Computers in Human Behavior 25 (2009) 354–359

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Using video and static pictures to improve learning of procedural contents

Amaël Arguel *, Eric Jamet

CRPCC - Laboratoire de Psychologie Expérimentale, Université Rennes 2 Haute-Bretagne, Place du Recteur Henri Le Moal, 35043 Rennes Cedex, France

ARTICLE INFO

Keywords:
Learning
Multimedia
Instructional visualization
Cognitive load

ABSTRACT

Animations and videos are often designed to present information that involves change over time, in such a way as to aid understanding and facilitate learning. However, in many studies, static displays have been found to be just as beneficial and sometimes better. In this study, we investigated the impact of presenting together both a video recording and a series of static pictures. In experiment 1, we compared 3 conditions (1) video shown alone, (2) static pictures displayed alone, and (3) video plus static pictures. On average the best learning scores were found for the 3rd condition. In experiment 2 we investigated how best to present the static pictures, by examining the number of pictures required (low vs. high frequency) and their appearance type (static vs. dynamic). We found that the dynamic presentation of pictures was superior to the static pictures mode; and showing fewer pictures (low frequency) was more beneficial. Overall the findings support the effectiveness of a combination of instructional animation with static pictures. However, the number of static pictures, which are used, is an important moderating factor.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

The use of computers for educational purposes has become increasingly common. Recent studies on multimedia presentations have produced various recommendations for helping designers to use multimedia with efficiency in various learning environments (see Mayer, 2005, for a review). This can be seen in the case of temporal contexts depicting continuous changes over time. In some experiments, using dynamic visualizations such as an animation or video could help learners build a more relevant internal representation of the content presented than static visualizations allow. Tversky, Morrison, and Betrancourt (2002) explained this effect by the "*congruence principle*" that occurs when the external representation presented by the learning material is close in nature to the internal representation needed for a relevant understanding of the content. Hence, the use of dynamic visualizations such as animations (Betrancourt, 2005; Tversky et al., 2002), sequential displays (Jamet, 2008; Jamet & Arguel, 2008), or video (Zacks & Tversky, 2003) are potentially well suited to learning content possessing temporal factors as a dimension (i.e., phenomena involving change over time). However, some studies have not shown dynamic representations to be consistently superior to static representations (Mayer, Hegarty, Mayer, & Campbell, 2005; for reviews see Betrancourt & Tversky, 2000; Höffler & Leutner, 2007; Park & Hopkins, 1993).

1.1. Dynamic visualizations

According to Betrancourt and Tversky (2000), the term "animation" refers to any representation which generates a series of frames, so that each frame appears as an alteration of the previous one, and represents an evolution in time. Thus, the term animation can refer to a rapid succession of pictures as in a cartoon, to animations made with a computer, or video clips made with a camera. Using animations in a learning environment can potentially present several advantages over static representations (e.g., Höffler & Leutner, 2007; Park & Hopkins, 1993). Firstly, because animations are able to use information from an analogical point of view (i.e., using an iconic depictive representation rather than a symbolic description representation), they can help the viewer to build relevant internal representation (Schnotz & Bannert, 2003). This seems to be particularly true when learning materials with high levels of visuo-spatial content, such as configurations of three-dimensional physical systems (Hegarty, 2005), or descriptions of the layout of several elements on a map, such as atmospheric systems (Lowe, 1993). Secondly, animation is by definition a rapid succession of pictures indicating a series of movements, manifestations and disappearances of graphic elements. Hence animations can be easily adapted to depict dynamic information involving changes over time because of the similarity in relation with time (Tversky et al., 2002). Thirdly, because animations are continuous, they give more information than a series of static pictures would. Thus, by explicitly showing the micro-steps needed between each important change, animations can be adapted for presenting continuous phenomenon because the learner is not required to infer

* Corresponding author.
E-mail address: amael.arguel@univ-rennes2.fr (A. Arguel).

