Instr Sci (2010) 38:471–485
DOI 10.1007/s11251-009-9117-6

Learning from animation enabled by collaboration

Cyril Rebetez · Mireille Bétrancourt · Mirweis Sangin · Pierre Dillenbourg

Published online: 18 November 2009

Abstract Animated graphics are extensively used in multimedia instructions explaining how natural or artificial dynamic systems work. As animation directly depicts spatial changes over time, it is legitimate to believe that animated graphics will improve comprehension over static graphics. However, the research failed to find clear evidence in favour of animation. Animation may also be used to promote interactions in computer-supported collaborative learning. In this setting as well, the empirical studies have not confirmed the benefits that one could intuitively expect from the use of animation. One explanation is that multimedia, including animated graphics, challenges human processing capacities, and in particular imposes a substantial working memory load. We designed an experimental study involving three between-subjects factors: the type of multimedia instruction (with static or animated graphics), the presence of snapshots of critical steps of the system (with or without snapshots) and the learning setting (individual or collaborative). The findings indicate that animation was overall beneficial to retention, while for transfer, only learners studying collaboratively benefited from animated over static graphics. Contrary to our expectations, the snapshots were marginally beneficial to learners studying individually and significantly detrimental to learners studying in dyads. The results are discussed within the multimedia comprehension framework in order to propose the conditions under which animation can benefit to learning.

Keywords Multimedia · Animation · Learning · Collaboration · Dynamic mental model

C. Rebetez (✉) · M. Bétrancourt
Tecfa, Faculty of Psychology and Educational Sciences, University of Geneva,
Acacias 54, 1227 Carouge, Switzerland
e-mail: Cyril.rebetez@unige.ch

M. Sangin · P. Dillenbourg
Craft, Federal Institute of Technology of Lausanne, Lausanne, Switzerland

Introduction

Given their dynamic nature, animated graphics are often considered as ideally suited for conveying explanations about dynamic phenomena that involve change over time, like mechanical, biological, or meteorological systems. As evidenced by the popularity of computer animations in web sites providing teaching resources, this assumption is mostly accepted by educational practitioners. However, research did not provide consistent evidence that animation improves the understanding of dynamic phenomena. In this paper, we claim that animation can help the understanding of dynamic systems, provided that the instructional material and the learning setting are designed in order to overcome the cognitive limitations of animation. Before examining when and why animation may support comprehension, it is necessary to consider the cognitive processes involved when using multimedia instruction for understanding how dynamic systems work.

Models of multimedia comprehension

In this paper we define multimedia instruction as instruction containing both symbolic (verbal information, formula, etc.) and analogical information (graphics, pictures, schemas, etc.). Originally derived from the domain of text comprehension, models of multimedia comprehension assume that verbal and pictorial information are first processed separately before being integrated in a common representation, or mental model. Schnotz and Bannert (2003) proposed that multimedia information is processed in two distinct paths, symbolic and analogical. In the symbolic path, semantic processing is applied to verbal information and leads to a propositional representation. On the analogical path, visual information is first organized according to perceptive rules into a visual image. The propositional representation and the visual image are then integrated into a unique analogical structure, the mental model.

Derived from multimedia information and from previous knowledge, the mental model makes it possible to generate new information and inferences during and after reading the instruction. Mayer's *selection, organisation and integration model* (Mayer 2001, 2005) also postulates a similar combination of ascending and descending processes and a dual processing principle. Information is first selected in two distinct sensory channels according to whether information is conveyed in a visual or auditory mode. Verbal and visual elements are then organized in two distinct representations, irrespective of their initial sensory mode, that are finally integrated in a single mental model. Mayer's model particularly focuses on the assumption that working memory capacity is limited for novel elements (Baddeley 1986). A large body of research using this model established principles for designing multimedia instruction in order to avoid cognitive overload at different stages of processing (Mayer 2005; Mayer and Moreno 2003).

A specific case is the construction of mental models of dynamic systems, which involve spatial changes over time. According to Narayanan and Hegarty (2002), constructing a dynamic mental model is a five-step process. A first organisation leads to static mental models (one verbal and one visual), followed by the identification of referential links between modalities. The dynamic mental model is achieved after the identification of cause-effect relationships and a final integration. It is then possible to mentally simulate the system in motion, and therefore to infer its functioning under different conditions.

According to the mental model theories reviewed above, it is legitimate to believe that animations are ideally suited to support the construction of a dynamic mental model, since spatio-temporal relations in the mental model can be directly mapped to spatio-temporal changes in the display, saving the learner from engaging in a cognitively demanding mental simulation. However, the literature does not support this assumption. Considering only the studies that involved carefully designed experiments comparing animated and static graphics, Tversky et al. (2002) reported that most studies found little if no benefit from animated graphics over static ones on learning outcomes. Several studies, designed in agreement with principles derived from multimedia learning theories, suggest that animated graphics may not improve learning compared to their static equivalents (e.g., Catrambone and Seay 2002; Hegarty et al. 2003; Lowe 2003). For example, Scheiter et al. (2006) used an hypermedia environment to teach probability theory and showed that the frequent use of animations led to higher learning times and a decrease in performance compared to a static-pictures condition. A recent meta-analysis (Höffler and Leutner 2007) using the results of 26 studies comparing static and animated visualizations found that animation was beneficial when the learning content involved procedural knowledge, but not when it involves declarative knowledge.

