International Journal of Engineering Education Vol. 29, No. 4, pp. 1013–1023, 2013
Printed in Great Britain

0949-149X/91 $3.00+0.00

Animated vs. Illustrated Software Tutorials: Screencasts for Acquisition and Screenshots for Recalling*

DRAGOLJUB NOVAKOVIĆ, NEDA MILIĆ and BRANKO MILOSAVLJEVIĆ
Faculty of Technical Sciences, University in Novi Sad, Trg Dositeja Obradovica 6, 21000 Novi Sad, Serbia. E-mail: milicn@uns.ac.rs

Previous studies in the domain of software learning did not reveal consistent results considering which type of visualisation is more efficient and effective for mastering software procedures. In this study, two independent units (A and B) for two procedures in the MS Excel application were made in two versions: an illustrated (screenshots) and an animated one (screencasts). Using a 2×2-experimental design with the visualisation type (illustrated vs. animated) and counterbalanced learning order (A, B vs. B, A) as independent variables, 72 participants were randomly assigned to one of the four groups. For unit A, the animated version was assigned to groups 1 and 3, and the illustrated version to groups 2 and 4, and vice versa for unit B. The main conclusion deducted from results and confirmed with participants' choice of preferable version is that animated tutorials are finest learning solutions for software procedure acquisition, while illustrated tutorials are foremost for procedure recall.

Keywords: software tutorial; cognitive load theory; 'animation deficit'; acquisition; recalling

1. Introduction

Software learning is a cognitively demanding task, often resulting in frustrating experiences and time losses for both experienced and novice users [1–3]. With a growing variety and complexity of software interfaces and features, the usage of static or dynamic visualisations in the domain of software learning support is inevitable. Animated tutorials, showing how a software expert performs tasks within an authentic environment, have generated a growing interest over the last two decades as a promising tool for mastering software [4], since users find them comfortable and motivating [1, 2, 5–7].

1.1 Related studies

Although an intuitive assumption about animated tutorials is that they facilitate mastering software procedures compared to other versions of instructional media-textual explanations with or without static visualisations [5, 8, 9], research results concerning their learning value are, in some cases, less affirmative. Instructional media that learners enjoy and prefer are not necessarily the ones that lead to the best learning results [2].

For evaluating a certain instructional design, retention cannot be separated from acquisition performance [10].

In the initial acquisition of software procedures, the benefits from animated tutorials are strongly evidence-based, since numerous studies have shown that subjects learning with animated tutorials have a faster and a more accurate immediate performance [8, 9, 11–15]. However, there are contradictory results about the retention and transferability of knowledge with animated tutorials indicated in delayed testing sessions [11, 12, 14].

Palmiter [11] found that, in the initial test, learners who studied animated tutorials acquired procedures in significantly less time and with fewer errors than the learners using textual instruction did. In a one-week delayed test, the same groups (animation-only and animation with narrated explanation) had difficulty in recalling acquired procedures and took longer to perform both identical and similar tasks. This phenomenon concerning users' short-term performance advantage, but a significant deterioration in the long-term retention performance and transfer, is described as an 'animation deficit' [11, 12].

Explaining the 'animation deficit', Palmiter defined the 'mimicry model' [11], according to which the learners appear to mimic animated demonstrations by passively memorising the procedure steps and copying them to software environment without interpreting the steps and, thus, without understanding the task. On the other hand, the processing of text manuals is more demanding since it includes a 'referential step'—the mapping of verbal concepts in the text to the objects and/or actions in the interface, leading to lower immediate performance, but better procedure encoding in the long-term memory [11]. Other explanation for 'animation deficit' is that watching an animation may create an illusion of understanding [2, 16], since students perceive it as a form of entertainment and an easy way of learning, which results in a superficial processing of the content [12, 17, 18].

* Accepted 21 April 2013.

'Animation deficit' effect, however, was not replicated in later studies [6, 8, 9, 15]. Lipps et al. [15] noticed that the drawbacks of animated demonstrations have only been shown for the tasks that were different from those in the acquisition phase, while working on identical tasks would not lead to deterioration in retention performance. Still, this conclusion indicates the lack of knowledge transfer as a sign of mimicry. To avoid mimicry and encourage learners to process animated demonstrations more deeply, Atlas et al. [6] used animated demonstrations which showed procedures different from those which had to be performed in the testing session. In addition, the animation was accompanied by narrated explanations presenting conceptual information, which was not redundant to the animation. In their study, the animated demonstration group even performed better in subsequent tests. In a study by Harrison [9], it was confirmed that the usage of visualisations, either static or dynamic, in on-line tutorials supports the users in acquiring procedures. A slight, no significant advantage for dynamic over static visualisations was also discovered in their study. The study of Spannagel et al. [8] demonstrated the advantage of animated demonstrations over text manuals with static visualisations. The animation group yielded better results than the text manual group in all tasks during the acquisition phase and in the post-test retention phase, and were more satisfied with the instructional materials than the text group.

The influence of software animated tutorials, and animated instruction material in general, on learning outcomes can be analysed within the theoretical framework of Cognitive load theory.