0747-5632/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.chb.2008.12.014

how phenomena change from one step to the next (Betrancourt & Tversky, 2000). Fourthly, recent findings seem to indicate that the advantage of using animations instead of series of statics could be especially relevant for depicting some human-motor skills. Thus, authors showed that learning such contents as tying a knot or making a paper-folding can be improved by using a video-based material rather than series of statics (Ayres, Marcus, Chan, & Qian, 2009; Wong et al., 2009). Nevertheless, despite such potential advantages animations have not been found to be that effective in learning environments (Betrancourt & Tversky, 2000; Park & Hopkins, 1993).

Several explanations for the relative ineffectiveness of animations have been proposed. One of them is related to the "*congruence principle*" formulated by Tversky et al. (2002). This principle can explain the failings of animation in cases where the content to be learnt is not dynamic and/or a dynamic internal representation of the content is not essential for learning. Hence, from a cognitive load theory (see Sweller, 2005) perspective using dynamic representation such as animations, in situations where static representation alone would be sufficient, can lead to an increase in extraneous cognitive load. By presenting too much information at the same time, extraneous cognitive load is created leading to poorer learning (Sweller, 2005; Sweller & Chandler, 1994). In this case too much visual information is unhelpful.

Another possible explanation for the ineffectiveness of animations is their transient nature. As animations present dynamically temporal information at a constant rate, it can be difficult for the learner to sufficiently process information that is visible only for a short time before vanishing (see Ainsworth & VanLabeke, 2004; Ayres & Paas, 2007). Thus, holding important information in working memory while constructing a coherent internal representation with complementary information could increase the cognitive load (Sweller & Chandler, 1994). In such situations, the learner has no control over the pace of presentation of disappearing information. The comparison between animation and a series of static pictures is analogous with the comparison between an aural presentation and a printed text presentation. In the latter case, learners can read some passages quickly and others slowly, can compare several passages if necessary and have the possibility of re-reading any section if they misunderstand a specific element. In contrast, while watching animations or listening to aural speeches, learners cannot afford to miss any important information because it would be a permanent loss, and therefore, having perceived specific information as important it has to be kept active in working memory before being integrated with other information. Hence retaining and integrating information is very resource intensive on working memory.

There are two possible means to avoid the problems related with the transient nature of animation. First, in some cases, the disappearance of past information could be avoided by displaying key static pictures from the animation to remove the transience of some information elements (Rebetez, Bétrancourt, Sanguin, & Dillenbourg, 2005). This would permit learners to keep a visual trace of past events and allow them to review earlier information as necessary. In most cases, however, the dense content of animations makes this kind of presentation impractical as the visualization may become perceptually overloaded and, by consequence, very unclear for viewers. Second, learners can be given the possibility to control the pace of information with a "slider bar" or simply a "stop" and "play" button (Betrancourt, 2005; Hasler, Kersten, & Sweller, 2007). While using this "*interactivity principle*", the effectiveness of instructional animations can be improved. By giving the control to learners, they can avoid missing information and can slow down the pace of the learning material when it becomes more difficult to understand. In addition, while learning a cause-and-effect system from a presentation, user interactivity can be beneficial by allowing the segmentation of the presentation into chunks that will be more easily organized into a mental model (Mayer & Chandler, 2001).

The act of controlling the pace of a learning material can be problematic by itself. Indeed, using an interface to control pace is another activity the learners must cope with at the same time as learning the document. This added activity could be demanding in attention and detract from the principal task of learning the document (Hegarty, 2005). Also, the interface requires extra skills. Learners must master its use, which may be particularly problematic for people who do not usually use computers. Furthermore, asking learners to control the pace of the learning material is similar to asking them to have a specific relationship with it in which they identify its most relevant information (Hegarty et al., 2007). Consequently since interactivity involves strategies (Lowe, 1999), the lack of appropriate strategies could lead the learners to the problem of needing to identify and select relevant information, and lead to increased cognitive load (Schwan & Riempp, 2004).