Why and when animation helps understanding dynamic systems

Schnotz and Rasch (2008) develop the idea of three possible effects of animated pictures on the learner. In the first, *facilitating effect*, animations can facilitate the construction of a dynamic mental model, mainly by preventing the learners from having to engage in demanding mental simulation. The second effect, called *enabling*, refers to the potential animated graphics have for allowing the comprehension of dynamic systems that novice learners are unable to mentally simulate. According to Tversky et al. (2002), the main benefit of animated over static graphics is to convey the microsteps between larger steps, specifically the precise spatial-temporal actions of components. Many of the static graphic displays portray the coarse segments whereas the animations portray both the coarse and fine segments. Without the microsteps, novice learners would simply not be able to form an accurate dynamic mental model. Moreover, since animation shows the microsteps of the dynamics process, learners do not have to mentally infer how the system functions. Cognitive resources can thus be allocated to processing the verbal and pictorial information, ensuring a better memorization. The third effect described by Schnotz and Rasch (2008) is negative to learning and is called *inhibiting effect*. As change over time is directly perceived from animation and does not have to be inferred, the use of animated pictures can inhibit the learners from mentally animating the dynamic phenomenon, leading to a shallow processing of information. Though animation directly conveys the succession of steps and transformations, the underlying conceptual dynamics still required active integration processes by the learner. Similarly, Lowe (2004) described the *underwhelming effect as* the risk that the direct visualization of dynamic information in animation may induce an illusion of understanding and consequently lead to a cognitive involvement withdrawal. This concern agrees well with Mayer's (2005) focus on the need of learner to be active, also underlined by Ainsworth and Van Labeke (2004).

Another drawback of animation mentioned in the literature deals with the delivery of dynamic information. As animation involves change over time, information is transient and cannot be reinspected, contrary to a series of static pictures (Tversky et al. 2002). As a result, some critical information can be missed or inaccurately perceived, especially because novice learners tend to focus on perceptually salient changes that are not necessarily conceptually relevant (Lowe 2004). Moreover, animation imposes a heavy working memory load, since each and every change needs to be memorized (initial point, type of change, inter-dependency, etc.). According to the cognitive load theory (Sweller 2003), the cognitive demand due to the processing of the instructional material (extraneous load) should be kept to a minimum in order to save resources for the construction of the mental schema (germane load) and for the processing of the content (intrinsic load).

Given these cognitive difficulties imposed by animation, we propose to use recapitulative snapshots of critical steps of the dynamic phenomenon depicted in the animation, in order to offload working memory and to support the construction of a dynamic mental model. Since a graphic representation of the critical steps remains on the screen, the initial and final position of objects can be reinspected. The learners can adopt a "piecemeal strategy" (Hegarty and Sims 1994) and construct first local representations that are then integrated in a coherent mental model. Catrambone and Seay (2002) found that providing frames of critical steps of a dynamic process was as effective as animation to explain the functioning of computer algorithms, except for complex transfer problems where animation was slightly more effective.

In a preliminary study, Bétrancourt et al. (2003) used a translated version of Mayer and Chandler's (2001) instructional material about the formation of lightning. In one condition, eight snapshots of the main steps of the process were displayed above the commented animation. Participants studying with the animated version outperformed participants studying with the static version. But participants studying the material with snapshots did not perform differently from participants without the snapshots. However, the relative low complexity and absence of concurrent changes in the lightning instructional material might be responsible for the lack of effects. We assume that snapshots can still be useful when visual changes are concurrent and sufficiently complex to overwhelm learner's attention and working memory capacity.

Individual and collaborative learning situation

The collaborative learning setting may be a good candidate to overcome the *underwhelming effect*, a major drawback of animation described by Lowe (2004). The assumption underlying collaborative learning theories is that learning is improved through learners' efforts to integrate their different perspectives. Learners studying collaboratively have to negotiate meanings, share and compare their points of view and construct common knowledge. A common representation of the learning task has to be constructed and maintained all along the task, by a process called grounding (Clark and Brennan 1991; Roschelle and Teasley 1995). Dillenbourg (1999) claimed that the benefit of collaboration for learning is a "side-effect" of the cognitive involvement needed to build and maintain a shared representation of the problem at hand. The grounding process can be supported through *artefacts*, external elements used as symbols or information sources to create an array of accessible knowledge around an individual and his peers (Moore and Rocklin 1998). The research in computer-supported collaborative learning (CSCL) focuses mainly on designing computerized artefacts that support the elaboration and updating of the common representation.