1.2 Theoretical framework of cognitive load theory

Cognitive load theory (CLT) [19–21] focuses on how constraints on the human working memory in order to help determine how, when and why some instructional materials are more effective and efficient for learning than others, with the end goal of defining instructional design guidelines that will align with human learning processes. According to CLT [20], a learner experiences three types of cognitive load by maintaining and processing information in the working memory: intrinsic (ICL), extraneous (ECL) and germane (GCL). The ECL and GCL are imposed by the design of instructional materials, where ECL is caused by unnecessary cognitive demands imposed by instructional design which hinder learning, while GCL is caused by the information and activities that foster acquisition. The ICL is imposed by the number of novel interactive information elements in the instructional material that have to be simultaneously processed in the working memory, which implies that the amount of ICL is influenced by the level of previous knowledge and cannot be manipulated by instructional design. The total cognitive load experienced during learning is additively composed of these three load types and, for learning to be successful, the instructional design should not cause a cognitive overload [19, 22]. From a CLT perspective, the optimal learning state happens when ECL is reduced and GCL increased. Every reduction of ECL caused by adequate instructional design does not directly implicate an increase in GCL, although it leaves 'more room' for germane processing to happen [23].

The large body of cognitive-load-based research has already provided us with valuable instructional design principles [23, 24]. The basic design principle, 'modality principle', recommends that text should be presented in a narrated, rather than written, form, whenever the visualisations, static or dynamic, are the focus of a textual explanation and both are presented simultaneously [25]. The rationale for this recommendation is that, when visualisations are combined with simultaneous written textual explanations, students are forced to split their visual attention during learning, which yields ECL and has a detrimental effect on learning performance [25, 26]. In addition, another CLT design principle, 'redundancy principle', indicates that redundant narrated and written textual explanations are less effective than narrated alone [27].

The design principles made so far, however, solely make predictions about learning with visualisations and text in general, but they do not make any specific predictions concerning comparisons between dynamic and static visualisations in terms of learning effects. The reason for that is currently unclear state in the literature. In a review by Tversky, Bauer-Morrison, and Bétrancourt [28], most studies had shown no advantages of dynamic compared to static visualisations or in the case where dynamic visualisations seemed superior to the static ones, it was revealed that dynamic visualisations had conveyed more information or had involved interactivity. However, a meta-analysis by Höffler and Leutner [29] revealed a significant advantage of dynamic over static visualisations in the acquisition of procedural knowledge and a medium-sized overall advantage.

Designing effective multimedia learning material requires more investigations on the cognitive processing of visualisations and on the difficulties experienced by learners. It was expected that dynamic visualisations would help learners mentally animate a dynamic process or a procedure [30], resulting in the reduction of cognitive load, namely ECL, compared to a situation in which a dynamic process or a procedure has to be reconstructed from a series of static visualisations [7]. If the learners have weaker previous knowledge or/and spatial abilities, static visualisations result in an increase of ECL and harbour the risk of an incomplete or inadequate mental model of dynamic changes [31]. In those cases, dynamic visualisations, by showing changes directly, may offload working memory [32], which either facilitate cognitive processes or enable deeper cognitive processes that otherwise would not have been possible (the 'facilitating' and 'enabling' function) [33]. For learners with higher previous knowledge or/and spatial abilities, the benefits of using dynamic rather than static visualisations may be less pronounced and these learners might even be bothered by dynamic visualisations containing redundant and unnecessary information that they know or believe to know, which might induce ECL [34].

An important tool in cognitive load theory is a self-rating instrument, the mental effort scale [22], and the difficulty scale, [35] for measuring the global cognitive load imposed during learning phases. Though it is still conflicting whether the perceived difficulty presents total cognitive load [35], the sum of ICL and ECL [36], or just ECL [32], the desired state is a low-perceived difficulty, since the instructional material of the group which achieved the same test performance with less amount of cognitive load is with a higher instructional efficiency [22].

2. Methods and materials

2.1 Participants and experimental design

The participants of this investigation were 72 undergraduate students (39 male, 33 female; age M = 20.2) from the Department of Graphic Engineering and Design, the Faculty of Technical Sciences (www.ftn.uns.ac.rs). In terms of previous experience with the software chosen for experimental learning material (Microsoft Excel), the sample can be considered satisfyingly homogeneous since all participants have successfully passed Excel beginner course in the previous semester which represents sufficient condition for the comprehension of units. The 72 participants were randomly assigned to one of the following four groups, each containing 18 participants, which resulted from a 2×2-experimental design (see Table 1) with visualisation type (illustrated-static vs. animated-dynamic) and learning order (A,B vs. B,A) as independent variables:

- Group 1 firstly learned unit A using animated tutorial and then unit B learning illustrated tutorial;
- Group 2 firstly learned unit A using illustrated tutorial and then unit B learning animated tutorial;
- Group 3 firstly learned unit B using animated tutorial and then unit A learning illustrated tutorial;
- Group 4 firstly learned unit B using illustrated tutorial and then unit A learning animated tutorial.

2.2 Learning material

The two units used in the study are demonstrating two independent data processing procedures of commercial spreadsheet application Microsoft Excel—What if analysis (Unit A; See Appendix A) and Advanced filtering (Unit B; See Appendix B). For the purpose of investigation, each unit is made in two versions: the illustrated ('screenshots' plus written textual explanation) and the animated ('screencasts' plus written textual explanation), explaining concisely the utility of a particular procedure and demonstrating the steps toward execution, in the form of worked examples. The worked examples used in learning units are tied up to realistic tasks ('worked example principle') [23], so that mastering these tasks goes beyond pure software training [8]. For example, in unit A it is explained what type of task can be done with certain software, how to apply the Goal-seek analysis procedure steps (e.g., choosing menu Data/What if analysis section/Goal-seek, filling the fields of dialogue window with appropriate data etc.), and the underlying concepts (e.g. that this analysis allows reverse computing of the input value that corresponds to desired output value).