Because using an interface and actively selecting relevant and transient information is cognitively demanding, in some cases the learners prefer not to use this possibility. For example, a study by Hasler et al. (2007) found that a group with learner control (able to stop an animation) learnt better than a group without learner control (unable to stop an animation), in spite of rarely using the interactive facility. How can this result be explained since all learners were confronted with the same visualization and its transient information? Perhaps this difference can be explained by the instructions given to learners before the learning phase. For example, Hegarty and her colleagues found better results for learners asked to *mentally animate* a system (a flushing cistern) from static diagrams before the learning phase than for those with no specific instructions (Hegarty, Kriz, & Cate, 2003; Hegarty, Narayanan, & Freitas, 2002). They concluded that people often learn more effectively if they are more active in the learning process. Similarly, in Hasler et al.'s study, the experimenter asked some participants to only watch the learning material, but asked others to use the interface of control if necessary. In this situation, we could envisage that the participants would not have the same perception of the learning material since the ones offered interface control have to actively partake in the multimedia presentation in order to gauge if a break was needed, even though they may not have used this possibility.

1.2. Combining video and static pictures in learning procedural content

Documents presenting procedural contents describe the evolution of a phenomena or a succession of actions over time and have several characteristics. The procedures are characterized by the existence of a beginning and an end, and between these two extreme points, there are a succession of steps describing each action or some steps of the procedure. The order of these steps is very important and an inversion can disable the execution or the comprehension of the procedure as a whole. Therefore, using multimedia presentations instead of paper-based documents could potentially provide an alternative way for helping the learning of procedural contents (Brunyé, Taylor, Rapp, & Spiro, 2006).

To keep the advantages of animations (i.e., for conveying temporal information) and reduce their limitations, an alternative format of presentation is proposed. In this format, animations are accompanied by both spoken text and static pictures. In this way, the depiction of micro-steps from the procedure and its natural development are maintained. Moreover, the static pictures are visible throughout the animation and act to limit the transience of the animation (see Fig. 1). The negative impact of the *split attention effect* (Ayres & Sweller, 2005; Chandler & Sweller, 1991; Sweller & Chandler, 1994) between static pictures and animation

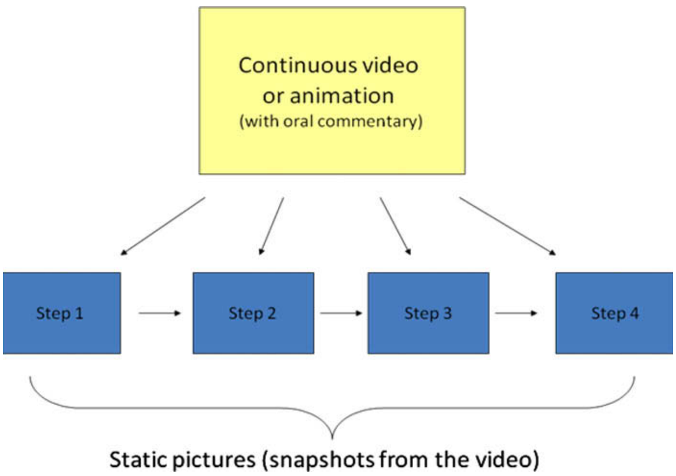

Fig. 1. Schematic representation of the learning material.

can be avoided by using pictures that are exact copies of frames from the video, so there is no need to integrate this identical information. The spoken text is also useful because it provides additional information to the visual modality and permits learners to hold their visual attention on the pictorial visualization. If the text were written, the learners would have to share their visual attention between the text and the animation and/or static pictures. With a spoken text, the integration of these various types of information can be improved according to the modality principle (Mayer & Moreno, 1998).

The third medium in this format is the pictures. Presented in series, the static pictures remain visible until the end of each video, hence limiting the transience of any one piece of information. Consequently, the learner is able to consult them repeatedly, and at anytime, thus reducing extraneous cognitive load caused by having to hold information in working memory while simultaneously processing new information Finally, the added static pictures can replace the need for pace-control while providing permanent pieces of information that do not require learners to act on the learning material.

