APPLIED COGNITIVE PSYCHOLOGY
Appl. Cognit. Psychol. **21**: 713–729 (2007)
Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/acp.1345

Learner Control, Cognitive Load and Instructional Animation

BÉATRICE SUSANNE HASLER[1]*, BERND KERSTEN[2] and JOHN SWELLER[3]

[1]*Department of Psychology, University of Zurich, Zurich, Switzerland*
[2]*Department of Psychology, University of Bern, Bern, Switzerland*
[3]*School of Education, University of New South Wales, Sydney, Australia*

SUMMARY

In order to investigate the influence of learner-controlled pacing in educational animation on instructional efficiency, three versions of an audio-visual computer animation and a narration-only presentation were used to teach primary school students the determinants of day and night. The animations were either system-paced using a continuous animation, learner-paced using discrete segments or learner paced using 'stop' and 'play' buttons. The two learner-paced groups showed higher test performance with relatively lower cognitive load compared to the two system-paced groups, despite the fact that the 'stop' and 'play' buttons were rarely used. The significant group differences regarding test performance were obtained only for more difficult, high element interactivity questions but not for low element interactivity questions. Copyright © 2007 John Wiley & Sons, Ltd.

Dynamic visualisation, such as computer animation, has become increasingly popular in computer-based education across a range of subject domains and education levels (Lowe, 2004). However, animation is frequently used only for its aesthetic or attention gaining effect, and we know little about how animation needs to be designed in order to facilitate learning (Plötzner & Lowe, 2004). Studies that have compared animation and static visuals have provided inconsistent results and failed to support the assumption of a general superiority of animated over static visuals (for a review see Bétrancourt & Tversky, 2000). It appears that under some circumstances, animation can hinder rather than improve learning (Mayer, Hegarty, Mayer, & Campbell, 2005). Animation may impose greater cognitive processing demands than static visuals because information is frequently transient if critical objects and their relations disappear during the animation. When viewing transient animation, learners not only need to integrate new information with knowledge they have stored in long-term memory but also with what they have been presented in earlier parts of the animation because previously presented information does not remain visually available. Mentally integrating the currently presented visual information of a transient animation with previous information may be difficult because of the temporal limitations of working memory. As Tversky, Morrison, and Bétrancourt

*Correspondence to: Béatrice Susanne Hasler, Department of Psychology, University of Zurich, Binzmuehlestrasse 14/27, 8050 Zurich, Switzerland. E-mail: b.hasler@psychologie.unizh.ch

(2002) state, animations often violate the 'apprehension principle' of good graphics according to which 'animations are often too complex or too fast to be accurately perceived' (p. 247). The question arises whether the cognitive load imposed by transient instructional animation can be ameliorated by an appropriate instructional design. This instructional issue is addressed by cognitive load theory (CLT) research.

COGNITIVE LOAD THEORY AND INSTRUCTIONAL COMPUTER-BASED ANIMATION

CLT provides a framework for instructional design based on the structure and function of human cognitive architecture (Clark, Nguyen, & Sweller, 2006; Paas, Renkl, & Sweller, 2003; Sweller, 1988, 1999, 2003, 2004, 2005a; van Merriënboer & Sweller, 2005). Human cognitive architecture may be described as a natural information processing system which can be specified in several ways (Sweller, 2003, 2004; Sweller & Sweller, 2006). In this paper, we will use five basic, underlying principles.

1. *Information store principle*: Natural information processing systems include a very large store of information required to govern the activity of relevant entities. In the case of human cognition, that store consists of long-term memory.
2. *Borrowing and reorganising principle*: Long-term memory is built primarily by imitating other people, listening to what they say or reading what they write. In other words, information is borrowed from the long-term memory of other people. The process involves constructive reorganisation in that new information must be combined with previous information using a constructive process.
3. *Randomness as genesis principle*: The borrowing and reorganising principle does not create new information except insofar as reorganisation results in novelty. Novelty comes from random generation followed by tests of effectiveness during problem solving.
4. *Narrow limits of change principle*: Because of the randomness inherent in both reorganising and random generation and test, large, effective changes to the information store are unlikely. Accordingly, change must be very slow. A limited working memory ensures limited changes to long-term memory (Miller, 1956; Peterson & Peterson, 1959).
5. *Environmental organising and linking principle*: Working memory is only limited when dealing with novel information. When dealing with previously organised information from long-term memory, there are no known limits to working memory. Large amounts of information can be brought into working memory from long-term memory and used to determine activity in the environment (Ericsson & Kintsch, 1995).

This cognitive architecture determines which instructional activities are likely to be effective and which are not. In general terms, the purpose of instruction is to increase knowledge in the information store—long-term memory—and the most effective way of doing so is by directly presenting learners with information via the borrowing and reorganising principle. Once the information is in the information store of long-term memory, it can be used via the environmental organising and linking principle to direct activity. Thus, in the current experiment, animation is used to provide learners with information rather than discover it themselves via the randomness as genesis principle. Following the narrow limits of change principle, that information was structured to reduce working memory load. These principles provide the cognitive base for CLT and can be used to determine which instructional procedures are likely to be effective.

Learning materials can differ substantially in the extent to which they impose cognitive load. Instructional formats that violate the narrow limits of change principle by imposing a high demands on learners' working memory capacity may result in ineffective transfer of information to the long-term memory store due to the unavailability of cognitive resources. Consequently, in accordance with the narrow limits of change principle, CLT aims to reduce unnecessary cognitive load to make more resources available for the learning process (Bannert, 2002).