Since a recent meta analysis [28] revealed cases where the superiority of the learning material with animated graphics compared to the one with static graphics was caused by the lack of equivalence between animated and static graphics in their content- animations were conveying more information,

Table 1. The graphical presentation of experimental design

Unit	Unit A		Unit B	
	Learning order of units (A,B or B,A)			
Modality (type of visualisation)	**1st**	**2nd**	**1st**	**2nd**
Animated tutorial	Group 1	Group 3	Group 4	Group 2
Illustrated tutorial	Group 2	Group 4	Group 3	Group 1

the special attention was invest to make animated and illustrated tutorial versions same regard to information content. The main difference between versions is the movement of the mouse cursor. The written textual explanations were the same in both versions and not redundant to visualisations, and the screenshots in the illustrated tutorials were taken from the animation. An equivalence procedure was used to make the text and demonstrations for each task as identical in content as possible [8].

Both the illustrated and the animation demonstrations are web-based and designed with a commercial authoring tool, TechSmith Camtasia, according to basic instruction design principles from Cognitive Theory of Multimedia Learning [37] and Cognitive Load Theory [19–21], which should result in reduced ECL and enhanced GCL:

- Tutorials will be more effective if they are segmented into smaller sections. Segmentation has been proposed as a means to reduce the high load occurring due to the transience of animations [37–39]. The determination of experimental units' lengths was based on recommendations for e-Learning lessons to 2–5 minutes [37], due to working memory's restrictions on how much information can be hold at one time. The units consisted of a brief introductory overview and procedure steps' demonstration.
- Tutorials will be more effective if the learner has control over the presentation. The self-pacing of both versions was comparable in the sense that participants could play, pause, stop and reread or replay the unit. In this manner, one of the potential drawbacks of dynamic visualisations, their transient nature [16], was ruled out.
- Tutorials will be more effective if the key information is cued or signalled. Both versions contained visual marks for mouse movement and mouse or key down events and highlighted active part of graphic interface.

In the case of unit A (see Appendix A), tutorial versions had textual explanation of approximately 160 words (in Serbian language), where the illustrated version had 8 static key screenshots and the animated version lasted 1:40 minutes. In the case of unit B (see Appendix B), tutorial versions had textual explanation of approximately 350 words (in Serbian language), where the illustrated version had 10 static key screenshots and the animated version lasted 2:30 minutes.

2.3 Procedure

Based on a pre-experimental interview, 72 participants with no experience with the tutorial units' content were chosen for participation. The participants were randomly assigned to one of four groups and were instructed to acquire as much information as possible and be as fast as possible in task performance. The testing was conducted in two sessions (the between-sessions interval was one week).

The experimental procedure of the first session had the following steps:

- participants took a previous knowledge test and spatial memory test (Corsi block tapping test);
- each of the participants viewed the version corresponding to his/her treatment group on the separate computer in duration of 5 minutes;
- after the acquisition phase, participants had to answer questions about the perceived difficulty of the software procedure presented in instructional materials;
- participants were given 5 minutes for completing individually on separate computers a near-transfer task (similar to the task from worked example-testing of procedure acquisition);
- participants were again given 5 minutes for completing a far-transfer task (a task for testing the comprehension of basic concepts);
- participants again answered the question about which type of instruction material they preferred and were allowed to comment on their choice.

The same procedure was used for both units.

After the testing phase of the second lessons, participants answered the question about which type of instruction material they preferred and were allowed to comment on their choice.

The second session's procedure steps were as follows:

- participants were given 5 minutes for completing individually on separate computers a near-transfer task (similar to the task from worked example-testing of procedure recalling);
- participants viewed again both unit tutorials and answered again the question about which type of instruction material they preferred and were allowed to comment on their choice.

2.4 Research goal and questions

Since learning the software usually puts the tremendous burden on cognitive load and previous investigations comparing dynamic and static visualisations are giving an inconsistent and often contradictory picture of learning outcomes, an important question still remains: which type of instructional material is an optimal support for learning software procedures and under which conditions?

For that purpose, in this study, the efficiency and effectiveness of software tutorials with dynamic visualisations ('screencasts') and with static visualisations ('screenshots') were comparatively investigated in usage environments that 'mirror' the complexity of the real world (real usage conditions).

Since the stepping-stone in this field is Palmiter's 'animation deficit' effect [11], caused by the mimicry of animation tutorials without textual explanation or with narrated explanation, this study investigates if written procedural text added to animation can provide a meaningful supplement to animation and facilitate retention and later transfer, although this does not align with the 'modality principle' mentioned in 1.1.

The experiment in this study had two testing sessions: an immediate and a week delayed. Furthermore, every participant had learnt two tutorial units—one unit with the illustrated and one with the animated tutorial and the learning order of units was counterbalanced, so that inter-individual variance was maximally reduced.

The research questions drawn from the considerations above are:

- Since two tutorial units are independent in terms of their content, will differences in learning performance and subjective measures of the cognitive load between groups 1 and 3, and groups 2 and 4 appear? In other words, does the learning order moderate or not dependable variables of this study?
- Will the groups using an animated tutorial show superior behaviour in mastering new procedures concerning performance time and the necessity for additional help (access to instructional material) initially after learning and also a week later?
- Will the self-rated perceived difficulty be lower for the groups using animated demonstrations?
- What would be the choice of preference (animated or illustrated version) for students in the phase of procedure acquisition and a week later, in the procedure recalling?

