

Experimental Research on The Effectiveness of Animated Instruction

Liying Zhao¹ and Tao Huang

Abstract In this experiment(N=69), we investigated the learning effect of static graphic vs. animation, as well as the effect of user-control dynamic presentation. The results show that system-control animation has the highest cognitive load and the worst learning effect. The learning outcome of static picture and user-control animation are all better than system-control animation.

Keywords animation; static graphic; user control; cognitive load

1 Introduction

Animation has become increasingly popular in computer-based education across a range of subject domains and education levels. It is easy to intuitively forecast that animations might be particularly effective for its attention gaining effect, as well as special learning situations that involve motion. However, studies that comparing animation with static visuals have provided inconsistent results and failed to confirm its superiority over static visualization in improving learning. On the one hand, Höfler & Leutner (2007)^[7] made a meta-analysis of 26 primary studies that concerning the instructional effectiveness and efficiency of animation and static graphic. They found a medium-sized overall advantage of dynamic over static visualizations. The mean weighted effect size on learning outcome is $d=0.37$. In contrast, in a review by Tversky, Bauer-Morrison, and B érancourt (2002)^[20] most of the studies failed to show any advantages of dynamic compared to static visualizations, if they had, it could be attributed to more detailed or additional information presented in dynamic visualizations as compared to static versions of experimental materials. Koroghlanian and Klein (2004)^[10] also found in their experiment that participants who received animation as opposed to static illustrations spent more time studying the information, with no corresponding gains in terms of learning. Mayer (2005)^[14] found the learning result of static graphics are the same as or even better than the animation in explaining the process of mechanical system.

In order to explain the relative ineffectiveness of animation, researchers have proposed several possible reasons and cognitive load is regarded as the most significant factor (Ayres, Kalyuga, Marcus, and Sweller, 2005)^[1]. Cognitive load is used in cognitive psychology to illustrate the load related to the executive control of working memory. Ayres (2007, 2009)^{[2][3]} have argued that animations place greater cognitive load demands on the learner than corresponding static presentations. If information disappears from the screen as the animation progresses, learners may be forced to process current

¹L.Y. Zhao (✉)

Education Science and Technology Institute, Zhejiang University of Technology
Hangzhou, China
e-mail: zly@zjut.edu.cn

T. Huang

Education Science and Technology Institute, Zhejiang University of Technology
Hangzhou, China
e-mail: 897754164@qq.com

information while trying to remember or recall previous information. In these circumstances, animation produces an extraneous cognitive load. In contrast, static graphical displays can be revisited a number of times by the learner, hence eliminating the problem associated with transitory information. To compensate for the high cognitive load caused by transitory effect, a logical strategy is to allow the learner to stop the animation. Mayer and Chandler(2001)^[13] found that learners achieved higher scores if they can use the ‘stop’ and ‘play’ buttons to control the pace of the animation. Hasler et al. (2007)^[6] showed that providing learners the option to pause an animation on the day-night cycle improved transfer performance. Remarkably, in their study learners hardly used the control options but still got better learning performance, which they attributed to a deeper cognitive involvement of these learners. Tabbers et al.(2010)^[19] also successfully replicated the interactivity principle in terms of better transfer, However, this coincided with a large increase in time-on-task. But on the other side, some other studies have found rather detrimental effects of giving learners control over the pace and order of an animation. For example, Lowe (2004)^[12] showed that novices studying an interactive animation on weather maps did not use the interactive features in a very effective way. Moreover, Schnotz et al. (1999)^[18] found that interactive animated pictures were less effective than static pictures and argued that the interactive features only increased cognitive load and did not foster learning. In sum, research on instructional animations has not yet provided uniform evidence for the assumed positive effects of learner control in instructional animations.

The present study aims to continue exploring the different instructional effectiveness of static graphic, system-control animation and user-control animation, and further analyze their influences on learner’s cognitive load under the research framework of cognitive load theory.