Collaborative learning has rarely been used in multimedia learning research. Schnotz et al. (1999) designed two experiments involving either an individual or a collaborative setting (dyads of participants). The instructional material was either an interactive animation or a static representation that explains time zones on earth. In the individual learning setting, students using these animations showed better detail encoding than the ones using static pictures, but no difference was found on a mental simulation task. In the collaborative setting, participants achieved poorer performance when learning from the dynamic material than from the static material (for both detail encoding and mental simulation tasks). The authors suggested that participants in dyads experienced a cognitive overload with the materials used since they had to handle both the collaboration and the information flow of the animation. However, in Schnotz et al. (1999) experiment, the learning material was an interactive system which asked for a high level of user control. With an animation involving a minimal user control, collaboration might intensify the processing of the animation and thereby foster learning outcomes. Another issue with the experiment of Schnotz et al. (1999) might be related to grounding: As information in the animation was transient, dyads would suffer from the lack of permanent references, or artefacts, to ground their mutual understanding, while they could use the permanent static representation as a grounding artefact. Thus, the conditions for "optimal collaboration load" (Dillenbourg and Bétrancourt 2006) were not fulfilled in this experiment.

In this research, we designed an experimental study to investigate the effect of animation when an external memory device (recapitulative snapshots) is provided or not, in individual and collaborative learning settings. We claim that animation can be beneficial for learning under specific conditions currently investigated in the multimedia learning literature. First, animation will be more adequate than static pictures if the depicted phenomenon is intrinsically dynamic, which means that its understanding implies the construction of a dynamic mental model. Second, the instructional material has to be designed to support the individual processes involved in the construction of a dynamic mental model. Following Mayer and Chandler's (2001) recommendation, we used a segmented animation with a limited control device (to run the next sequence) in order to give learners time to integrate each chunk of information before presenting the next one. Finally, the learning setting has to be arranged to avoid cognitive underwhelming. One solution is to involve engagement from the learner through the learning task. In this study, we use a collaborative setting to ensure a deeper cognitive involvement of the learners through the elaboration of a shared understanding of the situation (Dillenbourg 1999). It is important to note that the goal of this experiment is not to investigate the conditions under which collaboration is optimally designed, but to use collaboration as a particular learning condition for studying multimedia material.

Based on these considerations, our hypotheses are the following: First, animation will improve learning as compared with static pictures since it provides the microsteps of the dynamic process that novices may not be able to mentally infer from static pictures (*enabling effect*). Second, collaboration is not expected to provide any benefit overall, since it was not scaffolded in a detailed manner. However, learning from animation will particularly benefit participants studying in dyads, since collaboration induces a higher cognitive involvement in the task, thus preventing the underwhelming risk. The higher cognitive involvement of dyads will be reflected through longer time spent studying the learning material than participants in individual setting. Third, the presence of snapshots will have an overall positive effect since they offload working memory. We expect an interaction between the presence of snapshots and the collaborative setting since dyads will use the snapshots as artefacts to help grounding their mutual understanding. Accordingly, the snapshots will benefit more to the dyads than to the individual learners.

Method

Participants and design: One-hundred and sixty first year university students were distributed depending on three factors in a between-subjects experimental design (8 groups). The first factor was the type of multimedia instruction (static/animated): The learning material was either a series of 12 static pictures or a series of 12 animated sequences. The

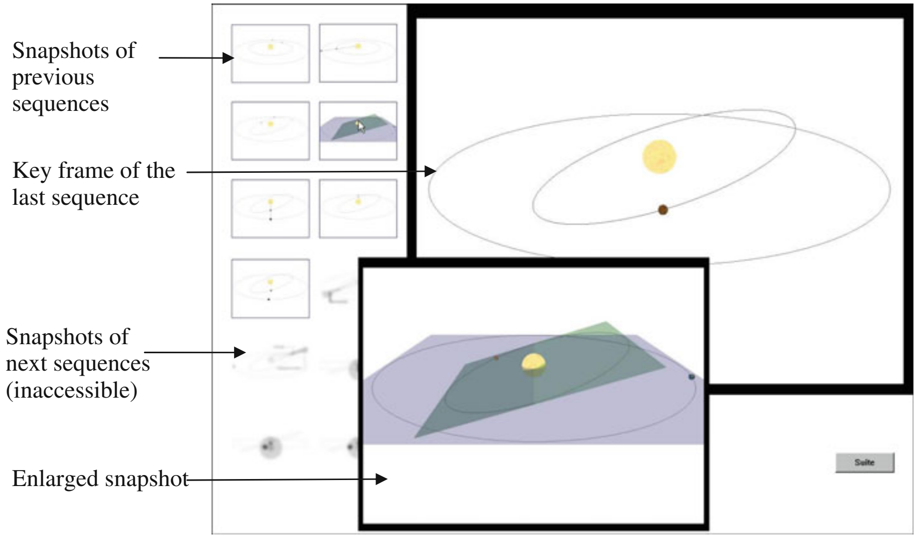

Fig. 1 Screenshot of the display during the pause after the seventh sequence of the astronomical animation in condition with snapshots. *Note*: The key picture of previous sequences could be accessed by moving the mouse over the corresponding snapshot during the pauses between two sequences

key picture from each animated sequence was used for the static version. The same audio explanations were used for both versions of the material.

The second factor was the presence of snapshots (with/without) depending on whether a snapshot of each sequence was displayed alongside the presentation or not (see Fig. 1).

The third factor was the learning setting (individual/collaborative) depending on whether learners studied the instructional material individually or in dyads.