Types of cognitive load

The cognitive load imposed depends on both the intrinsic nature of the learning material (intrinsic cognitive load) and on how that material is presented (extraneous cognitive load). Intrinsic cognitive load cannot be modified by instructional design (other than through learning) without compromising understanding because it is intrinsic to the material being dealt with. The level of intrinsic cognitive load depends on the number of elements that must be processed simultaneously in working memory in order to understand the instructional material (Pollock, Chandler, & Sweller, 2002; Sweller, van Merriënboer, & Paas, 1998). The elements of low-element interactivity material can be understood and learned serially, without consideration of any other elements because the elements are independent and do not interact. Processing low-element interactivity material does not place high demands on cognitive processing. In contrast, the elements of high-element interactivity material cannot be understood in isolation because the relations between the interacting elements are essential for understanding. For understanding to occur, all interacting elements must be processed simultaneously in working memory (Sweller, 2003).

Extraneous cognitive load refers to unnecessary, ineffective cognitive load that is determined by the way the information is presented (Sweller, 1994). It does not contribute to learning but instead, reduces working memory capacity available for learning activities (Bannert, 2002). In contrast to intrinsic cognitive load, extraneous cognitive load can be altered by instructional design.

According to the 'element interactivity effect' (Chandler & Sweller, 1996; Sweller et al., 1998), instructional techniques to reduce extraneous cognitive load are only relevant when dealing with material that imposes high intrinsic cognitive load. When intrinsic cognitive load is high, a limited working memory becomes critical. If material is low in element interactivity, and therefore, imposes a low intrinsic load, the additional load due to inadequate instructional designs may not overburden working memory capacity and may not hinder learning. It can be expected that cognitive load effects disappear when dealing with low element interactivity material (Sweller, 2003, 2004). However, the level of element interactivity is relative to the level of learners' expertise (Kalyuga, Ayres, Chandler, & Sweller, 2003). For experienced learners, a whole set of interacting elements may be incorporated into a schema and treated as a single element in working memory (Chi, Glaser, & Rees, 1982). Learners may not experience a high cognitive load when dealing with well-learned material that consists of many interacting elements. In contrast, novice learners may need to process each of the elements and their interactions individually. When high element interactivity material is presented in an inappropriate

DOI: 10.1002/acp

instructional format that imposes a high extraneous load, then the learning task is likely to overload learners' working memory resulting in poor performance.

Paas et al. (2003) emphasised a third source of cognitive load that refers to the effort required for the construction of schemas (germane cognitive load). Whereas both a high intrinsic load and high extraneous load interfere with learning, germane cognitive load enhances learning. It occurs when free working memory resources are actively devoted to learning activities. As is the case for extraneous cognitive load, germane cognitive load can be influenced by instructional design. As the three types of cognitive load are additive, the total cognitive load (extraneous load plus intrinsic load plus germane load) imposed by an instructional design must stay within the learner's working memory limits in order to be effective. Therefore, appropriate instructional designs decrease extraneous cognitive load but increase germane cognitive load (Kirschner, 2002). Van Merriënboer (1997) states that instructional designs intended to increase germane cognitive load have to withdraw the learners' attention from processes that are not relevant to learning and direct their attention toward relevant processes.

Instructional effects and design principles

Several instructional design procedures have been derived from CLT based on the cognitive architecture described above and the three types of cognitive load and their interactions. Each procedure assumes that the aim of instruction is to increase information in the long-term memory store in order to permit that information to subsequently govern activity via the environmental organising and linking principle. In order to facilitate the storing of relevant information, the borrowing and reorganising principle should be the governing learning procedure, with the narrow limits of change principle as exemplified by a limited working memory used to reduce extraneous cognitive load.

Only the cognitive load effects relevant for the current study are described below. For a detailed overview the reader may refer to Sweller (2003, 2004).

The split-attention effect occurs when learners are presented with multiple sources of information that are unintelligible in isolation and must be mentally integrated in order to derive meaning from instruction (Ayres & Sweller, 2005; Kalyuga, Chandler, & Sweller, 1999; Mayer & Moreno, 1998). This process is likely to cause a high extraneous load. In learning with computer animation, spatial split-attention is likely to occur when learners have to mentally integrate animated pictures with simultaneously presented on-screen text. Due to the transient nature of information in many instructional animations, temporal split-attention also is likely to occur when learners have to mentally integrate current information with previous information that is no longer available in visual or auditory form. Appropriate instructional designs have to take into account the negative consequences of the split-attention effect, for example, by physically integrating the multiple sources of information (Chandler & Sweller, 1992; Sweller, Chandler, Tierney, & Cooper, 1990), or by using verbal or visual cues (such as colour cueing) to direct the learners' attention to relevant parts of the instruction (Kalyuga et al., 1999). A common technique to avoid split-attention consequences is the use of dual-mode presentation (see the modality effect, below).

The redundancy effect

If simultaneously presented sources of information are fully intelligible in isolation, and one source merely recapitulates the information contained in the other, a high extraneous

DOI: 10.1002/acp

cognitive load may be imposed through the unnecessary processing of redundant material (Kalyuga et al., 1999; Sweller, 2005b; Yeung, Jin, & Sweller, 1998). Redundancy is a crucial issue in multimedia learning because frequently, instructions have not been adapted to learners' expertise in the subject area. Whereas some information might be redundant for one learner, it might be necessary for a less knowledgeable learner.