2 Method

2.1 Design

In this experiment, The independent variable is three versions of computer-based presentation: static graphic, system-control animation, user-control animation. The static graphic is a serial of explanatory pictures. A continuous version of the animation is system-paced, which is 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. User-control animation allows the learners to decide when to stop or continue the animation according to their cognitive need by using ‘stop’ and ‘play’ buttons. The dependent variables are test performance (retention and transfer test) and subjective rating scale: mental effort rating, difficulty rating and usability rating. Cognitive load theory identifies three categories of cognitive load, which are extraneous cognitive load, intrinsic cognitive load and germane cognitive load. Extraneous cognitive load refers to unnecessary, ineffective cognitive load that is determined by the way the information is presented. Intrinsic cognitive load is the ‘natural’ load imposed by the information that must be acquired and it is influenced by the material’s complexity. Germane cognitive load is the cognitive resources being devoted to schema acquisition and automation (Fred Paas, Alexander Renkl, John Sweller, 2003)^[16]. According to previous research (Opfermann, M., Gerjets, P., Scheiter, K., 2007)^[15], difficulty rating is a sensitive indicator of the intrinsic cognitive load, mental effort rating can reflect germane cognitive load, while usability rating can be used to evaluate extraneous cognitive load.

2.2 Materials

2.2.1 Learning Materials

The Learning materials are about the process of HIV infection. The static graphic version contains 8 explanatory pictures and every four of them is in one big page (see Fig.1), so consisting two pages. System-paced animation is presented without pauses and at the end of the animation, a 'restart' button is given for the learners to replay the animation (see Fig.2). User-control animation is set with 'pause' button, so the learner can stop the animation as they will. Once pressed, the 'pause' button will turns into a 'play' button and learners can use it to restart the animation(see Fig.3). Both of the static graphic and animations have the same text for explanation. The material was designed and produced by using Macromedia Flash software.