2.5 Measures

To sum up, the effect of visualisation type and learning order is assessed with the following dependent variables, classified into two groups:

(1) behaviour in the immediate-acquisition phase
   - completing rates of near-transfer tasks with and without additional help (access to instructional material);
   - near-transfer task performance time (sum of time working on task and time of additional instructional);
   - completing rates of far-transfer tasks;
   - assessment of cognitive load (the perceived difficulty);
   - choice of preference.

(2) behaviour in the delayed-recalling phase
   - completing rates of near-transfer tasks with and without additional help;
   - near-transfer task performance time;
   - choice of preference.

2.5.1 The previous knowledge test

The prior knowledge test assessed participants' basic knowledge on operations associated with developing, formatting and using a spreadsheet. It consisted of ten tasks (e.g. writing values in appropriate cells, writing standard mathematical and logical operations using basic formulas and functions, copying formulas) extracted from ECDL Module 5 Practice exercise book. One credit was assigned to each correct answer, resulting in a maximum of 10 points.

2.5.2 Spatial ability test

Spatial short-term memory was measured using the Corsi block tapping test [40].

2.5.3 Subjective measure of cognitive load

Total cognitive load was measured by subjective ratings of perceived difficulty [22] on a 9-point Likert scale.

2.5.4 Transfer tasks

Experimental testing contained two near-transfer tasks (one on immediate testing and one on delayed testing) and one far-transfer task for both lessons. The near-transfer tasks determined whether participants acquired all necessary procedure steps and whether they were able to apply the acquired procedures to new scenarios, similar to those in the worked example task (see appendices A and B). The far-transfer tasks required participants to understand underlying concepts of the procedure and make a generalisation about procedure usage.

3. Presentation

3.1 Control variables

No statistically significant between-group differences were detected in prior knowledge, $F(3, 71) < 1$, $p > 0.05$, or spatial memory, $F(3, 71) < 1$, $p > 0.05$.

3.2 Data analysis

Since the given near and far-transfer tasks were scored with 1 for mastering the whole task and 0 for an incomplete solution, Chi-square test for independence was used. For comparing performance times, a two-factorial ANOVA with type of visualisation and learning order as between-subject factors was used. First, the performance behaviour in the immediate testing session was analysed. Second, the results regarding the performance in

Table 2. Overview of dependent variables and statistical analysis results

Dependent variables	Mean value (standard deviation)				Statistic analysis (significance)
Unit A Immediate performance	**Group 1 (animated)**	**Group 2 (illustrated)**	**Group 3 (animated)**	**Group 4 (illustrated)**	**Main effect of type of visualisation**
Near transfer task					
Competing rates	18	18	18	18	χ^2 = 7.15, p < 0.01**
Without additional help	16	9	16	12	
With additional help	2	9	2	6	
Performance time	46.9 (16.1)	57.2 (15.5)	47.4 (14.7)	55.1 (12.3)	F = 6.75, p = 0.011**
Far transfer task					
Completing rates	16	13	15	12	χ^2 = 2.01, p > 0.05
Perceived difficulty	2.89 (1.02)	4.06 (1.16)	3.06 (0.99)	3.72 (0.83)	F = 14.846, p = < 0.001***
Unit B Immediate performance	**Group 1 (illustrated)**	**Group 2 (animated)**	**Group 3 (illustrated)**	**Group 4 (animated)**	**Main effect of type of visualisation**
Near transfer task					
Competing rates	18	18	18	18	χ^2 = 4.88, p < 0.05**
Without additional help	10	15	10	13	
With additional help	8	3	8	5	
Performance time	68.3 (14.6)	59.2 (19)	68.2(15)	60.5 (18)	F = 4.46, p = < 0.05**
Far transfer task					
Completing rates	11	15	13	12	χ^2 = 2.009, p > 0.05.
Perceived difficulty	3.67 (0.77)	4.83 (1.15)	4.17 (0.98)	4.67 (1.19)	F = 11.64, p = < 0.001***
Choice of preference	14(A)/4(I)	13(A)/5(I)	15(A)/3(I)	15(A)/3(I)	79.2% (A)/20.8% (I)
Unit A Delayed performance	**Group 1 (animated)**	**Group 2 (illustrated)**	**Group 3 (animated)**	**Group 4 (illustrated)**	
Near transfer task					
Competing rates	18	18	18	18	
Without additional help	15	11	13	10	χ^2 (1, n = 72) = 2.3, p > 0.05
With additional help	3	7	5	8	
Performance time	53.5 (16.9)	52.9 (11.1)	55.11 (17.3)	55.7 (7.3)	F < 1, p > 0.05
Unit B Delayed performance	**Group 1 (illustrated)**	**Group 2 (animated)**	**Group 3 (illustrated)**	**Group 4 (animated)**	
Near transfer task					
Competing rates	18	18	18	18	
Without additional help	10	9	8	11	χ^2 = 0.223, p > 0.05
With additional help	8	9	10	7	
Performance time	67.0 (14.6)	88.8 (31.1)	70.0 (15.3)	80.2(28.6)	F = 8.307, p < 0.01***
Choice of preference	6(A)/12(I)	6(A)/12(I)	11(A)/7(I)	10(A)/8(I)	45,8% (A)/54.2 % (I)

* Small Cohen's effect; ** Medium small Cohen's effect; *** Large small Cohen's effect [41].

the delayed testing session were presented. The results were reported separately for units A and B. In the end, the participants' preferable choice for learning new software procedures was analysed. Table 2 gives an overview of dependent variables and statistical analysis results.