Material: The learning material consisted of two successive multimedia instructions. The first one exposed how and why sometimes Venus can be seen in front of the Sun (Venus transit, see Fig. 1 for a snapshot). The second multimedia instruction explained the geological phenomenon of rift and subduction (see Fig. 2 for a screenshot). Each multimedia instruction was divided in 12 sequences representing a critical step in the phenomenon as defined by experts in the domain. After each multimedia sequence, the presentation paused and participants had to click on a button to run the next sequence.

Two versions of the material were designed according to the type of multimedia instruction. In the static version, we used the most informative and complete frame of each sequence. In the animated version, the sequences were short animations. Since the description and depiction of elements varied in complexity, the duration of each sequence could vary between 8 and 35 s. The audio commentary was identical for the static and animated version. The presentation durations of the static pictures were equal to their audio commentaries while the length of the animated sequences was occasionally longer than their audio commentary. As a result, the total duration for the static version of the multimedia instruction was shorter than for the animated one. The total duration of the animated sequences for Venus transit was $3'42''$ in contrast to $2'32''$ in the static version; this large difference was due to the depiction of three possible transits which took a long time to display in the dynamic presentation, but only a short sentence to describe. The duration of the sequence for the rift instruction was more constant across versions, $3'20''$ for the static version and $3'30''$ for the animated one.

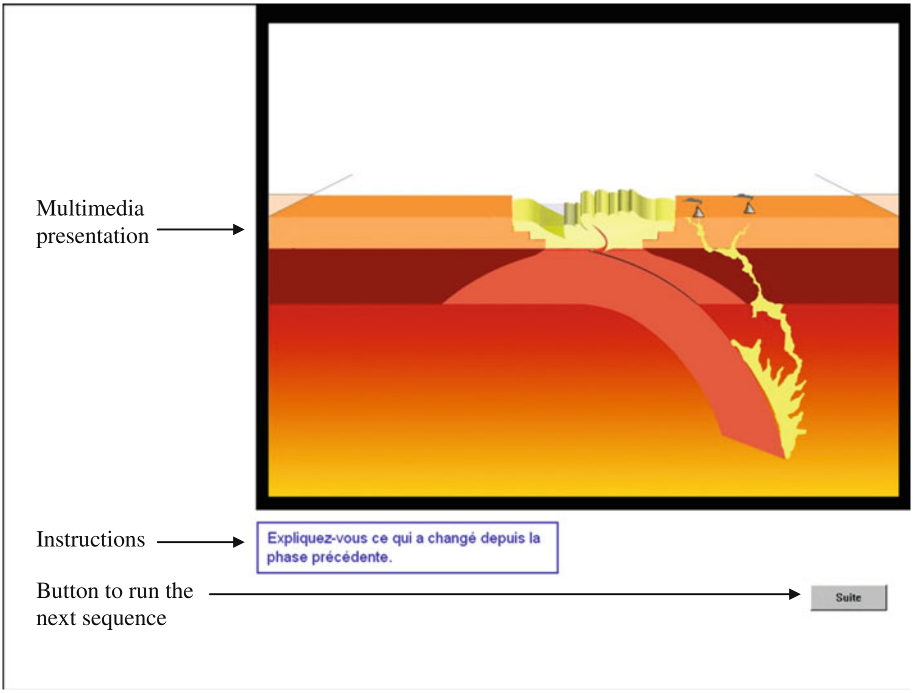

Fig. 2 Screenshot of the display during the pause after the eleventh sequence of the geological animation in a condition without snapshots. *Note*: The presentation could not be controlled, nor could be viewed a second time. A prompt appeared during pauses to encourage participants to "figure out what has happened in the last sequence". A button was provided to run the next sequence

Two displays were designed in order to implement the snapshots factor. In the display without snapshots, the static and animated pictures had the same size (800 × 600 pixels), and were displayed on the top-right part of the interface. In the version with snapshots, 12 empty areas for each multimedia instruction were provided on the left side of the screen (see Fig. 1). After each sequence was shown (static or animated according to the type of multimedia presentation), a snapshot of this frame (320 × 200 pixels) appeared in one of the empty areas. Participants could enlarge the snapshot (640 × 480 pixels) by moving the mouse over the corresponding field, but only during the pauses between sequences. The snapshots in the static version were the same as those in the animated version.

Dependent measures: We assessed learning outcomes, elaboration time, and subjective cognitive load as dependent measures. To measure learning outcomes for the astronomy instruction, participants had to answer a knowledge test with 16 multiple choice questions. The correct answers to nine of these questions were explicitly depicted or explained during the presentation and assessed learners' retention of elements. Seven questions involved transfer of knowledge; they evaluated the same principles but applied to other phenomena (e.g., a moon eclipse). For the geology instruction, 9 multiple-choice items were used to measure retention, and another 7 to assess transfer performance. Learners always received one point for a correct answer. Retention and transfer score from both instructions were added up into a total retention score and a total transfer score and were then transformed into percentages of correct answers for easier interpretation.

To assess subjective cognitive load, participants were asked to fill in a simplified version of the NASA-TLX (Hart and Staveland 1988), where they had to rate their mental demand, temporal demand, performance, effort and frustration each on a scale from 0 (low) to 100 (high). The scores from both materials were added up for the analysis.