According to findings on *event cognition*, learners should construct an internal representation composed in several discrete steps rather than in a smooth and continuous manner (Newtson, 1973; Zacks, Tversky, & Iyer, 2001). Using a series of static pictures in addition to videos, could produce an external representation of procedures closer to the mental representation expected of learners than that offered when only videos are presented. In this case, the distance is reduced between the external representation carried by the learning material and the internal representation of the content described. So, consistent with the *congruence principle* (Tversky et al., 2002), the combined material should be more efficient in promoting learning than visualizations composed only of videos. The aim of the experiments described in this paper is to test this hypothesis, by using multimedia presentations dealing with first aid materials and procedures.

2. Experiment 1

2.1. Hypothesis

In order to test the efficiency of our alternative format, composed of animated and static pictures with spoken text, we compared it with two other formats: a video recording with spoken commentary and a series of static pictures accompanied by the same spoken commentary. We hypothesized that the alternative (mixed) format would produce better learning outcomes than the video, because the static pictures added to the video would limit the transience of some information, and the alternative format would be more effective than the static presentation, because the static pictures would not provide enough information to the learners given the relatively complex nature of the materials to be learnt.

2.2. Method

2.2.1. Participants

Fifty-one undergraduate students (42 females and nine males with a mean age of 19.9 years; $SD = 2.23$) from the University of Rennes 2 (France) acted as voluntary participants. None had prior experience with the learning materials (first aid instructions) or procedures.

2.2.2. Materials

Five video clips were selected from the commercially available DVD "*Les gestes qui sauvent*" ("the saving gestures") edited by the French Red Cross (La Croix-Rouge Française, 2003). Each of these video clips dealt with a specific technique of first aid. The first two explained suffocation treatment for adults (video 1) and for infants (video 2), and the final three videos explained haemorrhage prevention techniques using direct compression and pressure bandages (video 3), pressure points (video 4), and tourniquets (video 5). Except for video 2 (infant suffocation), all the video clips were performed by actors playing victims, rescuers, and witnesses, in realistic dramatisations. Video 2 was a cartoon made with a flash-animation. All the videos were accompanied by spoken commentary by an off-camera voice that described and explained the actions performed by the rescuer at the same time as presented on screen. The total time needed for viewing the five videos was about 6 min. Between each video clip, the participant had to click on a "continue" button to start the next video clip. The video clips were displayed in a 380 × 280 pixel window centred on the screen.

For two conditions, we generated static pictures by taking snapshots of the video clips. These pictures corresponded with the most important steps of each procedure as determined by two experts. Thus, three pictures (from videos 2 and 4) or four pictures (from videos 1, 3, and 5) were associated with each video clip. At the beginning of each sequence, the pictures were displayed together on 248 × 188 pixel windows on the screen just under the video window (see Fig. 2).

2.2.3. Experimental design and procedure

The multimedia presentations were developed using Adobe Director software and presented on four PC computers, each with a 17-in. monitor set at 1024 × 768 pixel resolution. Each computer was equipped with earphones, keyboard, and mouse. There were three conditions using a between-subjects design. In the first condition, only the static pictures without the video were presented; the second condition presented only the video without the static pictures; the third condition included both the pictures and the video. All conditions were accompanied by the same spoken commentary. The participants were randomly distributed between these three conditions. After a short welcome, each participant was asked to fill in a paper-based questionnaire developed in order to detect if participants had either a high level of first aid knowledge or a very limited experience with computers. It should be noted that after analysis, no participants were removed from the study because nobody was specialist of first aid or novice with computer interfaces.

Next, the participants were seated in front of a computer in a room divided into cubicles. They were told to attend closely because after viewing the learning material, questions concerning

Fig. 2. Snapshot of the material during the learning phase of the video 1.

it would be asked. A maximum of four participants could participate in the study at the same time. Viewing the whole material took about 6 min and only one viewing was allowed.

Then, a paper-based questionnaire composed of ten open questions was given to participants (e.g., "What do you have to do when someone is suffocating? What is the procedure?" or "You find someone severely bleeding so you compress the wound. But you are alone and without a phone, so how can you raise the alert?"). These questions were made to test both comprehension and recall of the learning material but the test material did not differentiate between the two. The participants could use all the time they needed to answer the questionnaire. No access to the multimedia presentation was allowed during this test.

