction effect is accompanied by a main effect of pacing on listing superfluous details. With learner paced presentations, students listed less irrelevant elements in the structure task (see Table 4). In summary, the learner paced, visual text group showed the best performance (high reproduction of relevant items paired with a low rate of superfluous details) and the system paced, visual text group was the worst (low reproduction of relevant items paired with a higher rate of superfluous details). The auditory groups lay in between. Thereby, listing correct and superfluous details correlated positively. The system paced, auditory text group concurrently reproduced more relevant as well as irrelevant items, the learner paced, auditory text group listed less relevant as well as irrelevant items.

In respect of labelling pictures, a significant main effect of pacing occurred concurrently with a significant pacing by modality interaction (see Table 3). Table 4 shows that these effects are due to the worse performance of the learners in the system paced group using visual texts, in comparison to the system paced, auditory text group, which was as successful in learning as the two learner paced groups. Thus, learner pacing only fostered the labelling of pictures when visual texts were used in instructions (post-hoc Welch test; $t(50.70) = 3.38$, $p < .001$, one sided), for auditory texts, pacing did not affect labelling performance (post-hoc t-test; $t(54) = -.09$, ns). Alternatively, one can state that a clear modality effect with system paced presentations was revealed (post-hoc t-test; $t(55) = 2.31$, $p \leq .01$, one sided), whereas with learner paced presentations, modality effect vanished (post-hoc t-test; $t(51) = -1.03$, $p > .15$, one sided).

Both pacing and text modality had no significant effects on transfer performance (see Table 3), resembling that learning groups did not differ in their mean transfer performance (see Table 4). Accordingly, a pacing by modality effect did not occur either.

Study times differed between treatment instructions. For learner paced instructions, students learned significantly faster with visual than with auditory text (see Table 4; post-hoc t-test; $t(51) = -5.76$, $p < .001$, $h_p^2 = .39$). The mean study times of the system paced instructions lay between these means and were also different (see Table 4; post-hoc t-test; $t(55) = -3.17$, $p < .003$, $h_p^2 = .15$). Variances appeared due to transfer rate of the WWW. The means of the learner paced instructions also differed significantly from the means of the system paced instructions (post-hoc Welch tests; $p < .001$).

Discussion

System versus learner pacing

Pacing had the expected effects on learning with visual texts, but not with auditory texts. Allowing learners to pace the instruction with visual texts reduced the extraneous cognitive load while learning, and increased retention performance, success in labelling pictures, and structural knowledge. Thus, learner pacing proved to be more effective than system pacing of presentations with visual texts. In order to explain this result, it was assumed that learners chose a pace that matched their individual speed of comprehension (inclusive of a compensation of split attention effect). In this experiment, the learners chose a pace that was much faster than the presentation pace installed for the system paced instructions.

For auditory texts, the assumed positive effect of pacing could only be found in respect of extraneous load. In respect of performance, the system paced and learner paced instructions proved to be equally effective for all performance tasks, except the structural task, but it is supposed that the structural knowledge of the system paced group is as good as that of the learner paced group: reporting both more relevant and irrelevant items reflects a comparably bad structural knowledge as reporting both less relevant and irrelevant items does. Additionally, the learner paced instruction was the longest in aspects of study time.

But what does this really mean? Previously, learner pacing meant mainly that study time was extended, i.e. having more time to select relevant information and to organise and integrate it into a mental model (Mayer & Chandler, 2001; Tabbers, 2002; Schmidt-Weigand, 2006). The study durations for instructions in this experiment show that learner pacing led to a prolongation of study time for instructions with auditory texts, but for visual texts, study time was decreased, in comparison to system paced instructions. In other words, although learners studying system paced instructions with visual texts had more time to circumvent spatial split attention, and to extract relevant information, they were outperformed by learners studying self paced instructions with visual texts. The faster average work through of learner paced instructions using visual texts might indicate that learners could more easily compensate spatial split attention and extract relevant information within the system paced presentation periods per information unit, so that the extra study time (system forced presentation time minus time needed for processing) was waiting time that interrupted processing.

Hence, it is proposed that learner pacing has a different status for instructions with visual or auditory texts. For auditory texts, it is proposed that learner pacing means, like before, extended time for learning, which often results in an increased likelihood of organising and integrating information adequately. For visual texts, it is suggested that learner pacing means either extended or decreased time for learning. Extended study time is used for reducing spatial split attention, in order to improve the selection of relevant information and, consecutively, the construction of a mental model. Decreased study time results in a reduction of temporal split attention, in order to increase the likelihood of relevant information that must be mentally integrated being held concurrently in working memory. In general, it is assumed that individually pacing an instruction compensates split attention effects, thus mental effort while learning is diminished, information processing is optimised, and performance is improved.

Pacing and modality effects

Modality effects on overall cognitive load, intrinsic load and extraneous load were not observed. Overall, it would have been very surprising if intrinsic load was affected by modality, as the complexity of the learning material did not vary. Looking at extraneous load, it is not surprising that load reports did not differ between treatment groups, as presentation pace was moderate, and therefore not very load demanding. Especially, visual split attention seems to play no role in learning, because system paced instructions were not very speedy and learner paced instructions could be adjusted in order to prevent it.