The modality effect refers to superior instructional efficiency when visualisations are presented with simultaneous narration instead of written text. Dual-mode presentations are a popular instructional technique to overcome the negative consequences of split-attention designs (Moreno & Mayer, 1999; Mousavi, Low, & Sweller, 1995).

Pacing and segmentation

As an instructional technique to overcome the problems associated with high extraneous load due to processing transient information, it has been suggested that learners control the pace of the presentation or that the instructions are divided into meaningful segments (Mayer & Chandler, 2001; Mayer, Dow, & Mayer, 2003; Moreno & Valdez, 2005; Tabbers, Martens, & van Merriënboer, 2004). Pacing enables learners to adapt the pace of the presentation to their individual cognitive needs, for example, by providing them with control options that can be used to pause and continue the instruction at any time. Segmentation allows learners to view a presentation in discrete segments rather than as one continuous presentation, by automatically pausing between logical segments.

Research on pacing and segmentation in multimedia instruction has provided inconsistent results. Whereas some authors have obtained positive effects for learner-controlled pacing and segmentation (Ertelt, Renkl, & Spada, 2005; Mayer & Chandler, 2001; Mayer et al., 2003), others found limitations of the pacing and segmentation effects or even a reverse effect with system-paced instruction being beneficial for learning (Moreno & Valdez, 2005; Tabbers et al., 2004).

Rationale of the experiment

Based on the inconsistent results obtained in multimedia learning regarding pacing and segmentation, the aim of the present study was to investigate the influence of two different forms of learner-controlled pacing during a transient, audio-visual animation on instructional efficiency. An experiment was conducted comparing two learner-paced versions of an animation that either used pre-defined segments or allowed the learner to divide the animation into segments by providing 'stop' and 'play' buttons, with two system-paced versions which presented either a continuous animation, or a continuously presented narration-only version neither of which allowed any form of learner control. The narration-only version served as a control condition to determine whether the animations provided redundant information or whether both sources of information were needed for the material to be fully intelligible.

Superior instructional efficiency, that is, higher test performance with relatively lower cognitive load, was predicted for both the learner-paced groups (segments and stop-play) compared to the system-paced groups (continuous and narration-only). Both the system-paced conditions were expected to impose a high extraneous cognitive load due to the transient nature of the presentation and the lack of learner control. In contrast, learner control, as provided in the stop-play and segments versions of the animation was expected to reduce cognitive load and enhance learning. According to the element interactivity effect, the superiority of the learner-paced groups over the system-paced groups regarding

DOI: 10.1002/acp

test performance should be particularly pronounced in test questions that tap high element interactivity material. Reduced or no test differences were expected on low element interactivity questions.

METHOD

Participants

The participants were 72, male primary school students aged 9 to 11, of a boy's college located in Sydney, Australia. The students were familiar with computer-based education, since the use of laptops is an essential part of their curriculum. They had not yet been taught the subject matter but nevertheless some had prior knowledge of the subject domain. They were randomly assigned to four groups of 18 participants each.

Materials

The materials consisted of four versions of a computer-based presentation that explained the causes of day and night. Three of the versions consisted of narrated animations that differed in their degree of learner control while the fourth was a narration-only version. The participants viewed the computer-based presentation on a laptop with a 17' monitor in full screen mode using Macromedia Flash player, and speakers which were placed on each side of the laptop. The material was designed by the first author using Macromedia Flash software.

A continuous version of the animation and a narration-only version were system-paced. These versions were presented without pauses and without giving the learner an opportunity to control the presentation in any form, apart from replaying the presentation after it ended. The segmented and stop-play versions of the animation were learner-paced and differed in their level of learner control over the pace of the animation. The segmented version allowed students to view the animation as a sequence of eleven pre-defined segments. The animation automatically stopped after each segment, and the students were required to click a 'play' button to continue the animation. The stop-play version provided 'stop' and 'play' buttons to pause the animation at any time, and turn it into a still image, which permitted them to view the animation at their own pace.

The animation was designed with simultaneous audio narration, spoken by a male voice. Simultaneous narration was used in order to maximise the modality effect and avoid the negative consequences associated with 'split-attention' designs. The audio commentary was identical for each of the four versions. The information was transiently presented in the animation, that is, the animated elements appeared, and (at least partly) disappeared over time. The audio component, of course, was transient as well.

The animation consisted of 2260 frames and was presented at a rate of 10 frames per second. In the segments version of the animation, the mean number of words in each of the 11 segments was 47.1 ($SD = 17.4$). The mean time to play one segment was 19.1 seconds ($SD = 6.8$), and the mean speed of the narration was 2.5 words per second ($SD = 0.34$). The length of segments was chosen at around 20 seconds each to take the duration limitation of working memory into account. According to Peterson and Peterson (1959), information decays within about 20 seconds without rehearsal. It took 3 minutes and 45 seconds to play the entire presentation in the system-paced versions. For all four groups the total duration of the learning phase was held constant at 10 minutes to ensure that any group differences

DOI: 10.1002/acp

were not due to different study period lengths. Within the 10-minute period each time the presentation stopped, the learner-control students were asked to restart it. After 10 minutes, the presentation stopped automatically. At each point during this learning phase the available time left of the 10 minutes was displayed on the screen for all groups.