2.3. Results

The assessment questionnaires were randomly blind corrected and global scores were generated for each participant. These global scores were marked out of 25 and were equally composed with all questions relative to each sequence of the material. A one-way ANOVA found significant difference between the three groups, $F(2,50) = 22.41$, $MSE = 11.4$, $p < .001$, $\eta^2 = .483$. A linear contrasts analysis indicated significant differences between means for each pair of conditions (all $p < .05$). The poorest results came from the "pictures only" group ($M = 8.71$, $SD = 2.93$), and the learners assigned to the "video only" group produced intermediate results ($M = 11.41$, $SD = 4.37$) with a medium effect size ($d = 0.73$) in relation to the "pictures only" group (Cohen, 1988). The best results were produced by the participants assigned to the "video plus pictures" condition ($M = 16.35$, $SD = 2.55$) which represent a large effect size ($d = 2.78$) in relation to the "picture only" group as well as in relation to the "video only" group ($d = 1.38$).

2.4. Conclusion

As predicted, the alternative format (video plus pictures) outperformed the two other formats. Because static pictures should compensate for the transience of animations, they were found to be effective when added to videos. However, the static pictures alone seemed to be unable to convey enough information as the group viewing only the pictures performed the worst of all conditions, even when compared to a traditional animation (video). Beyond the assumption of the limitation of transience when adding static pictures to videos, there could be another explanation of this result. The series of static pictures added to videos could be beneficial because the steps depicted in pictures were actually the most crucial steps of the procedure described. As supposed in event cognition, learners are prone to construct mental representation of events as composed by several discrete steps rather than a continuous manner. Therefore, the precise number of pictures shown should be important. In a second experiment, we investigated the impact of the actual number of supporting pictures shown. We also tested the manner in which these pictures were presented by using either a static or dynamic presentation of the static pictures.

3. Experiment 2

3.1. Hypothesis

According to event cognition, when learners are asked to segment a procedural video in the finest grain (i.e., with a high number of sections), they obtain better results than when they are asked to segment on more coarse grain (Hanson & Hirst, 1989). We argue that presenting many static pictures in addition to a video presentation should lead to a similar advantage. In this case, the external representation would be closer to the expected internal representation, and thus contribute to fostering the congruence principle of the presentation. So, we expected to observe a superiority of the presentation of more pictures (high frequency condition) over the condition where fewer pictures were shown (low-frequency condition). With regard to the second factor (i.e., the type of appearance), we expected the dynamic appearance of pictures to outperform the static presentation. We based this assumption on the temporal contiguity effect that considers the benefit of synchronising complementary information from both the visual and the auditive channel (Ginns, 2006; Mayer, 2001). In the static condition, all pictures were presented simultaneously from the beginning of each video. In the dynamic condition, the pictures were presented one-by-one in synchronisation with the video and its spoken commentary. This sequential presentation could serve for guiding visual attention on the learning material (Jamet, Gavota, & Quaireau, 2008). In this case, the visual exploration during learning is helped because each picture can be observed only when the corresponding information is available. As each picture remained visible from its appearance to the end of the video, the advantage of maintaining transient information should be preserved. Thus, we expected to observe an advantage of presenting dynamically the pictures in each series, rather than statically.

3.2. Method

3.2.1. Participants

Seventy two undergraduate students (56 females and 16 males with a mean age of 20.7 years; $SD = 2.35$) from the University of Rennes 2 (France) volunteered to participate in this study.