3.3 Lesson A

3.3.1 Immediate testing session

Completing rates (near transfer task)

All participants (n = 72) mastered the near transfer task in predicted period. However, a Chi-square test for independence (with Yates Continuity Correction) indicated significant association between groups presented with animated (Groups 1 and 3) and illustrated (Groups 2 and 4) demonstration in completing near transfer task without additional access to instructional material, χ^2 (1, n = 72) = 7.15, p < 0.01, pfi = –0.347 (medium effect). A Chi-square test between separate groups showed significant difference between number of participants who master the task without additional help in Groups 1 and 2, χ^2 (1, n = 36) = 4.71, p < 0.05, phi = –0.42 (medium effect), and Groups 3 and 2, χ^2 (1, n = 36) = 4.71, p < 0.05, phi = –0.42 (medium effect).

Performance time (near transfer task)

The 2 × 2 ANOVA on performance time showed only a significant main effect of the visualisation type F(1,71) = 6.75, p = 0.011, η_{p^2} = 0.09 (medium effect) indicating that learners presented with dynamic visualisations performed faster than learners presented with static visualisations. No main effect of learning order and interactive effect of the two factors were found.

Completing rates (far transfer task)

Although number of participants from animated groups who solved far transfer task in predicted time frame was larger, a Chi-square test for independence (with Yates Continuity Correction) indicated no significant association between groups presented with animated (Groups 1 and 3) and illustrated (Groups 2 and 4) demonstration in completing near transfer task, χ^2 (1, n = 72) = 2.009, p > 0.05.

3.3.2 *Delayed testing session*

Completing rates (near transfer task)

All participants (n = 72) mastered the near transfer task in predicted time frame. However, a Chi-square test for independence (with Yates Continuity Correction) revealed no significant association between groups presented with animated (Groups 1 and 3) and illustrated (Groups 2 and 4) demonstration in completing near transfer task without additional access to instructional material, χ^2 (1, n = 72) = 2.3, p > 0.05.

Performance time (near transfer task)

The 2 × 2 ANOVA on performance time showed only no significant effect F (1, 71) < 1, p > 0.05.

3.4 *Lesson B*

3.4.1 *Immediate testing session*

Completing rates (near transfer task)

All participants (n = 72) mastered the near transfer task in predicted time frame. A Chi-square test for independence (with Yates Continuity Correction) indicated significant association between groups presented with animated (Groups 2 and 4) and illustrated (Groups 1 and 3) demonstration in completing near transfer task without additional access to instructional material, χ^2 (1, n = 72) = 4.88, p < 0.05, pfi = –0.289 (medium effect).

Performance time (near transfer task)

The 2 × 2 ANOVA on performance time showed only a significant main effect of the visualisation type (F(1,71) = 4.46, p = 0.038 < 0.05, η_{p^2} = 0.062 (medium effect) indicating that learners presented with dynamic visualisations performed faster than learners presented with static visualisations. No main effect of learning order and interactive effect of the two factors were found.

Completing rates (far transfer task)

Although number of participants from animated groups who solved far transfer task in predicted time frame was larger, a Chi-square test for independence (with Yates Continuity Correction) indicated no significant association between groups presented with animated (Groups 2 and 4) and illustrated (Groups 1 and 3) demonstration in completing near transfer task, χ^2 (1, n = 72) = 2.009, p > 0.05.

3.4.2 *Delayed testing session*

Completing rates (near transfer task)

All participants (n = 72) mastered the near transfer task in predicted time frame. However, a Chi-square test for independence (with Yates Continuity Correction) revealed no significant association between groups presented with animated (Groups 1 and 3) and illustrated (Groups 2 and 4) demonstration in completing near transfer task without additional access to instructional material, χ^2 (1, n = 72) = 0.056, p > 0.05.

Performance time (near transfer task)

The 2 × 2 ANOVA on performance time showed only a significant main effect of the visualisation type F (1, 71) = 8.307, p = 0.005 < 0.01, η_{p^2} = 0.109 (large effect, Cohen) indicating that learners presented with static visualisations performed faster than learners presented with dynamic visualisations. No main effect of learning order and interactive effect of the two factors were found.

3.5 *Perceived difficulty*

The 2 × 2 ANOVA on perceived difficulty of unit A showed only a significant main effect of the visualisation type F(1,71) = 14.846, p = <0.001, η_{p^2} = 0.179 (large effect, Cohen, 1988), indicating that the subjects using dynamic tutorials perceived procedure as significantly easier than the subjects using static tutorials.

The 2 × 2 ANOVA on perceived difficulty of unit B showed only a significant main effect of the visualisation type F(1,71) = 11.64, p = <0.001, η_{p^2} = 0.146 (large effect, Cohen, 1988) indicating that the subjects using dynamic tutorials perceived procedure as significantly easier than the subjects using static tutorials.

3.6 *Choice of preference*

The 79.2% of participants have chosen animation as a preferred way to learn software after the testing phase. A Chi-square test for independence indicates no significant association between groups and their choice of preference, χ^2 (1, n = 72) = 3.05, p = 0.38 > 0.05.

However, after a week, the animated tutorial was a choice of preference for only 45.8% of participants.

A Chi-square test for independence indicates no significant association between groups and their choice of preference a week later, χ^2 (1, n = 72) = 4.64, p = 0.2 > 0.05.