Finally, students' elaboration time was recorded. After each sequence, the presentation paused and the participants were invited to take as long as they wished to reflect (individual setting) or discuss (collaborative settings) on the instruction. The time spent during these pauses was measured and called elaboration time. We chose to analyse elaboration time instead of total learning time since the duration of the instruction could vary across conditions (see material section) and was not under participants' control.

Procedure: All participants started the experiment with a short test on their knowledge in astronomy. Then, participants in the collaborative learning condition were seated in dyads in front of the same computer. Moreover, participants in the snapshots condition were explained the use and role of these snapshots. After a short introduction to the phenomenon of Venus transit, they were invited to study the first instruction. During the pauses, an onscreen instruction asked them to explain what had occurred in the last step. The goal of this instruction was to allow a reflective pause for learners in the individual setting and to stimulate an interaction between learners in the collaborative setting. Sequences could only be studied once; participants had no control over the pace and direction of the instruction except for the duration of the pauses between them, thus the total study time could vary across learners. When the learning session was over, participants were invited to fill in the simplified version of the NASA-TLX to assess their subjective load. Subsequently, participants answered the knowledge test with 16 multiple choice questions assessing retention and transfer performance.

Once the first part of the experiment was over, participants were invited to study the second multimedia instruction. Following the same procedure, they started with a pre-test on their knowledge of geology, followed by a quick introduction, the presentation (in the same experimental condition as the first one), the simplified nasa-tlx and lastly the retention and transfer tests.

Results

Learning performances

A multivariate analysis of variance was conducted with learning setting, snapshots and type of multimedia instruction as between-subjects factors and retention and transfer scores as dependent measures. Raw scores were collapsed across the two materials and the percentage of correct answers was used for comparison between our conditions (see Table 1). Significant overall differences were found between the animated and the static instruction, $F(2, 151) = 5.06$, $p < .01$. No significant differences were found between the conditions with and without snapshots, $F(2, 151) = 0.06$, *ns*, nor between the individual and the collaborative learning setting, $F(2, 151) = 0.52$, *ns*. The interaction between the type of instruction and the learning setting was statistically significant, $F(2, 151) = 5.02$, $p < .01$, as well as the interaction between the snapshots factor and the learning setting, $F(2, 151) = 4.50$ $p < .05$. The interaction between the type of instruction and the snapshots was not significant, $F(2, 151) = 0.69$, *ns*. The triple interaction was not significant either $F(2, 151) = 0.99$, *ns*. In the following, only the results of the ANOVAs performed

Table 1 Mean and standard deviation of percentage of correct answers by type of presentation, with and without snapshots

Learning setting	Static				Animated			
	Without		With		Without		With	
	M	SD	*M*	SD	*M*	SD	*M*	SD
Retention (%)								
Individual	50.00	15.04	54.97	13.47	61.90	15.24	56.35	15.64
Collaborative	53.40	11.77	51.39	14.52	59.72	17.74	60.10	13.57
Transfer (%)								
Individual	45.49	22.36	57.89	15.41	50.34	12.26	51.70	16.74
Collaborative	50.00	12.00	43.93	13.37	63.93	16.63	57.14	18.70

for the retention and transfer scores that correspond to the aforementioned overall effects will be described.

The mean retention and transfer scores are presented in Table 1. The analyses of variance performed on retention and transfer scores showed a significant main effect of the type of presentation. Participants in dynamic condition performed better than participants in static condition, both in retention scores, $F(1, 152) = 9.18$, $MSE = 217.67$, $p < .01$, Cohen's $d = .49$ (Cohen 1988), and transfer scores, $F(1, 152) = 6.25$, $MSE = 265.33$, $p < .05$, $d = .38$.

Concerning transfer scores, a significant interaction between the type of instruction and the learning setting was found, $p(1, 152) = 7.62$, $p < .01$. Participants in the collaborative setting showed better transfer performances in the dynamic condition than in the static condition $F(1, 76) = 15.10$, $MSE = 242.68$, $p < .01$, $d = .87$, whereas no significant effect was found in individual setting, $F(1, 76) = .03$, *ns*. Moreover, participants learning from a dynamic presentation scored higher in the transfer test when learning in a collaborative setting than individually, $F(1, 80) = 7.16$, $MSE = 265.12$, $p < .01$, $d = .57$, whereas the difference was not significant when learning from a static presentation, $F(1, 72) = 1.60$, *ns*.

The interaction between the snapshot factor and the learning setting was significant with regard to the transfer scores, $F(1, 152) = 6.65$, $MSE = 265.33$, $p < .05$. Participants learning individually performed better with than without the snapshots, but the difference was only marginally significant, $F(1, 76) = 3.28$, $MSE = 287.99$, $p = .074$, $d = .39$. Contrary to our expectations, participants learning collaboratively obtained inverse results; that is, higher transfer scores without snapshots than with them, $F(1, 76) = 3.39$, $MSE = 242.68$, $p = .07$, $d = .39$.

None of the other interactions was statistically significant. In particular, the two above-mentioned significant interactions concerning the transfer scores were not present on the retention scores (type of instruction $\times$ learning setting, $F(1, 152) < 1$; snapshots $\times$ learning setting, $F(1, 152) < 1$).