3.2.2. Materials

The apparatus was identical to Experiment 1, and the material presented to participants was based on the "video plus picture" condition from Experiment 1. The two factors tested were static vs. dynamic presentations of the supporting pictures and low vs. high frequency. Static vs. dynamic presentations concerned the manner by which the pictures appeared positioned under the video on the computer's screen. In the static condition, all pictures were presented simultaneously on the screen at the beginning of the animation and for the entire duration of each of the five clips constituting the presentation. Under the dynamic condition, the pictures appeared sequentially at the specific times corresponding to the occasions when they occurred in the procedure, as described by the video and the spoken commentary. After appearing, each picture remained visible until the end of each video clip. The factor "frequency" was related to the number of pictures presented. In the low-frequency condition, only the pictures corresponding to the most crucial steps of each procedure were displayed, that is three pictures (videos 2 and 4) or four pictures (videos 1, 3, and 5) per video clip. In the high frequency condition, in addition to the pictures from the low-frequency condition, additional pictures corresponding to the intermediate steps between pairs of major steps were added. In this way, six pictures (videos 2 and 4) or eight pictures (videos 1, 3, and 5) per video clip were presented in the high frequency condition. In all conditions, spoken commentaries were identical.

3.2.3. Experimental design and procedure

Both static vs. dynamic and low vs. high frequency conditions were between-subjects factors. Each participant was randomly distributed into one of the four experimental groups. The apparatus, the preliminary questionnaire, and the assessment test were identical to Experiment 1.

3.3. Results

After analysing the preliminary questionnaire, two participants were eliminated from the study because they produced a sufficiently high score to be considered as expert in first aid, leaving 18 participants in both the "static/low frequency" and "dynamic/low frequency" groups and 17 participants in both the "static/high frequency" and "dynamic/high frequency" groups. As in Experiment 1, test scores were obtained from a paper-based questionnaire at the end of the study. A 2×2 ANOVA indicated a main static vs. dynamic effect with a better performance for the dynamic ($M = 16.4$, $SD = 2.83$) over the static condition ($M = 14.2$, $SD = 3.72$), $F(1,66) = 8.32$, $MSE = 9.97$, $p = .005$, $d = 0.66$. The main effect of low vs. high frequency indicated a significantly better performance for the low frequency ($M = 16.2$, $SD = 3.34$) than for the high frequency condition ($M = 14.4$, $SD = 3.36$), $F(1,66) = 5.92$, $MSE = 9.97$, $p = .01$, $d = 0.54$. The interaction between these two factors was not significant, $F(1,66) = 2.66$, *ns*.

3.4. Conclusion

The learners from the dynamic group outperformed learners of the static condition at the post-test questionnaire. As expected, the dynamic appearance of pictures seemed to improve learning of the material. That is particularly true when learners were in the high frequency condition, despite the fact that no statistically significant interaction was found, only an interaction trend. When assigned to the high frequency condition, the effect of the factor is strong in favour of the dynamic group. The second simple effect observed was surprising because it was against our hypothesis. For the segmentation frequency condition, learners from the low-frequency group outperformed those from the high frequency group. As suggested by the interaction trend, this effect is particularly strong with learners from the static groups compared with learners from the dynamic groups.

4. General discussion

In the first experiment, evidence for the beneficial use of the combination of video and static pictures from the video was found. Learners of the hybrid group performed better in the assessment phase than those from the video-only group. Thus, the addition of static pictures in the interface seems to be particularly beneficial, and is an improvement of the original video-only format. This point is very interesting because it shows that the original commercial presentation is not the best manner for conveying this type of information. Our hybrid format has caused a better learning of the verbal description of the procedures. The combination of video and static pictures seems to be the crucial point. Evidence for this important combination was provided by the result found when comparing the only pictures group with other formats. The lowest post-test scores were obtained by pictures-only group. We can conclude that the static pictures are particularly useful when added to video but when the pictures are presented without video, they are not beneficial. Moreover, the presentation of the video alone produced better results then the presentation of the series of pictures alone. Thus, in our experiment, the dynamic information carried by the videos seems to be necessary for learning the material. The pictures can only be useful when associated with videos because they are complementary.