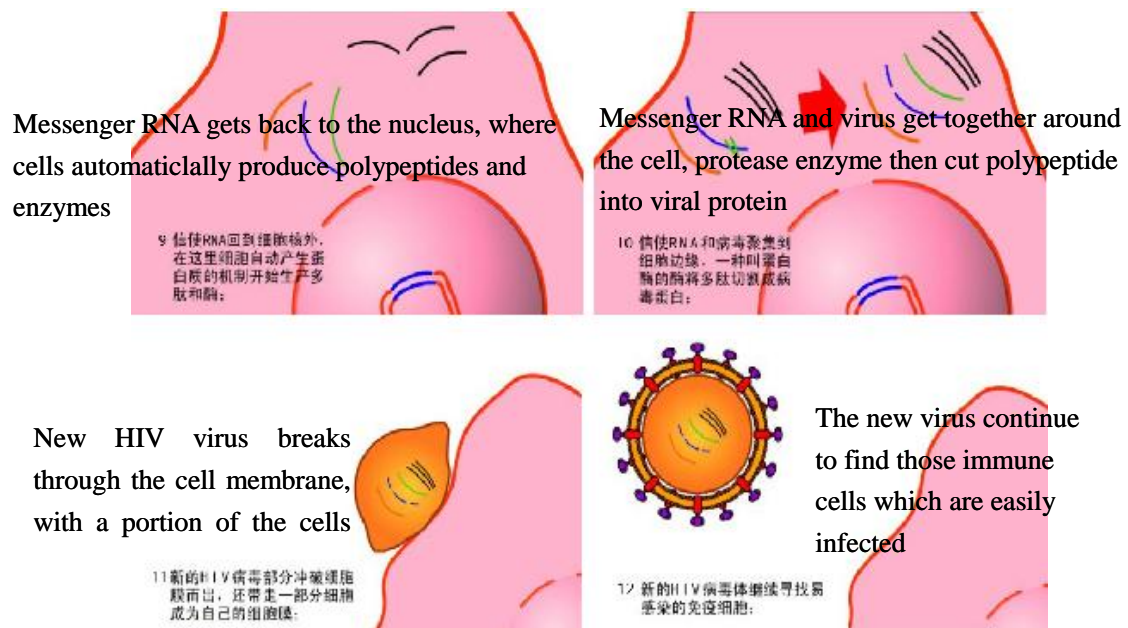


Fig.1 Example of Static graphic

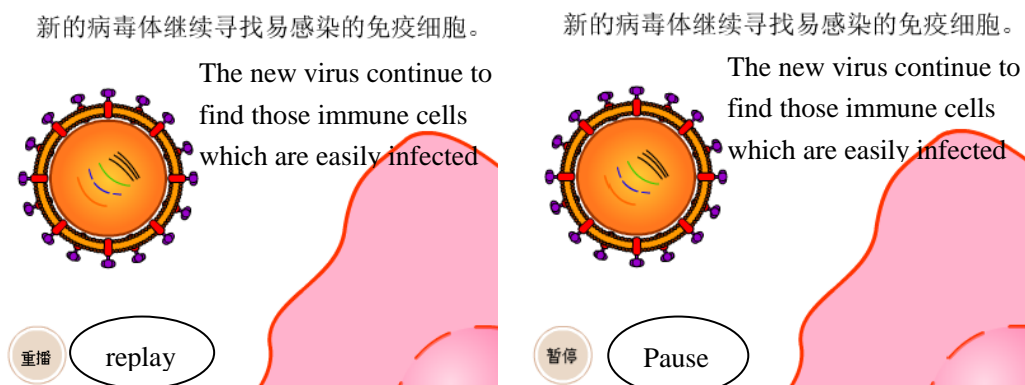


Fig. 2 Example of animation without control

Fig. 3 Example of user control animation

2.2.2 Test Materials

Test materials included the pretest, post retention test, post transfer test, subjective mental effort rating,

difficulty rating and usability rating.

The pretest was to assess learners' prior knowledge of HIV, so as to make sure whether the participants know little about what they are going to learn, and whether the three groups were at the same starting point of learning. The pretest included 6 questions, such as "which kind of cell does AIDS virus mainly infect?" The post retention test and transfer test were to test the students' mastery degree of the knowledge after the learning. Retention test consisted of 6 fill-in questions and all the answers can be found in the instructional materials. The questions were like: "Viral enzyme refers to _____, it can transfer virus RNA to DNA". Transfer test was to assess whether participants have taken sense of the knowledge and if they can apply what they have learned to related problems, so even the participants can't recall all the details, it will not influence their transfer test. The test included the following questions such as: "what is the main function of T-Helper Cells?" This is an open-ended question with no definitely correct answer. We gave score from 0-6 according the matching extent of the participant's answer with the acceptable answer.

Difficulty rating, usability rating and mental effort rating can be seen as the tests of the three different kinds of cognitive load. We use nine-point rating scales from 1 to 9 for participants to make the rating. In difficulty rating, participants were given the question "please choose a number from 1 to 9 to represent the difficulty of the material you have just learned, point 1 means very easy and point 9 means very difficult." In usability rating, the question is "please choose a number from 1 to 9 to represent the usability of the material you have just learned, point 1 means very convenient and point 9 means very difficult." In mental effort rating, the question is "please choose a number from 1 to 9 to represent the mental effort you invested in studying the material, point 1 means very low and point 9 means very high."

Learning materials were exhibited in computer, and all the tests were presented in paper.

2.3 Procedure

69 college students were recruited as participants in the study and they were all from Zhejiang University of Technology, China. All the participants were aged 18-22, with normal or corrected eyesight and have the basic ability to operate the computer. Before the experiment, none of the students were allowed to have strenuous exercise. All the participants were volunteered to this study. Before the experiment, they were given a gift so as to enhance their enthusiasm of participation.