From subjects whose preferred choice was the animated tutorial, immediately after the acquisition 25 (42.1%) changed their opinion a week later, whereas only 1 subject who had preferred illustrated versions initially chose animated version later.

4. Discussion

The crossover design of experiment helps to keep variability low, and, thus, the validity of the results high, while still allowing for smaller-than-usual subject groups [42]. In the current study, there was no moderating role of learning order with respect to learning outcomes and experienced cognitive load.

Table 2 gives an overview of all relevant results. For both units, learning with the animated tutorial predominantly led to a better performance than learning with illustrated tutorials during the initial testing session—the animation groups were faster at solving the near-transfer task than the illustrated groups, and significantly smaller portions of animated group needed additional help. A fast and accurate initial performance for the animation groups when compared to the illustration groups can be explained with a reduced amount of cognitive demands since they did not have to spend time interpreting the steps. The perceived difficulty, motivation and user's choice of preferable version, measured immediately after the learning phase, mirror the initial hypothesis that users find animated tutorial an easier and a more motivating way of procedure acquisition. According to their comments, subjects supported with animated tutorial are more motivated and feel more confident in working with software, since animations show the authentic representation of the software's environment.

The lower perceived difficulty for animated groups who had better performance in the immediate testing session indicates consistently higher instructional efficiency of the animated tutorial compared to the illustrated one in acquiring software procedures. However, the animation tutorial users consistently became slower between the immediate and the delayed testing sessions in solving tasks similar to the worked example. In contrast, users of illustrations groups, in the case of both units, performed at a steady pace between the immediate and the delayed testing sessions and, as in the case of unit B, even outperformed the animation groups' users at the delayed test session.

Although a superficial processing of procedures when watching the animation material can also be a possible explanation why subjects learning with animated tutorials took consistently less time to initially acquire the interface procedures, but deteriorated in later performance, the insignificant difference between groups in far-transfer tasks' completing rates does not confirm that assumption. If the animation tutorial users had mimicked the tasks during learning, they would not have been capable of deeper comprehension [20], which they were. These results are noteworthy since they contradict the findings of Palmiter's [11] studies.

Since no statistically significant difference in the far-transfer was found, a possible explanation why learners from animation groups had difficulty in recalling acquired procedures can be found in the content browsing inflexibility of animated tutorials.

Changing the choice of the preferred version in favour of illustrated tutorials, subjects also evaluated the animated versions negatively concerning their later recalling uses.

Based on the comments left in the delayed testing session, animation tutorials were now judged as a waste of time, since subjects who used additional help had difficulties in overcoming the pacing deficiencies and redundant information that they already knew or believed to know. On the other hand, the same subjects noticed that illustrated tutorials could be accessed more easily on demand and scanned through until the information needed was found. The obtained results are validated also by the CLT design principles: 'expertise reversal principle' [34], according to which methods adequate for novices lose their positive effects or even have a negative impact on experienced learners who already have the understanding of the material being presented; and 'guidance fading principle', according to which support should be faded out with increased expertise [43].

The research findings also add support to the notion that the presence of written procedural explanations with visualisations does not facilitate mimicry and supports the subjects in acquiring accurate comprehension of the procedure steps.

5. Conclusion

The current state-of-the-art in the field of software learning indicates the necessity for a further investigation of software tutorial forms and their efficiency and effectiveness in real usage conditions. Given the additional cost and expertise needed to create and deliver animated tutorials, it is worthwhile to ask whether static illustrations with written procedural text can promote software learning that is as good or better.

The study results indicate a promising sign for the future diffusion of animated tutorials in educational and work environments, although they are not a panacea for every learning phase.

The main conclusion deducted from participants' performance scores and confirmed with the participants' choice of preferable learning media is that animated tutorials are better choice for software procedure acquisition, while illustrated ones are better for later recalling of already acquired procedures.

These findings have both a theoretical goal of contributing to the cognitive load theory and a practical goal of recommending design guidelines for optimal instructional material in every learning phase. The contribution to the theory can be found in new interpretations of: 'expertise reversal principle', where animation tutorial, representing the adequate instructional form for novices in the acquisition phase, loses its positive effects or even hampers the recalling of learned software procedure; and 'guidance fading principle', where highly detailed screencasts can be faded out and replaced with screenshots in the phase when learners have good-enough understanding of the procedures. The strength of this work can be seen in complex experimental design which includes both objective and subjective measures captured during testing sessions.

However, there are several limitations of this study.

Firstly, only animated versions with the written explanation were used in the study.

The direction of future researches should be the comparative investigation of tutorials with written explanation only, with narrated explanation only, and their combination, in order to examine the validity of 'modality' and 'redundancy' principles in the case of both animated and illustrated software tutorials.

Secondly, in the study was used only one software (MS Excel) with tasks demanding simple GUI operations.

Since the tasks in this study did not involve the aspect of motion beyond the movement of the cursor, the type of task should be investigated systematically in a further study. It can be hypothesised that a successful transfer depends on the type of software or the type of task to be solved and that more pronounced differences would be found between the illustrated and animated conditions for the GUI procedural tasks involving the movement and transformations of an actual object.

Finally, the participants' task performances within a software environment were not recorded.

The screen records of the task performances would give us further insight whether the performance deterioration of groups using animated versions was caused by the browsing inflexibility of animated tutorials.

Acknowledgments—This work was supported by the Serbian Ministry of Science and Technological Development, Grant No.:35027 'The development of software model for improvement of knowledge and production in graphic arts industry'.