Since participants learned in dyads and then answered questionnaires separately, the independence of measurements is arguable. To rule out this possible critique, we performed a multilevel analysis and found no significant difference in the explained variance when either taking or not taking the composition of the groups into account (only the type of instruction factor was significant at the fixed level, $\beta = .68$, $p < .05$ for retention and $\beta = .95$, $p < .05$ for transfer). Moreover, standard deviations were relatively similar in

Table 2 Participants' mean and standard elaboration time by type of presentation, with and without snapshots

Learning setting	Static				Animated			
	Without		With		Without		With	
	M	SD	*M*	SD	*M*	SD	*M*	SD
Elaboration time (s)								
Individual	7.69	5.40	9.63	3.95	7.52	4.33	10.89	6.85
Collaborative	23.22	14.62	28.54	12.95	24.13	15.76	30.22	15.70

both learning settings (individual learning retention $M = 55.97$, $SD = 15.23$; transfer $M = 51.34$, $SD = 17.20$; collaborative learning: retention $M = 56.32$, $SD = 14.84$; transfer $M = 53.93$, $SD = 17.04$). Consequently, we considered our measurements as independent and performed statistical analysis accordingly (i.e., ANOVA).

Elaboration time

After each sequence, the presentation paused and the participants were invited to take as long as they wished to reflect (individual setting) or discuss (collaborative settings) on the instruction. The time spent during these pauses was measured and called *elaboration time*. We chose to analyse elaboration time instead of total learning time since the duration of the instruction could vary across conditions (see material section) and was not under participants' control. As expected, participants in the collaborative learning conditions spent considerably more time between the sequences than participants in the individual learning conditions, $F(1, 152) = 99.55$, $MSE = 123.94$, $p < .001$, $d = 1.59$, see Table 2 for details. Participants with snapshots also spent more time elaborating between sequences than participants without snapshots, $F(1, 152) = 5.62$, $MSE = 123.94$, $p < .05$, $d = .33$. There was no significant interaction concerning the elaboration time. No significant correlations were found between elaboration time and retention ($r = .07$, *ns*) or transfer scores ($r = .09$, *ns*) either. Thus, the increase in time was not directly related to an increase in performance.

The time spent studying the snapshots (i.e., elaboration time with an open snapshot) was relatively small ($M = 53.38$ s, $SD = 79.30$) as compared to the total elaboration time ($M = 427.76$ s, $SD = 342.08$). Participants with snapshots spent on average an eighth of their elaboration time with an open snapshot. Our other experimental conditions had no effect on the usage of snapshots. In particular, contrary to our expectations, participants in the collaborative learning setting did not spend more time with opened snapshots than participants in the individual learning setting, $F(1, 78) = 1.94$, *ns*.

Subjective evaluation of mental workload

The simplified NASA-TLX scores were quite similar across our conditions (see Table 3). A MANOVA revealed no significant overall differences between any of our experimental conditions. No significant differences were found between static and dynamic presentations $F(5, 148) = 0.74$, *ns*, between conditions with or without snapshots, $F(5, 148) = 0.46$, *ns*,

Table 3 Results of the five NASA-TLX scales depending on the experimental condition

Learning setting	Static				Animated			
	Without		With		Without		With	
	M	SD	*M*	SD	*M*	SD	*M*	SD
Mental demand								
Individual	155.25	34.20	146.48	34.44	135.55	44.69	144.83	34.41
Collaborative	152.24	24.14	151.05	34.56	139.51	44.10	149.23	27.26
Temporal demand								
Individual	116.61	38.41	105.00	32.82	103.62	37.55	100.07	29.39
Collaborative	110.97	38.36	98.02	37.18	95.27	44.54	115.57	28.70
Performance								
Individual	90.88	22.95	91.54	33.29	93.56	37.36	91.62	42.30
Collaborative	92.04	36.80	90.49	36.51	73.28	26.24	85.75	32.61
Effort								
Individual	117.79	32.28	111.20	30.84	110.44	35.16	115.14	38.91
Collaborative	103.61	35.17	104.12	33.95	95.19	27.74	102.46	37.14
Frustration								
Individual	94.13	26.99	97.15	29.50	96.48	32.58	90.32	41.63
Collaborative	90.23	42.03	97.96	35.01	97.27	32.17	74.94	38.02

N.B. scores from both instructions were added up (min 0, max 200)

nor between individual and collaborative learning setting $F(5, 148) = 1.43$, *ns*. Moreover, no interaction was significant (all ps $>$ .05). Accordingly, no ANOVA were performed on the subjective load scales.