Participants were randomly assigned to three groups. Before the experiment, participants received the following instruction: "You are going to learn an instructional material about the process of HIV infection, then you should answer some questions about what you just learn. So please keep your mind on the instruction and finish those tests carefully". Then, the paper-and-pencil pretest was conducted to measure participants' actual level of prior knowledge. After all the participants finished prior knowledge test, they were asked to start the multimedia learning. In order to exclude the influence of learning time on performance, all three groups were given 8 minutes and it is enough for all the participants to finish the study. After viewing the multimedia lesson, the subjective rating scale was presented first, then the retention test, and finally the transfer test. When the students had finished the tests, they were kindly thanked for their participation. The posttest lasted about 15 minutes and the entire experiment lasted about 30 minutes. All the data was analyzed in SPSS19.0.

3 Results

3.1 Prior Knowledge

Three participants were excluded because they failed to finish the prior test and posttest (both were 0), we think they were not well motivated. We finally got 66 participants, 23 in static graphic group, 22 in system-paced animation group, and 21 in user-control animation group. The prior knowledge of all the participants on the subject was rather low ($M=1.77$), and the three groups did not differ significantly on prior knowledge ($F(2,63)=0.208, p>0.05$).

3.2 Learning Performance

Tabel 1 shows the means and the standard deviations in the post-test. In the three groups, the retention and transfer test score have the following pattern: system-control group < static graphic group < user-control group. An analyses of variance (ANOVA) on the retention scores was further conducted, and a significant difference was found among these three groups ($F(2,63)=6.641, p<0.01$). Further analysis was conducted by using the least significant difference (LSD). The result indicates the retention scores of system-control group is significantly lower than the static graphic group ($p<0.01$) and user-control group ($p<0.01$). While no significant difference was found between static graphic and user-control animation group ($p>0.05$). The analysis outcome of transfer test scores is just like the retention test. Significant difference was found in the three groups ($F(2,63)=3.651, p<0.05$). LSD analysis shows that significant difference exists between system-control group and static graphic group ($p<0.05$), as well as the static graphic group and user-control group ($p<0.05$). No significant difference was found between static graphic and user-control group ($p>0.05$).

Table 1 The means (M) and standard deviations (SD) in the post-test

	Static graphic		System-control		learner-control	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
post-retention test	3.48	0.730	2.82	1.097	3.81	0.873
post-transfer test	2.83	1.114	2.23	0.752	2.86	0.655

3.3 Subjective Rating

The means and the standard deviations in the subjective rating are showed in table 2. An ANOVA for difficulty rating revealed no difference among the three groups ($F(2,63)=1.498, p>0.05$). An ANOVA for the mental effort also shows no difference ($F(2,63)=1.039, p>0.05$). However, we found a very significant difference in the usability rating among the three groups ($F(2,63)=18.161, p<0.01$). LSD Further shows that the usability of system-control group is significantly lower than the other groups ($p<0.001$), while the rating of the static graphic group and the user-control group was quite the same with no significant difference ($p>0.05$).

Table2 The means (M) and standard deviations (SD) in subjective rating

	Static graphic		System-control		User-control	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Difficulty	5.78	1.126	6.36	1.255	6.05	0.973
Mental Effort	6.70	1.521	6.27	1.453	6.86	1.108
Usability	5.57	1.237	7.14	1.082	5.05	1.244