References

1. G. Palaigeorgiou and T. Despotakis, Known and Unknown Weaknesses in Software Animated Demonstrations (screencasts): A Study in Self-Paced Learning–Settings, *Journal of Information Technology Education*, **9**(1), 2010, pp. 81–98.
2. T. Despotakis, G. Palaigeorgiou and I. Tsoukalas, Students' attitudes towards animated demonstrations as computer learning tools, *Educational Technology & Society*, **10**(1), 2007, pp. 196–205.
3. J. Lazar, A. Jones and B. Shneiderman, Workplace user frustration with computers: An exploratory investigation of the causes and severity, *Behaviour and Information Technology*, **25**(3), 2006, pp. 239–251.
4. B. Shneiderman and C. Plaisant, *Designing the user interface: Strategies for effective human-computer interaction*, 4th edn, Boston, MA: Addison-Wesley, 2005.
5. J. M. B. de Souza and M. Dyson, Are animated demonstrations the clearest and most comfortable way to communicate on-screen instructions? *Information Design Journal*, **16**(2), 2008, pp. 107–124.
6. R. Atlas, L. Cornett, D. M. Lane and H. A. Napier, The use of animation in software training: pitfalls and benefits, in M. A. Quinones and A. Ehrenstein (eds), *Training for a Rapidly Changing Workplace: Applications of Psychological Research*, Washington: American Psychological Association, 1997, pp. 281–302.
7. M. Hegarty, Dynamic visualizations and learning: getting to the difficult questions, *Learning and Instruction*, **14**(1), 2004, pp. 343–351.
8. C. Spannagel, R. Girwidz, H. Lothe, A. Zendler and U. Schroeder, Animated demonstrations and training wheels interfaces in a complex learning environment, *Interacting with Computers*, **20**(1), 2008, pp. 97–111.
9. S. M. Harrison, A comparison of still, animated, or non illustrated on-line help with written or spoken instructions in a graphical user interface, *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, Denver, CO: ACM, 1995, pp 82–89.
10. R. A. Schmidt and R. A. Bjork, New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training, *Psychological Science*, **3**(1), 1992, pp. 207–217.
11. S. Palmiter, The effectiveness of animated demonstrations for computer-based tasks: A summary, model and future research, *Journal of Visual Languages and Computing*, **4**(1), 1993, pp. 71–89.
12. S. Palmiter and J. Elkerton, Animated demonstrations for learning procedural computer-based tasks, *Human-Computer Interaction*, **8**(3), 1993, pp. 192–216.
13. P. E. Waterson and O'Malley, Using Animated Demonstrations in Multimedia Applications: Some Suggestions Based upon Experimental Evidence, *Proceedings of the Fifth International Conference on Human-Computer Interaction*, New York, 1993, pp. 543–548.
14. M. P. Kerr and S. J. Payne, Learning to use a spreadsheet by doing and by watching, *Interacting with Computers*, **6**(1), 1994, pp. 3–22.
15. A. Lipps, J. Trafton and W. Gray, Animation as documentation: A Replication with Reinterpretation, *Proceedings for the 45th annual conference of the Society for Technical Communication Arlington*, http://www.stc.org/proceedings/ConfProceed/1998/PDFs/00006.PDF, Accessed 11 January 2013.
16. M. Bétrancourt, The animation and interactivity principles in multimedia learning. in R. E. Mayer (eds), *Cambridge handbook of multimedia learning* ,Cambridge University Press, New York, 2005, pp. 287–296.
17. A. Ertelt, A. Renkl and H. Spada, Making a difference—Exploiting the full potential of instructionally designed on-screen videos, *Proceedings of the International Conference on Learning Sciences*, 2006, pp. 154–160.
18. R. K. Lowe, Animation and learning: Selective processing of information in dynamic graphics, *Learning and Instruction*, **13**, 2003, pp. 247–262.
19. J. Sweller, Cognitive load during problem solving: effects on learning, *Cognitive Science*, **12**(2), 1988, pp. 257–285.
20. J. Sweller, J. J. G. Van Merriënboer and F. Paas, Cognitive architecture and instructional design, *Educational Psychology Review*, **10**(3), 1998, pp. 251–295.
21. J. J. G. Van Merriënboer and J. Sweller, Cognitive load theory and complex learning: Recent developments and