The content (11 segments) of the learning material contained both low and high element interactivity information. The first segment depicted the sun rising in the east and setting in the west. The following two segments provided low element interactivity information about the sun and the earth. The fourth segment introduced the axis between the earth's north and south poles. The remaining segments contained high element interactivity material depicting the daily changes between day and night and explained why it is day on one side of the earth while the opposite side has night and why we see the sun rising in the east and setting in the west. Appendix A gives an example of a segment low in element interactivity (segment 3) and one high in element interactivity (segment 10).

To test learning, a set of 13 questions was constructed to include both low and high element interactivity. These questions had the highest discrimination power of a set of 24 tested in a pilot study. According to Sweller et al. (1998), element interactivity can be determined by counting the number of interacting elements that a person at a particular level of expertise has to process..Questions tapping low element interactivity knowledge are assumed to require the simultaneous processing of three or fewer elements while those tapping high element interactivity knowledge are assumed to require the processing of four or more interacting elements. Based on these criteria, questions 1, 2, 3, 5, 6, 7, 8, and 9 were classed prior to the experiment as tapping low element interactivity material while questions 4, 10, 11, 12, and 13 were classed as high in element interactivity (see Appendix B). All 13 questions were presented on screen individually. Students were able to move sequentially through the question set by clicking the 'next question' button.

Paper answer booklets were provided where students were to write their answers to the questions, which were displayed on the screen. As some of the later questions contained information that could be used to answer earlier ones, the answer booklets contained separate sheets for each question and did not display the text but only the number of the question. The program did not allow students to return to previously presented questions.

To measure cognitive load, a subjective rating scale was provided on the first page of the students' answer booklets. The participants were asked how hard or easy it was for them to learn the material, and rated their subjectively experienced difficulty on a nine-point rating scale ranging from 1 'extremely easy' to 9 'extremely difficult'. Nine-point rating scales have been used successfully in other studies (Kalyuga, Chandler, & Sweller, 1998; Marcus, Cooper, & Sweller, 1996; Tindall-Ford, Chandler, & Sweller, 1997). Uni-dimensional scales, such as retrospective difficulty ratings, are a popular subjective cognitive load measurement technique because they are easy to use and do not interfere with the learning task (Paas, van Merriënboer, & Adam, 1994).

Design

A 4×2 design with four groups (continuous animation, segmented animation, stop-play animation, and narration-only presentation) and two levels of test performance (high element interactivity questions vs. low element interactivity questions) with repeated measures on the second variable was used. The four instructional formats were compared regarding test performance and mental load as the dependent variables. In order to determine the instructional efficiency of the four different instructional formats, relative

DOI: 10.1002/acp

efficiency scores were calculated as a joint function of mental load and performance measures, using Paas and van Merriënboer's (1993) technique. This approach is based on the z-transformation of raw cognitive load values and raw performance scores. Instructional efficiency (E) was determined using the formula $E = (z_{Performance} - z_{MentalLoad})/\sqrt{2}$.

Procedure

The experiment consisted of a learning phase followed by a test phase. The participants were tested individually. They were informed using standardised instructions that the investigator aimed to evaluate how primary school students learn with multimedia, and that they would be presented some information on the solar system. Subsequently, the participants were asked to answer two multiple-choice questions to assess whether they were familiar with the subject domain.

After the students completed the prior knowledge test, the narration-only and continuous groups were instructed to carefully listen to the content of the presentation, or carefully study the animation, respectively, in order to be able to answer questions later. In addition to this information, the stop-play and segments groups were instructed how and when to use the control options during the presentation. The segments group was informed that the presentation was divided into segments and would automatically stop after each segment and that a 'play' button would appear. They were further instructed that when the presentation stopped after a segment, to think of what they had seen in the previous segment and to click the 'play' button when they were ready to continue the animation. The stop-play group was told to use the 'stop' and 'play' buttons to pause the presentation when they needed to think of what they had previously seen, and that they were allowed to stop the animation at any time for as long as they wanted.

Directly following the learning phase, the instructional efficiency of the different versions of the material was measured. First, the subjects rated their subjective mental load, and second, they completed the 13 test questions.

RESULTS

Pause times between segments

The 18 students who were assigned to the segments group paused the animation on an average for 2.40 seconds ($SD = 1.62$) the first time they played the animation, and for a mean time of 2.07 seconds ($SD = 2.03$) the second time. The longest time they paused the animation was after the first segment ($M = 5.37$ seconds, $SD = 2.56$), and after the last segment ($M = 6.23$ seconds, $SD = 2.62$) the first time they played the animation. Ten students had to be told to use the 'play' button to continue the animation after the first segment, and seven had to be told to click the 'start' button to restart the animation when it ended. Questions regarding the content of the animation were not answered during the experiment.

Usage of the 'stop' and 'play' buttons

In the stop-play group, 11 out of 18 students did not use the 'stop' and 'play' buttons at all while studying the animation. Six students stopped the animation once to ask the

DOI: 10.1002/acp

experimenter a question. Three students asked a question at the beginning of the presentation to clarify if they had understood the instruction about what they were required to do. The remaining three students asked questions concerning the content of the animation which were not answered during the experiment. Only one student used the 'stop' and 'play' buttons four times. When he stopped the animation he merely commented on what he had just been presented (for example, 'This is where the sun is at lunchtime'.).