Discussion

Our study stands as one of the few experiments reporting a positive effect of animated graphics over static ones for conceptual understanding of dynamic systems. However, while animation was beneficial to retention both for learners studying individually and in dyads, its positive effect on transfer tests was observed only for participants studying in dyads. According to Mayer's (2005) or Schnotz and Bannert's (2003) model of multimedia processing, the retention performance reflects the coherence and accuracy of propositional and pictorial models, but not the quality of the mental model. Since animation directly depicts changes over time, our results showing benefits of animation for retention can be explained using Tversky et al. (2002) apprehension principle. Interestingly, the results are not as straightforward for transfer questions, which reflect the quality of the elaborated mental model. Indeed, the animated instruction led to higher transfer scores only for participants in the collaborative learning setting. These results are strongly in favour of an underwhelming effect of the animated instructions, as suggested by Lowe (2003). By depicting the microsteps of the process, the animated presentations facilitated higher detail encoding. However, in an individual setting, they may have inhibited a mental animation of the system, hindering the construction of a dynamic mental model. When additional cognitive processing induced by collaboration is involved it could be shown that the same animation would allow for the construction of a more coherent and accurate dynamic mental model.

When collaboration facilitates learning from multimedia instruction

Obviously, the additional cognitive processing involved in a collaborative situation also takes more time. In our study, dyads always spent more time studying the animation than individual learners. The efficiency of the collaborative learning setting is thus questionable. It is important to note that learning improvements in the collaborative setting could not be attributed to the extra time spent studying the material. First, there was no correlation between elaboration time and performance. Second, only dyads using the animated instructions performed higher in the transfer test.

Individual and collaborative learning conditions are different on more than one dimension. Creating and maintaining a shared representation is one aspect but so is the verbalization. Participants in dyads had to talk together, so they had to verbalize their representations. Having to explain to others could have an effect, even if it is often smaller than collaboration itself (Ploetzner et al. 1999). In order to verify this hypothesis, we built a control experiment using different levels of verbalization for individual learners (Berney et al. 2005). The results did not confirm the hypothesis of an effect of verbalisation *per se*, in the sense that learners who where instructed to verbalize or self-explain the changes did not succeed better than learners who did not have to verbalize during the elaboration phase. These complementary results support the interpretation of the differences found between the individual and collaborative learning settings in this study in terms of grounding and maintained representation.

Our results for collaborating learners are not consistent with those of Schnotz et al. (1999). In their study, dyads performed better when using static rather than animated graphics. Conversely, for individual learners, advantages were found when using animated instead of static material. However, Schnotz et al. (1999) used interactive animations, which learners had to explore and use to answer precise questions. In our setting, participants had no control over the presentation (except for running the next sequence). These differences with regard to the interaction demands imposed by the animation might create meaningful differences in the management and outcomes of the learning session, as will be discussed in the next section in greater detail.

Did snapshots help learning from multimedia instruction?

Contrary to our expectations, the presence of snapshots of previous sequences did not effectively improve learning performances. However, individual learners obtained marginally higher transfer scores with snapshots, whereas collaborative learners obtained marginally lower scores when snapshots were presents. This interaction is very interesting since we expected the snapshots to serve as helpful artefacts for learner dyads in their mutual explanations. Contrary to this expectation, the presence of recapitulative snapshots helped learners studying individually but was detrimental to learners studying in dyads. In the individual setting, the snapshots acted as memory cues for the succession of steps, thus supporting the construction of a dynamic mental model. In a collaborative setting, participants using snapshots had to interact both with their partner and with the instructional material. Collaborative decisions had to be handled to decide which snapshot to open and when. Our results suggest that this collaborative human-computer interaction management prevented an otherwise efficient collaborative learning setting to take place. Referring to the split-attention effect (Chandler and Sweller 1991), we use the expression *split-interaction effect* to account for the interfering management of two sources of interaction. This interpretation in terms of interference between the additional processing of information induced by collaboration and the additional management of a supplementary source of information (the snapshots) is congruent with the analysis of verbal productions of the peers (Sangin et al. 2008). The main findings of this analysis showed that pairs provided with snapshots produced fewer verbal interactions than pairs without snapshots. The verbal utterances were also categorized and their analysis confirmed that less information on the learning content and less interaction management utterances were made in the group with snapshots than in the group without. The two categories of verbal productions were also positively correlated with the learning performances.

In order to further investigate our split interaction hypothesis, we plan to design a follow up study using the learning setting (collaborative or individual) and the level of interactivity with the learning material (with or without) as independent variables. This design would be more appropriate to reveal the existence of such a split-interaction effect.

The learning benefit of collaboration is a “side-effect” of efforts made by both peers, all along their interaction, to maintain a shared representation of the task (Dillenbourg 1999). This argument agrees quite well with our findings that show that adding active information processing can prevent an underwhelming effect. Adding a collaborative learning setting to increase the processing depth may not apply to very complicated animations, when many elements and changes are shown at the same time. In this case, the overwhelming effect would be close to a cognitive overload and asking participants to maintain a shared representation all over would be too demanding. However, in our study, the animated instruction contained both dynamic and static presentations since the last frame of each sequence was displayed during the pauses. The snapshots also inserted static elements in the animated instructions. These design decisions were made in order to keep the different versions of the materials as realistic as possible to what would be used in an instructional presentation; removing the static elements in the animated presentation did not make sense from this perspective.