As was expected, modality effects on performance were moderated by pacing. Firstly, modality effect was evidenced with system paced presentations, in respect of

structural knowledge and labelling pictures. This accords with Ginns (2005) and a large number of experiments reported in the literature. For system paced instructions, which determine time on presented information, spoken texts are regarded to be more effective, because texts and pictures are perceived simultaneously; there is less danger of missing information or increased cognitive load due to split attention. Learners studying a system paced instruction, based on visual texts, normally suffer from dividing attention between texts and pictures. Contrary to that, above it was suggested that the learners in this experiment were hindered by temporally dividing attention between information units, thus experiencing learning disruption.

Secondly, modality effect on performance vanished for retention and labelling pictures, or even reversed for structural knowledge with learner pacing. This can be attributed to learner pacing compensating for the temporal split attention effect, and allowing a more flexible processing of information, respectively. This result supports the rare evidence of learning situations that allow the modality effect to vanish (Schmidt-Weigand, 2006; Tabbers, 2002, Exp. 3) or even to reverse (Stiller, 2007, Exp. 2; Tabbers, 2002, Exps. 1 and 4). On the other hand, the evidence for modality effects in favour of auditory texts with learner paced instructions is also growing (Atkinson, 2002, Exps. 1 and 2; Brünken et al., 2005, Exps. 1, 2, and 3; Mayer, Dow & Mayer, 2003, Exp. 1; Moreno & Mayer, 2002, Exps. 1 and 2; Moreno, Mayer, Spires & Lester, 2001, Exps. 4 and 5; O'Neil et al., 2000, Exp. 1; Stiller, 2007, Exps. 1, 4, and 5). In the future, there will be more need to determine the circumstances that contribute to vanishing or reversing of modality effects and how these effects may be explained.

Conclusions

Some special limitations of this experiment must be acknowledged. (1) The content covered is relatively limited, scientific, and factual, (2) study time was short, and (3) performance measures were not of the problem solving type. On this background, the learner paced instruction using visual texts was superior to the others. Either the learners performed better or they reached equivalent levels of expertise, but faster. Nevertheless, there could be other contributory factors that should be investigated in future studies like learner characteristics, types of pictorial presentations (dynamic vs. static, realistic vs. diagrammatic etc.), types of instructional systems (e.g. complex interactive systems like simulation systems vs. less interactive presentational systems like animation or film), complexity of content, and instructional goals (e.g. functioning vs. structure of a system). In general, knowledge about the conditions which accompany a superiority of visual over auditory texts in multimedia instructions is deficient. One precondition is that instructions are learner paced, but this is not sufficient to guarantee a superiority of visual texts. If we know the boundary conditions under which one text modality should be preferred over the other, we can more effectively design instructions.

References

- Atkinson, R. K. (2002). Optimizing learning from examples using animated pedagogical agents. *Journal of Educational Psychology, 94*, 416-427.
- Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler* [Statistics for human and social sciences]. Heidelberg: Springer.

- Brünken, R., Plass, J. L. & Leutner, D. (2004). Assessment of cognitive load in multimedia learning with dual task methodology: Auditory load and modality effects. *Instructional Science*, 32, 115-132.
- Brünken, R., Seufert, T. & Zander, S. (2005). Förderung der Kohärenzbildung beim Lernen mit multiplen Repräsentationen [Fostering coherence formation in learning with multiple representations]. *Zeitschrift für Pädagogische Psychologie*, 19, 61-75.
- Brünken, R., Steinbacher, S., Plass, J. L. & Leutner, D. (2002). Assessment of cognitive load in multimedia learning using dual task methodology. *Experimental Psychology*, 49, 109-119.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York, NY: Erlbaum.
- Craig, S. D., Gholson, B. & Driscoll, D. M. (2002). Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features, and redundancy. *Journal of Educational Psychology*, 94, 428-434.
- Dillon, A. & Gabbard, R. (1998). Hypermedia as an educational technology: A review of the quantitative research literature on learner comprehension, control, and style. *Review of Educational Research*, 68, 322-349.
- Ertelt, A., Renkl, A. & Spada, H. (2005). Learning a new computer application using on-screen videos. In A. Méndez-Vilas, B. González-Pereira, J. Mesa González & J. A. Mesa González (Eds), *Recent research developments in learning technologies* (pp. 780-785). Cáceres, Spain: Formatex.
- Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction*, 15, 313-331.
- Hasler, B. S., Kersten, B. & Sweller, J. (2007). Learner control, cognitive load and instructional animation. *Applied Cognitive Psychology*, 21, 713-729.
- Jeung, H. J., Chandler, P. & Sweller, J. (1997). The role of visual indicators in dual sensory mode instruction. *Educational Psychology*, 17, 329-343.
- Kalyuga, S., Chandler, P. & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, 13, 351-372.
- Kalyuga, S., Chandler, P. & Sweller, J. (2000). Incorporating learner experience into the design of multimedia instruction. *Journal of Educational Psychology*, 92, 126-136.
- Lepper, M. R. (1985). Microcomputers in education: Motivational and social issues. *American Psychologist*, 40, 1-18.
- Low, R. & Sweller, J. (2005). The modality principle in multimedia learning. In R. E. Mayer (Ed), *The Cambridge handbook of multimedia learning* (pp. 147-158). New York, NY: Cambridge UP.
- Mayer, R. E. (2001). *Multimedia learning*. New York, NY: Cambridge University Press.
- Mayer, R. E. (2005). Principles for managing essential processing in multimedia learning: Segmenting, pretraining, and modality principles. In R. E. Mayer (Ed), *The Cambridge handbook of multimedia learning* (pp. 169-182). New York, NY: Cambridge University Press.
- Mayer, R. E. & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology*, 93, 390-397.
- Mayer, R. E., Dow, G. & Mayer, S. (2003). Multimedia learning in an interactive self-explaining environment: What works in the design of agent-based microworlds? *Journal of Educational Psychology*, 95, 806-813.