Calculation of test scores and difficulty indices

Two independent raters marked the students' answers to the open test questions. One point was given for correct answers, half a point for a correct idea, and zero points for wrong answers. The inter-rater reliability indicated $r = 0.94$. To examine the internal consistency of all 13 test items, Cronbach's Alpha was calculated, and indicated 0.64. The internal consistency of the test questions appears satisfactory when the small number of test questions and students is taken into account. As might be expected the low element interactivity questions were found to be easier than the high element interactivity questions in this experiment. The difficulty indices as well as the level of element interactivity for each test question are depicted in Appendix B. The difficulty index was calculated using the percentage of right answers on each question, that is, a difficulty index of 1.0 would indicate that 100 per cent of the students correctly answered the question, whereas a difficulty index of 0 would indicate that none of the students knew the correct answer. Therefore, the higher the difficulty index, the easier was the question, and *vice versa*. A Spearman correlation indicated a significant correlation ($r_s = 0.59$, $p = 0.02$) between the difficulty index and level of element interactivity.

Test performance

It was predicted that both the learner-paced groups (segments and stop-play) would lead to higher instructional efficiency compared to both the system-paced groups (continuous animation and narration-only) on a subsequent knowledge test. Mean group scores for the overall test, and the high and low element interactivity questions based on the means of the two raters' scores, as well as the scores of the subjective difficulty ratings for each group are presented in Table 1.

Based on the element interactivity effect (Chandler & Sweller, 1996; Sweller et al., 1998), it was predicted that the segments and stop-play groups would perform better on high element interactivity questions than the continuous animation and narration-only groups. No differences were expected between the groups on low element interactivity questions. A 4 (groups) $\times$ 2 (levels of element interactivity) ANOVA test with repeated measures on the second variable revealed a significant main effect of test performance for groups, $F(3, 68) = 4.71$, $MSE = 11.28$, $p = 0.01$, $\eta^2 = 0.17$, and a significant main effect for element interactivity, $F(1, 68) = 366.70$, $MSE = 298.28$, $p < 0.001$, $\eta^2 = 0.84$. The interaction effect for group $\times$ element interactivity failed to reach significance, $F(3, 68) = 2.23$, $MSE = 1.81$, $p = 0.09$. Nevertheless, one-way ANOVA tests revealed a significant difference between groups on the high element interactivity questions, $F(3, 68) = 9.09$, $MSE = 10.31$, $p < 0.001$, $\eta^2 = 0.29$, but no statistically significant difference between groups on the low element interactivity questions, $F(3, 68) = 1.34$, $MSE = 2.78$, $p = 0.27$. A Tukey post hoc test indicated that the significant group differences regarding the high element interactivity questions were due only to a significant difference between

DOI: 10.1002/acp

Table 1. Means and standard deviations of test scores and subjective difficulty ratings

		Test scores							
		Overall test ($k^* = 13$)		Low element interactivity questions ($k^* = 8$)		High element interactivity Questions ($k^* = 5$)		Cognitive load (subjective difficulty rating)	
Group	*N*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Stop-play	18	9.03	2.46	5.90	1.65	3.13	1.08	2.78	1.44
Segments	18	8.69	2.32	5.49	1.48	3.20	1.16	2.61	1.58
Continuous	18	6.96	1.94	5.06	1.33	1.90	1.00	3.22	1.06
Narration	18	6.92	1.99	5.10	1.27	1.82	1.04	3.22	1.44
Total	72	7.90	2.35	5.39	1.45	2.51	1.23	2.96	1.39

$^*k$ = number of questions.

the learner-paced and the system-paced groups. The performance of the narration-only group and the continuous animation group did not significantly differ. Likewise, no significant differences regarding the test performance was found between the stop-play and the segments groups.

In order to explore if the superior test performance of the stop-play group was because some of the students stopped the animation at least once, two-tailed independent samples *t*-tests between the scores of the students who stopped the animation ('stop' group) and the scores of the students who did not stop the animation ('no stop' group) was conducted. The total test scores of the 'stop' group ($M = 9.77$, $SD = 1.98$) did not significantly differ from the total test scores of the 'no stop' group ($M = 8.57$, $SD = 2.71$), $t(16) = 1.01$, $p = 0.33$.

Mental load ratings

The differences regarding the reported cognitive load (i.e. subjective difficulty ratings) did not reveal a statistically significant result on a one-way ANOVA test, $F(3, 68) = 0.91$, $MSE = 1.76$, $p = 0.44$. This failure to find a significant effect regarding mental load ratings may be due to the fact that the learning material was reported as relatively easy by all experimental groups (mean score was 2.96 on a 9-point scale, $SD = 1.39$).

Instructional efficiency

Figure 1 shows the instructional efficiency scores obtained for the four experimental groups in high and low element interactivity questions.

A 4 (groups) × 2 (level of element interactivity) ANOVA test with repeated measures on the second variable revealed a significant main effect of instructional efficiency for groups, $F(3, 68) = 3.90$, $MSE = 7.56$, $p = 0.01$, $\eta^2 = 0.15$, and a significant interaction effect for group × element interactivity, $F(3, 68) = 3.31$, $MSE = 0.70$, $p = 0.03$, $\eta^2 = 0.13$. The main effect of element interactivity on the standardised learning efficiency scores (mean equals 0) was not, of course, significant. A one-way ANOVA test revealed a highly significant difference in instructional efficiency of the high element interactivity questions between the four versions of the presentation, $F(3, 68) = 6.18$, $MSE = 6.28$, $p = 0.001$, $\eta^2 = 0.21$. A