4 Discussion

The experiment explores the instructional effectiveness of animation. The results show that the retention performance of static graphic and user-control animation is much better than system-control animation. One possible reason for this is the transitory nature of animation. As a consequence of transience, information presented at one moment needs to be maintained in working memory for the learner to be able to integrate it with information presented later which imposes high cognitive load on working memory. Static picture don't have this problem since learners can go back and look it again. User-control animation allows the learners to decide when to start or stop the animation according their cognitive need and it also reduce the visual information they have to process at the same time. So these two methods both could avoid cognitive overload. The transfer test scores of static graphic group and user-control group are significantly higher than the system-control group. While user-control group have higher transfer scores than static picture group, but no significant difference was found between the two groups. Transfer test was to assess learner's ability to solve novel related problems using what they have learned, which is based on the sufficient processing and comprehension of the leaning materials. Continuous animations without pause are commonly associated with entertainment and may lead the learners passively watching the animation. Learner will see a dynamic progress without much mental effort and may result in an illusion of understanding (B é rancourt, 2005)^[4], as well as a less thorough engagement with the content. Static graphic and learner-control animation are better at promoting learning material's deeper processing than uncontrollable animation. As to static picture, Mayer et al. (2005)^[14] have argued that statics allow more active processing as learners are required to make inferences from frame to frame instead of making passive observations. In this research, the static pictures are not independent pictures. They are the key frames taken from the animation and every four of these frames are deployed in one page. Boucheix et al. (2009)^[5] made comparisons between integrated sequential static frame presentation and sequential independent static frame presentation and they found integrated presentation is better. They thought integrated-sequential-static-frame presentation allowed direct visual comparisons between the micro-steps of the dynamic mechanical system, produced an improvement in the functional mental model. There are two reasons that may explain why user-control animation can facilitate transfer. First, user-control animation allows learners to decide the learning process of animation. When to stop, start or pause, all need the participation carry on positive cognitive processing. Second, learners will feel more autonomy in user-control animation and it will increase learner's learning interest and initiative. So learners are willing to have a deeper learning and will comprehend the materials better.

As to the subjective rating, the results showed no significant difference in difficulty rating. Difficulty rating is normally regarded as the most sensitive indication of intrinsic cognitive load, created by learning material's complexity. So to the same learning material, despite of the difference in

presentation mode, there should be no significant difference. Mental effort rating is an index for evaluating germane cognitive load. According to the definition of cognitive load theory, germane cognitive load is the resources used in schema construction, so it mainly influence transfer test score. Higher germane cognitive load naturally lead to better transfer test score. This is the same as the result of this study. User-control group shows the highest mental effort and the transfer test scores are also the highest in the three groups. The system-control group shows the opposite result and this means the user-control animation group has a deeper learning than uncontrollable group. When comes to the usability rating, there are significant differences among the three groups. The rating scores of static graphic and user-control animation are significantly higher than system-control animation. The usability reflects the extraneous cognitive load of the learners, so from the result we know the presentation mode affect learner's extraneous cognitive load and system-control animation has the highest extraneous cognitive load. With high extraneous cognitive load, the germane cognitive load of the learners will relatively low and this is coincidence with the subjective mental effort rating and transfer test results in this experiment.

5 Conclusion

The first aim of our experiment was to replicate the interactivity principle, which has successfully been accomplished. The study indicates that the learning results of static graphic and user-control animation are better than system-control animation. Static graphic and user-control animation allow students processing learning information deeply, enhancing germane cognitive load and transfer performance. System-control animation has caused the highest extraneous cognitive load for its transitory nature. But we expect and believe that animation will be more apt than static visualizations at conveying a deeper understanding of dynamic domain once their potential drawbacks like transience and visual complexity were diminished, while potential advantages such as the direct depiction of dynamic features is present, it is partly corroborated in this study.

This research is just a preliminary study on the learning effectiveness of static graphic and different version of animations. With the animation widely used in multimedia learning, more and more researchers begin to consider when and why dynamic and static visualizations might be best suited. For example, Lijia Lin, Robert K. Atkinson(2011)^[11] found that visual cues are useful in enhancing learning effect of the animation. In addition, many researchers start to pay attention the interaction effect of animation learning and learner's characteristics. They found learner's differences traits such as spatial ability, pre-knowledge and learning style all have impact on the learning effect of animational instruction(Ploetzner R., Lowe R., 2012)^[17]. For example, Höfler(2010)^[8] found that for learners with weaker spatial abilities, dynamic visualizations might be especially helpful, whereas for learners with higher spatial abilities the impact of using dynamic rather than static visualizations was less pronounced. Kalyuga (2008)^[9] investigates the relationship between instructional effectiveness of animated diagrams and levels of learner expertise. He found novice learners benefited more from static diagrams while more knowledgeable learners benefit more from animated diagrams. Animation is a popular media in recent multimedia learning environment. However, only if it was properly used, it could produce good learning effect. More research is needed to pinpoint the characteristics of effective animation.