- Mayer, R. E. & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology*, 90, 312-320.
- Mayer, R. E. & Moreno, R. (2002). Aids to computer-based multimedia learning. *Learning and Instruction*, 12, 107-119.
- Moreno, R. (2007). Optimizing learning from animations by minimizing cognitive load: Cognitive and affective consequences of signaling and segmentation methods. *Applied Cognitive Psychology*, 21, 765-781.
- Moreno, R. & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology*, 91, 358-368.
- Moreno, R. & Mayer, R. E. (2002). Learning science in virtual reality multimedia environments: Role of methods and media. *Journal of Educational Psychology*, 94, 598-610.
- Moreno, R., Mayer, R. E., Spires, H. A. & Lester, J. C. (2001). The case for social agency in computer-based multimedia learning: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction*, 19, 177-214.
- Mousavi, S. Y., Low, R. & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology*, 87, 319-334.
- Niemiec, R. P., Sikorski, C. & Walberg, H. J. (1996). Learner-control effects: A review of reviews and a meta-analysis. *Journal of Educational Computing Research*, 15, 157-164.
- O'Neil, H. F., Mayer, R. E., Herl, H. E., Niemi, C., Olin, K. & Thurman, R. A. (2000). Instructional strategies for virtual aviation training environments. In H. F. O'Neil & D. H. Andrew (Eds), *Aircrew training and assessment* (pp. 105-130). Mahwah, NJ: Erlbaum.
- Paas, F. G. & Van Merriënboer, J. J. G. (1993). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human Factors*, 35, 737-743.
- Paas, F. G. & Van Merriënboer, J. J. G. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review*, 6, 351-371.
- Reeves, T. (1993). Pseudoscience in computer-based instruction: The case of learner control research. *Journal of Computer-Based Instruction*, 20, 39-46.
- Rinck, M. & Glowalla, U. (1996). Die multimediale Darstellung quantitativer Daten [Presenting quantitative data using multimedia]. *Zeitschrift für Psychologie*, 204, 383-399.
- Schmidt-Weigand, F. (2006). Dynamic visualizations in multimedia learning: The influence of verbal explanations on visual attention, cognitive load and learning outcome. [viewed 31 Jan 2006, verified 25 Apr 2009]. <http://geb.uni-giessen.de/geb/volltexte/2006/2699/>
- Schwan, S. & Riempp, R. (2004). The cognitive benefits of interactive videos: Learning to tie nautical knots. *Learning and Instruction*, 14, 293-305.
- Stiller, K. D. (2007). *Computerised multimedia learning. Modes of text presentation and access to text*. Hamburg: Verlag Dr. Kovac.
- Sweller, J. (1999). *Instructional design in technical areas*. Camberwell, Australia: ACER Press.
- Tabbers, H. K. (2002). *The modality of text in multimedia instructions. Refining the design guidelines*. Heerlen: Open University of the Netherlands.

Tindall-Ford, S., Chandler, P. & Sweller, J. (1997). When two sensory modes are better than one. *Journal of Experimental Psychology: Applied*, 3(4), 257-287.

Van Gerven, P. W. M. (2002). Efficient complex skills training into old age: Exploring the benefits of cognitive load theory. [viewed 14 Mar 2008]. http://www.personeel.unimaas.nl/p.vangerven/Index_files/Pascal%20van%20Gerven%20-%20Dissertation.pdf

Williams, M. D. (1996). Learner-control and instructional techniques. In D. H. Jonassen (Ed), *Handbook of research for educational communications and technology* (pp. 957-983). New York, NY: Macmillan.

<p>Klaus D. Stiller, Annika Freitag, Peter Zinnbauer and Christian Freitag Institute for Psychology, Technische Universität Dortmund Emil-Figge-Str. 50, 44221 Dortmund, Germany Author for correspondence: Klaus D. Stiller. Email: klaus.stiller@tu-dortmund.de Web: http://www.psych.uni-dortmund.de/</p>
--