DOI: 10.1002/acp

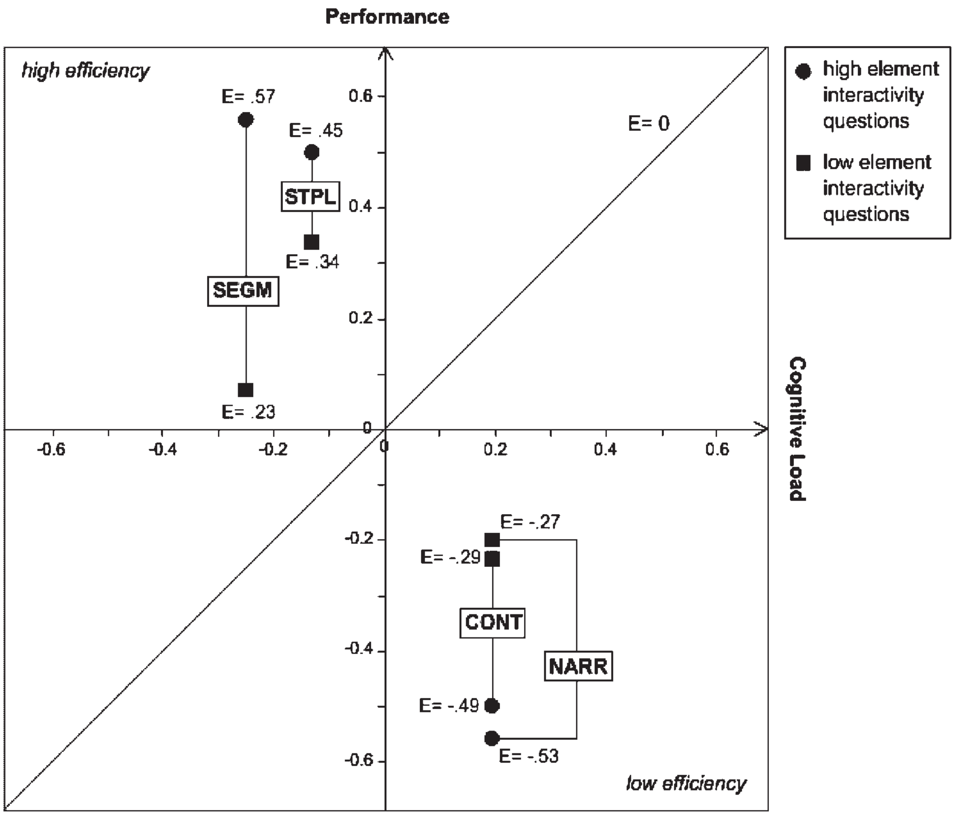


Figure 1. Instructional efficiency scores of the four instructional formats

Tukey post hoc test revealed that the significant difference regarding instructional efficiency between the groups was caused by a significant difference between the system-paced and the learner-paced groups. Again, the stop-play and segments groups did not differ significantly regarding their instructional efficiency but were more efficient as indicated by better learning performance with lower cognitive load respectively than the continuous and narration-only presentations which also did not differ significantly. A one-way ANOVA test of instructional efficiency on the low element interactivity questions revealed, as expected, no significant difference between the four versions of the presentation, $F(3, 68) = 1.75$, $MSE = 1.98$, $p = 0.17$. The ANOVAS of the efficiency scores yielded a similar pattern of results to those obtained on performance due to the fact that while the mental loads did not significantly differ between the groups, they were in the predicted direction.

DISCUSSION

The experiment described in this paper is based on the cognitive architecture intrinsic to CLT. That architecture indicates that effective instruction results in changes in long-term

DOI: 10.1002/acp

memory (information store principle) that then can be used to determine action via the environmental organising and linking principle. Changes in long-term memory are likely to be initiated by the borrowing and reorganising rather than the randomness as genesis principle and so learners should be presented information rather than asked to explore a problem space. Accordingly, in the current experiment, learners were presented information on the causes of day and night rather than being asked to search for and explore possible causes. Furthermore, once presentation rather than exploration is used, the issue of how information should be presented becomes critical with the need to consider working memory limits via the narrow limits of change principle. Ultimately, the manipulations of the current experiment were driven by these considerations.

The superior test and efficiency performance of the learner-paced animation groups compared to the narration-only group indicates that the animated pictures and the simultaneous narration do not provide redundant information imposing an extraneous working memory load. Both sources of information are necessary for more understanding to occur. Therefore, possible negative consequences of a redundancy effect can be excluded for the audio-visual animation in this study. The finding that the students who were presented the continuous version of the animation performed no better than the narration-only group indicates that animation without learner control is as effective as no animation in this domain. In addition, the analysis of learning outcomes regarding the level of element interactivity showed that providing simple forms of learner control when viewing transient animation might lead to better test performance for high element interactivity material but not necessarily for low element interactivity material. This result provides empirical evidence for the 'element interactivity' effect derived from CLT. Furthermore, this result is in line with Mayer and Chandler's (2001) findings of superior test performance of a segments group compared to a continuous group only on transfer but not retention tasks. These results provide empirical support for the assumption that learner-controlled pacing of an animated presentation is beneficial for higher-level schema construction.