Conclusion

Our results clearly support a possible benefit of animated pictures over a series of static ones for constructing an accurate mental model of a dynamic system. Nevertheless, this potential benefit is constrained to very specific learning settings. Our results suggest that supplementary cognitive processing such as a collaborative learning setting is necessary in order to benefit from the additional information about microsteps conveyed by a dynamic presentation. However, a collaborative learning setting—although effective—can be less efficient, since dyads needed more time to finish their learning phase than individual learners did. Future research will be carried out to investigate the possible split-interaction observed in this study, thus providing more elements for explaining how animations support understanding of dynamic processes and how collaborative learning can affect the efficiency of multimedia instructions.

Acknowledgments We are grateful to Gaëlle Molinari and the three anonymous reviewers for their thoughtful and helpful comments on the first version of this paper, and to Peter Gerjets and Katharina Scheiter who greatly improved the final version through multiple comments and proposals. This research was funded by the Swiss National Science Foundation (grant #11-68102.02).

References

Ainsworth, S., & Van Labeke, N. (2004). Multiple forms of dynamic representation. *Learning and Instruction, 14*(3), 241–256.

Baddeley, A. D. (1986). *Working memory*. Oxford: Clarendon Press.

Berney, S., Bétrancourt, M., & Rebetez, C. (2005). *Effect of verbalization and self-explanation on learning from multmedia instruction.* Paper presented at the Interlearn conference 05, Helsinki (Finland).

Bétrancourt, M., Dillenbourg, P., & Clavien, L. (2003). *Reducing cognitive load by delivery features in learning from computer animation.* Paper presented at the 10th EARLI biennale conference, Padova (Italy).

Catrambone, R., & Seay, A. F. (2002). Using animations to help students learn computer algorithms. *Human Factors, 44*, 459–511.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*, 293–332.

Clark, H. H., & Brennan, S. A. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. Teasley (Eds.), *Perspectives on socially shared cognition*. Washington: APA Books.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Dillenbourg, P. (1999). What do you mean by collaborative learning? In P. Dillenbourg (Ed.), *Collaborative-learning: Cognitive and conceptual approaches* (pp. 1–19). Oxford: Elsevier.

Dillenbourg, P., & Bétrancourt, M. (2006). Collaboration Load. In J. Elen & R. E. Clark (Eds.), *Handling complexity in learning environments: research and theory*. Amsterdam: Elsevier.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139–184). Amsterdam: Elsevier.

Hegarty, M., Kriz, S., & Cate, C. (2003). The roles of mental animations and external animations in understandig mechanical systems. *Cognition and Instruction, 21*(4), 325–360.

Hegarty, M., & Sims, V. K. (1994). Individual differences in mental animation during mechanical reasoning. *Memory and Cognition, 22*(4), 411–430.

Höffler, T. N., & Leutner, D. (2007). Intructional animation versus static pictures: A meta-analysis. *Learning and Instruction, 17*, 722–738.

Lowe, R. K. (2003). Animation and learning: Selective processing of information in dynamic graphics. *Learning and Instruction, 13*, 157–176.

Lowe, R. K. (2004). Interrogation of a dynamic visualisation during learning. *Learning and Instruction, 14*, 257–274.

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

Mayer, R. E. (2005). Introduction to multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 1–18). Cambridge: Cambridge University Press.

Mayer, R. E., & Chandler, P. (2001). When Learning is just a click away: Does simple interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology, 93*(2), 390–397.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 43–52.

Moore, J. L., & Rocklin, T. R. (1998). The distribution of distributed cognition: Multiple interpretations and uses. *Educational Psychology Review, 10*(1), 97–113.

Narayanan, N. H., & Hegarty, M. (2002). Multimedia design for communication of dynamic information. *International Journal of Human-Computer Studies, 57*, 279–315.

Ploetzner, R., Dillenbourg, P., Praier, M., & Traum, D. (1999). Learning by explaining to oneself and to others. In P. Dillenbourg (Ed.), *Collaborative-learning: Cognitive and computational approaches* (pp. 103–121). Oxford: Elsevier.

Roschelle, J., & Teasley, S. (1995). The construction of shared knowledge in collaborative problem solving. In C. E. O'Malley (Ed.), *Computer supported collaborative learning* (pp. 69–97). Heidelberg: Springer.

Sangin, M., Dillenbourg, P., Rebetez, C., Bétrancourt, M., & Mollinari, G. (2008). The effect of animations on verbal interaction in computer supported collaborative learning. *Journal of Computer Assisted Learning, 24*(5), 394–406.

Scheiter, K., Gerjets, P., & Catrambone, R. (2006). Making the abstract concrete: Vizualizing mathematical solution procedures. *Computers in Human Behavior, 22*, 9–26.

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction, 13*, 141–156.

Schnotz, W., Böckheler, J., & Grzondziel, H. (1999). Individual and co-operative learning with animated pictures. *European Journal of Psychology of Education, 14*(2), 245–265.

Schnotz, W., & Rasch, T. (2008). Functions of animations in comprehension and learning. In R. K. Lowe & W. Schnotz (Eds.), *Learning with animation: Research implications for design* (pp. 93–113). New York: Cambridge University Press.

Sweller, J. (2003). Evolution of human cognitive architecture. In B. H. Ross (Ed.), *The psychology of learning and motivation* (pp. 215–266). New-York: Academic Press.

Tversky, B., Morrison, J. B., & Bétrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human-computer Studies, 57*, 247–262.