future directions, *Educational Psychology Review*, **17**(2), 2005, pp. 147–177.
22. F. Paas, J. E. Tuovinen, H. Tabbers and P. W. M. Van Gerven, Cognitive load measurement as a means to advance cognitive load theory, *Educational Psychologist*, **38**(1), 2003, pp. 63–71.
23. T. Chong, Recent Advances in Cognitive Load Theory Research: Implications for Instructional Designers, *Malaysian Online Journal of Instructional Technology (MOJIT)*, **2**(3), 2005, pp. 106–117.
24. J. Sweller, Instructional design consequences of an analogy between evolution by natural selection and human cognitive architecture, *Instructional Science*, **32**(1), 2004, pp. 9–31.
25. J. Sweller and P. Chandler, Why some material is difficult to learn. *Cognition and Instruction*, **12**(3), 1994, pp. 185–233.
26. B. Park, R. Moreno, T. Seufert and R. Brünken, Does cognitive load moderate the seductive details effect? A multimedia study, *Computers in Human Behavior*, **27**(1), 2011, pp. 5–10.
27. R. E. Mayer, J. Heiser and S. Lonn, Cognitive Constraints on Multimedia Learning: When Presenting More Material Results in Less Understanding, *Journal of Educational Psychology*, **93**(1), 2001, pp. 187–198.
28. B. Tversky, J. Bauer-Morrison and M. Bétrancourt, Animation: Can it facilitate, *International Journal of Human–Computer Studies*, **57**(1), 2002, pp. 247–262.
29. T. N. Höffler and D. Leutner, Instructional animation versus static pictures: A meta-analysis, *Learning and Instruction*, **17**(1), 2007, pp. 722–738.
30. M. Hegarty and S. Kriz, Effects of knowledge and spatial ability on learning from animation, in R. K. Lowe and W. Schnotz (eds), *Learning with animation: Research and design implications*, Cambridge University Press, New York, 2007, pp. 3–29.
31. W. Schnotz and R. Lowe, A unified view of learning from animated and static graphics. in R. Lowe and W. Schnotz (eds), *Learning with animations: Research implications for design*, MA: Cambridge University Press, New York, 2008, pp. 304–356.
32. T. Kühl, K. Scheiter, P. Gerjets and J. Edelmann, The influence of text modality on learning with static and dynamic visualizations, *Computers in Human Behavior*, **27**(1), 2011, pp. 29–35.
33. W. Schnotz and T. Rasch, Enabling, facilitating, and inhibiting effects of animations in multimedia learning: why reduction of cognitive load can have negative results on learning, *Educational Technology: Research and Development*, **53**(3), 2005, pp. 47–58.
34. S. Kalyuga, P. Ayres, P. Chandler and J. Sweller, The expertise reversal effect, *Educational Psychologist*, **38**(1), 2003, pp. 23–31.
35. N. Cerpa, P. Chandler and J. Sweller, Some conditions under which integrated computer-based training software can facilitate learning, *Journal of Educational Computing Research*, **15**(4), 1996, pp. 345–367.
36. P. Gerjets, K. Scheiter, M. Opfermann, F. W. Hesse and T. H. S. Eysink, Learning with hypermedia: the influence of representational formats and different levels of learner control on performance and learning behaviour, *Computers in Human Behavior*, **25**(2), 2009, pp. 360–370.
37. R. E. Mayer, *The Cambridge handbook of multimedia learning*, Cambridge University Press, New York, 2005.
38. R. E. Mayer and R. Moreno, Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, **38**(1), 2003, pp. 43–52.
39. P. Ayres and F. Paas, Making instructional animations more effective: A cognitive load approach. *Applied Cognitive Psychology*, **21**(6), 2007, pp. 695–700.
40. R. Kassels, C. van Zandvoort, J. E. Martine, A. Postma, L. J. Kappelle and E. H. F. de Haan, The Corsi Block-Tapping Task: Standardization and Normative Data. *Applied Neuropsychology*, **7**(4), 2000, pp. 252–258.
41. J. Cohen, *Statistical power analysis for the behavioral sciences*, 2nd edn, NJ: Erlbaum, Hillsdale, 1988.
42. A. Mink, Conducting Repeated Measures Analyses: Experimental Design Considerations, http://ericae.net/ft/tamu/Rm.htm, Accessed 11 January 2013.
43. D. Leutner, Double-fading support—a training approach to complex software systems, *Journal of Computer Assisted Learning*, **16**, 2000, pp. 347–357.

Appendix A: Goal-seek analysis (Lesson A)

Example task: How much does point value need to increase to produce a gross wage of 152.114,70, with other conditions unchanged?

Near-transfer task (immediate test session): How many points are necessary to produce a net wage of 800,000, with other conditions unchanged?

Far-transfer task (immediate test session): On how many different ways can payment amount be 1,500,000, with constant point value?

Near-transfer task (delayed test session): How much should life insurance amount to produce a net wage of 991,000, with other conditions unchanged?

Appendix B: Advanced filtering (Lesson B)

Example task:

(a) Filter newspapers from the publisher 'Press' with prices less than 100.
(b) Filter newspapers from publishers 'Press' or 'Dnevnik'.

Near-transfer task (immediate test session): Filter newspapers with price less than 15 and percentage of discount 18.

Far-transfer task (immediate test session): Filter newspapers from the publishers 'Press' and 'Politika' with the prices greater than 30, but less than 60.

Near-transfer task (delayed test session): Filter newspapers with percentage of discount 17 and 19.

Dragoljub Novaković is a professor and Head of the Department of Graphic Engineering and Design (Faculty of Technical Sciences, University in Novi Sad). He received his PhD degree from the University in Novi Sad as a first PhD candidate in the field of Graphic Engineering and Design in the country. He was enrolled in over 40 national and international projects and he published over 300 scientific works and many university books. Currently, he is a leader of project 'The development of software model for improvement of knowledge and production in graphic arts industry', supported by the Serbian Ministry of Science and Technological Development.

Neda Milić is a teaching assistant at the Department of Graphic Engineering and Design (Faculty of Technical Sciences, University in Novi Sad). She received her MsC degree from the University in Novi Sad with award for the best student in generation. She is enrolled as a researcher in project 'The development of software model for improvement of knowledge and production in graphic arts industry', supported by the Serbian Ministry of Science and Technological Development.

Branko Milosavljević is currently an associate professor at the Department for Computing and Control Engineering (Faculty of Technical Sciences, University in Novi Sad). In 2003 his PhD thesis entitled 'Extensible Multimedia Information Retrieval System', defended at the Faculty of Technical Sciences, was awarded with reputable 'Mileva Maric Einstein' award for the best PhD thesis in the field of computing. He published over 100 scientific works and many university books.