Acknowledgments: This study is funded by Humanity and Social Science Youth foundation of Ministry of Education(11YJCZH246) in 2011. This work is also supported by funded project of 2011 outstanding young teachers in Colleges and universities in Zhejiang province

References

1. Ayres, P., Kalyuga, S., Marcus, N., et al. (2005). The conditions under which instructional animation may be effective. Paper presented at the International workshop Extending Cognitive Load Theory and Instructional Design to the Development of Expert Performance, August 29-30,2005. *Heerlen, the Netherlands: Open University of the Netherlands*.
2. Ayres, P., & Paas, F. (2007). Making instructional animations more effective: A cognitive load approach. *Applied Cognitive Psychology*, 21(6), 695-700.
3. Ayres, P., & Paas, F. (2009). Interdisciplinary perspectives inspiring a new generation of cognitive load research. *Educational Psychology Review*, 21(1), 1-9.
4. B é rancourt, M. (2005). The animation and interactivity principles, in R. E. Mayer (ed.) *Handbook of Multimedia*. (pp 287-296), Cambridge: Cambridge University Press
5. Boucheix, J. M., & Schneider, E. (2009). Static and animated presentations in learning dynamic mechanical systems. *Learning and instruction*, 19(2), 112-127.
6. Hasler, B. S., Kersten, B., & Sweller, J. (2007). Learner control, cognitive load and instructional animation. *Applied cognitive psychology*, 21(6), 713-729.
7. Hö ffler T. N., Leutner D. (2007). Instructional animation versus static pictures: A meta-analysis. *Learning and instruction*, 17(6): 722-738.
8. Hö ffler, T. N., Precht, H., & Nerdel, C. (2010). The influence of visual cognitive style when learning from instructional animations and static pictures. *Learning and Individual Differences*, 20(5), 479-483.
9. Kalyuga, S. (2008). Relative effectiveness of animated and static diagrams: An effect of learner prior knowledge. *Computers in Human Behavior*, 24(3), 852-861.
10. Klein, J., & Koroghlanian, C. (2004). The effect of audio and animation in multimedia instruction. *Journal of Educational Multimedia and Hypermedia*, 13(1), 23-46.
11. Lin, L., & Atkinson, R. K. (2011). Using animations and visual cueing to support learning of scientific concepts and processes. *Computers & Education*, 56(3), 650-658.
12. Lowe, R. (2004). Interrogation of a dynamic visualization during learning. *Learning and instruction*, 14(3), 257-274.
13. 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(2), 390.
14. Mayer, R.E., Hegarty, M., Mayer, S.,et al. (2005).When static media promote active learning: annotated illustrations versus narrated animations in multimedia instruction.*Journal of Experimental Psychology:Applied*,11(4), 256.
15. Opfermann, M., Gerjets, P., Scheiter, K.(2007) Online measurement of different types of cognitive load with an adjusted version of NASA-TLX[A]. Paper presented at *International conference on cognitive load theory*, 26-28.
16. Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational psychologist*, 38(1), 1-4.
17. Ploetzner, R., & Lowe, R. (2012). A systematic characterisation of expository animations. *Computers in Human Behavior*, 28(3), 781-794.
18. Schnotz, W., B ö ckheler, J., & Grzondziel, H. (1999). Individual and co-operative learning with interactive animated pictures. *European journal of psychology of education*, 14(2), 245-265.
19. Tabbers, H. K., & de Koeijer, B. (2010). Learner control in animated multimedia instructions. *Instructional Science*, 38(5), 441-453.
20. Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: can it facilitate?. *International journal of human-computer studies*, 57(4), 247-262.