Despite rarely using the 'stop' and 'play' buttons, the stop-play group performed better on high element interactivity questions than the continuous group that did not pause during the animation, and equally well as the segments group that did pause. The superiority of the stop-play group over the continuous group occurred even though the stop-play condition appeared not to differ from the continuous condition because most participants in the stop-play group also viewed the animation continuously. Only one out of 18 students in the stop-play group used the 'stop' and 'play' buttons in a regular and purposeful fashion. The question arises why the other students decided not to actively make use of the learner control option. Learners may have been unsure at what point it was appropriate to stop the animation. They were novices in this domain and at any given point did not know what information was following. In the absence of instructional information indicating when to stop, learners had to guess an appropriate stopping point. They had no choice but to use the randomness as genesis rather than the borrowing and reorganising principle. Rather than deciding to stop at what may have been a random point, almost all learners chose not to stop at all. Nevertheless, while the uncertainty of the randomness as genesis principle under, for example, discovery learning conditions, may result in a high extraneous cognitive load as learners search for possible problem solutions (Kirschner, Sweller, & Clark, 2006), in the current experiment, learners were not searching for problem solutions. They were searching for a possible stopping point and that search may have increased

DOI: 10.1002/acp

germane cognitive load as cognitive resources were devoted to the expository instructional material.

The participants in the stop-play group were instructed to learn the material and use the stop-button whenever they felt that they needed to stop the presentation in order to think of what they had seen previously, whereas the continuous group was simply told to learn the material. Learners may closely monitor the information in order to decide whether they need to think about it further by pausing the animation. They must also closely monitor where the animation might be paused. That monitoring may assist in schema acquisition whether or not learners pause the animation. The deeper cognitive processing of the instructional information in terms of a higher germane cognitive load is likely to result in better learning performance. In contrast, the continuous animation group may have treated the animation as little more than a movie and given it very little thought. Similar results were obtained by Moreno and Valdez (2005) who found that learner-controlled pacing was only effective when students were asked to evaluate their understanding before receiving corrective feedback from the system.

In summary, the results obtained in this study provide empirical support for the validity and successful application of principles derived from CLT in the context of multimedia learning. Furthermore, the findings of the experiment have a practical impact for instructional designers: Learner control, either in the form of pre-defined segments or by allowing the learners to pause the animation at any time, should be integrated in educational animation in order to improve instructional efficiency. However, the length of the segments, as well as the speed of the entire animation may play a critical role and require further investigation.

ACKNOWLEDGEMENTS

The authors wish to thank Graeme Watson, Headmaster at the Preparatory School of the Scots College, Sydney, and staff and students who participated in this experiment. Furthermore, we would like to thank Prof. Steffen-Peter Ballstaedt for his comments on an earlier version of the manuscript.

REFERENCES

Ayres, P., & Sweller, J. (2005). The split-attention principle in multimedia learning. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (pp. 135–146). New York: Cambridge University Press.

Bannert, M. (2002). Managing cognitive load—recent trends in cognitive load theory. *Learning and Instruction*, *12*, 139–146.

Bétrancourt, M., & Tversky, B. (2000). Effect of computer animation on users'performance: A review. *Le travail Humain*, *63*, 311–330.

Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, *62*, 233–246.

Chandler, P., & Sweller, J. (1996). Cognitive load while learning to use a computer program. *Applied Cognitive Psychology*, *10*, 151–170.

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 7–75). Hillsdale, NJ: Erlbanm.

Clark, R., Nguyen, F., & Sweller, J. (2006). *Efficiency in learning: Evidence-based guidelines to manage cognitive load*. San Fransisco: Pfeiffer.

DOI: 10.1002/acp

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*, 211–245.
Ertelt, A., Renkl, A., & Spada, H. (2005). Learning a new computer application using on-screen videos. In A. Méndez-Vilas, B. Gonzalez Pereira, J. Mesa González, & J. A. Mesa González (Eds.), *Recent research developments in learning technologies* (pp. 780–785). Cáceres, Spain: Formatex.
Kalyuga, S., Ayres, P., Chandler, P., Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, *38*, 23–31.
Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors*, *40*, 1–17.
Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, *13*, 351–371.
Kirschner, P. A. (2002). Cognitive load theory: Implications of cognitive load theory on the design of learning. *Learning and Instruction*, *12*, 1–10.
Kirschner, P., Sweller, J., & Clark, R. (2006). Why minimal guidance does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential and inquiry-based teaching. *Educational Psychologist*, *41*, 75–86.
Lowe, R. K. (2004). Interrogation of a dynamic visualization during learning. *Learning and Instruction*, *14*, 257–274.
Marcus, N., Cooper, M., & Sweller, J. (1996). Understanding instructions. *Journal of Educational Psychology*, *88*, 49–63.
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. T., & 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., Hegarty, M., Mayer, S., & Campbell, J. (2005). When static media promote active learning: Annotated illustrations versus narrated animations in multimedia instructions. *Journal of Experimental Psychology: Applied*, *11*, 256–265.
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.
Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.
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., & Valdez, A. (2005). Cognitive load and learning effects of having students organize pictures and words in multimedia environments: The role of student interactivity and feedback. *Educational Technology Research and Development*, *53*, 35–45.
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.
Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, *38*, 1–4.
Paas, F., & 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., van Merriënboer, J. J. G., & Adam, J. J. (1994). Measurement of cognitive load in instructional research. *Perceptual and Motor Skills*, *79*, 419–430.
Peterson, L., & Peterson, M. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology*, *58*, 193–198.
Plötzner, R., & Lowe, R. (2004). Dynamic visualisations and learning. *Learning and Instruction*, *14*, 235–240.
Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. *Learning and Instruction*, *12*, 61–86.
Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*, 257–285.
Sweller, J. (1994). Cognitive load theory, learning difficulty and instructional design. *Learning and Instruction*, *4*, 295–312.