In the second experiment, we studied the effect of the static pictures' features. We manipulated the number of pictures shown in each series, or in other words the segmentation frequency of the procedures described in the learning material. On this point, in accordance with the literature on event cognition, we hypothesised a superiority of the higher frequency condition (i.e., presenting many pictures). The data we have collected gave the opposite result: we observed a significant superiority for the low-frequency condition. Under this circumstance, we must consider a re-examination of the impact of the number of pictures during the learning. From the cognitive load point of view, showing too many pictures could increase extraneous load and consequently have a harmful effect on learning performance. This assumption is supported by the finding of absence of negative effect when the high frequency condition is crossed with the dynamic appearance of static pictures within each series. In this case, since the pictures were sequentially presented all along each sequence (i.e., from an empty screen, only the video had started then the pictures were presented one-by-one synchronized with video), there is not a perceptual overload during learning. Similarly, when fewer pictures are presented (in the low-frequency condition), the mode of appearance is not crucial. In our experiment, it is probable that in the low-frequency condition, the small number of pictures did not cause sufficient difficulty for visual exploration, even when presented in a static manner. Despite the interaction between the factors, mode of appearance and segmentation frequency not being statistically significant, the trend provides some plausible insights, which could be investigated further. For example, it is likely that, in the static condition, learners watch all of a series of pictures as soon as the video starts, without waiting for the appropriate time. Therefore, when they think about the pictures, they cannot watch the video and may lose some information. In a future study, it could be relevant to observe the process of visual exploration during learning by using online measures with an eye-tracking device (see Rayner, 1998, for a review).

Another aspect we have to consider is related with the level of cognitive load involved during learning. Using a subjective evaluation of level of cognitive load with a post-questionnaire, such as NASA-TLX (Tsang & Velasquez, 1996), Workload Profile (Hart & Staveland, 1988), or the 9-point mental rating scale developed by Paas (1992), should inform us of the difference in cognitive load across conditions (Paas, Tuovinen, Tabbers, & Van Gerven, 2003; Paas & Van Merrienboer, 1993; Paas, Van Merrienboer, & Adam, 1994). This type of measure may indicate whether the low assessment performances in some conditions are related with the level of cognitive load. It is possible that the act of presenting static pictures near videos could be beneficial because the pictures highlight to learners the crucial steps of an unknown procedure, rather than maintaining transient information. For novice learners it may be difficult for them to extract some relevant information among the continuous flow of the animation. If, the main function of static pictures was to stress the important points of the procedure, indicating these points by another way (e.g., colours coding, perceptual cues, keywords…) should produce similar results.

The results found in this study are promising. Much of the research in the field have compared static diagrams with animations only, and have not combined the two formats. However, combining them together under certain conditions have produced the best learning outcomes than either format presented individually. Consequently future research should be conducted to test the effectiveness of the hybrid approach to other learning domains.

References

Ainsworth, S., & VanLabeke, N. (2004). Multiple forms of dynamic representation. *Learning and Instruction, 14*, 241–255.

Ayres, P., Marcus, N., Chan, C., & Qian, N. (2009). Learning hand manipulative tasks: When instructional animations are superior to equivalent static representations. *Computers in Human Behavior, 25*, 348–353.

Ayres, P., & Paas, F. (2007). Making instructional animations more effective: A cognitive load approach. *Applied Cognitive Psychology, 21*, 695–700.

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.

Betrancourt, M. (2005). The animation and interactivity principles in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 287–296). New York: Cambridge University Press.

Betrancourt, M., & Tversky, B. (2000). Effect of computer animation on users' performance: A review. *Le Travail Humain, 63*, 311–329.

Brunyé, T. T., Taylor, H. A., Rapp, D. N., & Spiro, A. B. (2006). Learning procedures: The role of working memory in multimedia learning experiences. *Applied Cognitive Psychology, 20*, 917–940.

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

Cohen, J. (1988). *Statistical power analysis for behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Ginns, P. (2006). Integrating information: A meta-analysis of the spatial contiguity and temporal contiguity effects. *Learning and Instruction, 16*, 511–525.

Hanson, C., & Hirst, W. (1989). On the representation of events: A study of orientation, recall, and recognition. *Journal of Experimental Psychology: General, 118*(2), 136–147.

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. 139–183). Amsterdam: Elsevier Science.

Hasler, B. S., Kersten, B., & Sweller, J. (2007). Learner control, cognitive load and instructional animation. *Applied Cognitive Psychology, 21*, 713–729.

Hegarty, M. (2005). Multimedia learning about physical systems. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 447–465). New York: Cambridge University Press.

Hegarty, M., Keehner, M., Cohen, C., Montello, D. R., Lippa, Y., & Allen, G. L. (2007). The role of spatial cognition in medicine: Applications for selecting and training professionals. In *Applied spatial cognition: From research to cognitive technology* (pp. 285–315). Lawrence Erlbaum Associates Publishers.

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

Hegarty, M., Narayanan, N. H., & Freitas, P. (2002). Understanding machines from multimedia and hypermedia presentations. In J. Otero, A. C. Graesser, & J. Leon (Eds.), *The psychology of science text comprehension* (pp. 357–384). Lawrence Erlbaum Associates Publishers.

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

Jamet, E. (2008). Peut-on concevoir des documents électroniques plus efficaces? L'exemple des diaporamas. [Is it possible to improve multimedia presentation? The example of slideshows] *European Review of Applied Psychology*, **58**, 185–198.

Jamet, E., & Arguel, A. (2008). La compréhension d'un document technique multimédia peut-elle être améliorée par une présentation séquentielle de son contenu? [Can the learning from a multimedia technical document be improved with a sequential presentation of its content?] *Le Travail Humain*, **71**, 253–270..

Jamet, E., Gavota, M., & Quaireau, C. (2008). Attention guiding in multimedia learning. *Learning and Instruction, 18*, 135–145.

La Croix-Rouge Française (Producer), & ESV (Director). (2003). *Les gestes qui sauvent [The saving gestures]* [Motion picture]. Paris: ESV.

Lowe, R. K. (1993). Constructing a mental representation from an abstract technical diagram. *Learning and Instruction, 3*, 157–179.

Lowe, R. K. (1999). Extracting information from an animation during complex visual learning. *European Journal of Psychology of Education, 14*, 225–244.

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

Mayer, R. E. (Ed.). (2005). *The Cambridge handbook of multimedia learning*. New York: Cambridge University Press.

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

Mayer, R. E., Hegarty, M., Mayer, S., & Campbell, J. (2005). When static media promote active learning: Annotated illustrations versus narrated animations in multimedia instruction. *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.

Newtson, D. (1973). Attribution and the unit of perception of ongoing behavior. *Journal of Personality and Social Psychology, 28*, 28–38.

Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology, 84*, 429–434.

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*, 63–71.

Paas, F., & Van Merrienboer, J. J. G. (1993). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human Factors, 35*(4), 737–743.

Paas, F., Van Merrienboer, J. J. G., & Adam, J. J. (1994). Measurement of cognitive load in instructional research. *Perceptual & Motor Skills, 79*, 419–430.

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

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–422.

Rebetez, C., Bétrancourt, M., Sanguin, M., & Dillenbourg, P. (2005). *Collaborer pour mieux apprendre d'une animation. [Collaborating for better learning from animation]*. Paper presented at the 17e Conférence en Interaction Homme-Machine. Toulouse, France.

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

Schwan, S., & Riempp, R. (2004). The cognitive benefits of interactive videos: Learning to tie nautical knots. *Learning and Instruction, 14*, 293–305.

Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19–30). New York: Cambridge University Press.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition & Instruction, 12*, 185–233.

Tsang, P. S., & Velasquez, V. L. (1996). Diagnosticity and multidimensional subjective workload ratings. *Ergonomics, 39*, 358–381.

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

Wong, A., Marcus, N., Ayres, P., Smith, L., Cooper, G., Paas, F., et al. (2009). Instructional animations can be superior to statics when learning human motor skills. *Computers in Human Behavior, 25*, 339–347.

Zacks, J. M., & Tversky, B. (2003). Structuring information interfaces for procedural learning. *Journal of Experimental Psychology: Applied, 9*, 88–100.

Zacks, J. M., Tversky, B., & Iyer, G. (2001). Perceiving, remembering, and communicating structure in events. *Journal of Experimental Psychology: General, 130*, 29–58.