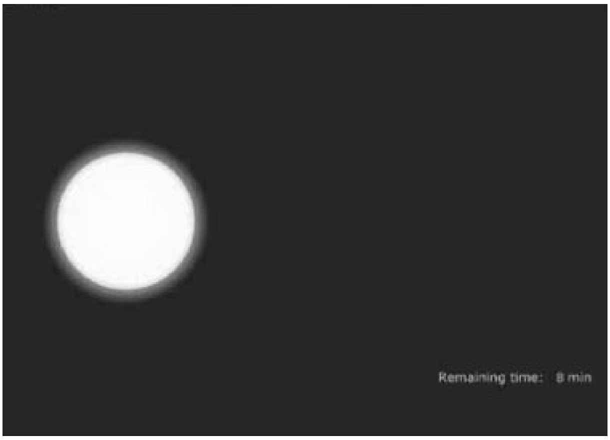
Sweller, J. (1999). *Instructional designs in technical areas*. Melbourne: ACER Press.
Sweller, J. (2003). Evolution of human cognitive architecture. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 43, pp. 215–266). San Diego: Academic Press.
Sweller, J. (2004). Instructional design consequences of an analogy between evolution by natural selection and human cognitive architecture. *Instructional Science*, *32*, 9–31.
Sweller, J. (2005a). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 19–30). New York: Cambridge University Press.
Sweller, J. (2005b). The redundancy principle. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 159–167). New York: Cambridge University Press.
Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load and selective attention as factors in the structuring of technical material. *Journal of Educational Psychology: General*, *119*, 176–192.
Sweller, J., & Sweller, S. (2006). Natural information processing systems. *Evolutionary Psychology*, *4*, 434–458.
Sweller, J., van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*, 251–296.
Tabbers, H., Martens, R., & van Merriënboer, J. J. G. (2004). Multimedia instructions and cognitive load theory: Effects of modality and cueing. *British Journal of Educational Psychology*, *74*, 71–81.
Tindall-Ford, S., Chandler, P., & Sweller, J. (1997). When two sensory modes are better than one. *Journal of Experimental Psychology: Applied*, *3*, 257–287.
Tversky, B., Morrison, J. B., & Bétrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human-Computer Studies*, *57*, 247–262.
van Merriënboer, J. J. G. (1997). *Training complex cognitive skills: A four-component instructional design model for technical training*. Englewood Cliffs, NJ: Educational Technology Publications.
van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, *17*, 147–177.
Yeung, A., Jin, P., & Sweller, J. (1998). Cognitive load and learner expertise: Split-attention and redundancy effects in reading with explanatory notes. *Contemporary Educational Psychology*, *23*, 1–21.

DOI: 10.1002/acp

APPENDIX

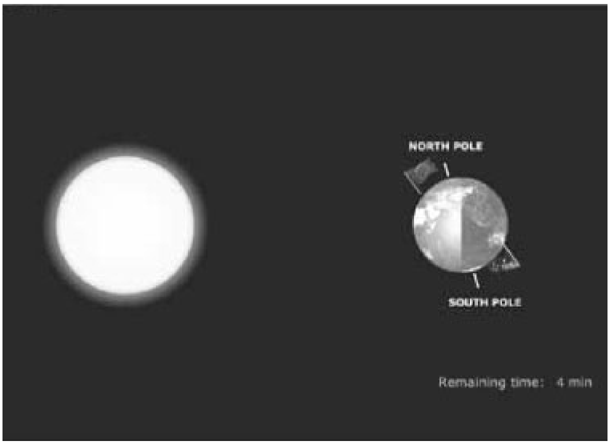
A. Examples for high and low element interactivity materials taken from the animated versions of the presentation

Segment 3 (low element interactivity material)



The sun is a star. It seems bigger and hotter than the other stars, but this is because it is closer to the earth than all the other stars. The sun gives us energy. Without the sun, life on earth would
not exist.

Segment 10 (high element interactivity material)



When it is daytime in Australia it is night-time on the other side of the world—for example, in Europe. When the sun rises in Australia it sets in Europe and it gets dark in Australia when people in Europe have sunrise.

DOI: 10.1002/acp

B. Test questions

		Level of element interactivity	Difficulty index (0 = most difficult, 1 = easiest)
1	What causes the sun to appear to move across the sky?	Low	0.66
2	Why can we not feel any movement when the earth is spinning?	Low	0.73
3	How many times does the earth turn around its axis in 1 week?	Low	0.73
4	If the earth were to rotate twice as fast around its own axis, how long would it take for one full rotation?	High	0.68
5	Does the earth spin from east to west or west to east?	Low	0.43
6	What causes the daily change between day and night?	Low	0.68
7	Where is the sun when it is midnight in Sydney? Add the sun to the graphic on your answer sheet.	Low	0.77
8	If it is 8 O'clock in the morning in Sydney, what time is it on the opposite side of the earth?	Low	0.75
9	Explain why Australia experiences daytime while Europe experiences night?	Low	0.65
10	Why does the sun rise in the east and set in the west?	High	0.33
11	If the earth were to spin from east to west, would the sun rise in the east or in the west?	High	0.53
12	Why does the sun rise and set first in Eastern Australia and later in Western Australia?	High	0.39
13	London is west of Paris. In which town does the sun rise first? Can you explain why?	High	0.58

DOI: 10.1002/acp