Relative effectiveness of animated and static diagrams: An effect of learner prior knowledge

Slava Kalyuga *

School of Education, The University of New South Wales, Sydney 2052, Australia

Abstract

This paper investigates the relationship between instructional effectiveness of animated vs. static diagrams and levels of learner expertise in the task domain of transforming graphs of simple linear and quadratic functions. It was demonstrated on many occasions that instructional formats that are effective for low-knowledge learners could be ineffective, or even deleterious, for high-knowledge learners, and vice versa (the expertise reversal effect). The levels of learner (university students) expertise in this study were measured using an online rapid diagnostic method, a rapid verification technique, that involves presenting learners with a series of possible solution steps reflecting various stages of the solution procedure and asking them to rapidly verify the suggested steps. The results indicated a significant interaction between levels of learner expertise and instructional formats. Novice learners benefited more from static diagrams than from animated diagrams, while more knowledgeable learners benefited more from animated rather than static diagrams. A theoretical explanation of the effect is suggested within the framework of cognitive load theory.

Keywords: Animations; Learner prior knowledge; Expertise reversal effect; Diagrams; Learner-tailored instruction

A number of recent studies have demonstrated significant interactions between levels of learner expertise and instructional methods. According to the expertise reversal effect, instructional techniques and procedures that are effective for novice learners may become ineffective, or even harmful, for more experienced learners, and vice versa. The effect was
demonstrated with such diverse instructional techniques as embedding textual explanations into diagrams to reduce learner split attention; using narrated verbal explanations synchronized with animated diagrams; providing learners with detailed worked-out examples of problem-solving steps, etc. (see Kalyuga, 2005, 2006b; Kalyuga, Ayres, Chandler, & Sweller, 2003 for overviews). The major instructional implication of the effect is the need to tailor the selection of instructional methods and the design of learning environments to changing levels of learner expertise.

The effect was explained within the framework of cognitive load theory (see Van Merriënboer & Sweller, 2005, for a recent overview) by considering processing limitations of our cognitive architecture and the role of organized knowledge base in human learning and performance. Our working memory is severely limited in duration and capacity when dealing with unfamiliar information (Baddeley, 1997; Miller, 1956; Peterson & Peterson, 1959). However, in familiar domains, the available knowledge base in long-term memory allows us to encapsulate large amounts of information in larger chunks that are treated as single elements, thus reducing working memory limitations. Human learning and performance alter significantly with the development of expertise in a domain. In the absence of relevant prior knowledge, novices have to deal with many novel elements of information that may easily overwhelm working memory capacity. Therefore, without considerable external instructional support, they may experience significant cognitive overload. On the other hand, if such detailed instructional support is provided for more experienced learners, the process of reconciling the related components of their available knowledge structures in long-term memory and externally provided guidance would likely to require additional working memory resources and could also increase unnecessary cognitive load. Consequently, less capacity could be available for new knowledge acquisition and performance improvement.

Animations have long been regarded as an essential part of new computer-based instructional technologies. There are serious theoretical and cognitive arguments in favor of greater effectiveness of animated rather than static images, and corresponding principles for designing and using animation in instruction (e.g., Mayer & Anderson, 1992; Reed, 2005; Rieber, 1990; Weiss, Knowlton, & Morrison, 2002). However, existing research literature does not provide compelling and clear empirical evidence about educational advantages of animations over static graphics. Some existing reviews concluded that animations are no more (and sometimes less) effective than the equivalent static graphics (e.g., Hegarty, Kriz, & Cate, 2003; Tversky, Morrison, & Betrancourt, 2002). The transient nature of animations and limited duration and capacity of working memory could be the major reason for these failures of animated instructions to demonstrate clear advantages (Chandler, 2004; Lowe, 1999).

When learning from animations, many new elements of information may need to be processed in very limited time. New information can be held in working memory for no more than few seconds (unless rehearsed). In order to construct an integrated mental representation of novel material in working memory, a learner needs to hold information from earlier frames of the animation while attending to the following frames. If the next frame is processed before the information from the preceding frames is incorporated into an organized knowledge structure, new information may interfere with the earlier information. Cognitive demands of processing new information while holding previous information in working memory could overwhelm working memory resulting in cognitive overload. In contrast, static graphics do not create such transitivity problems, because elements of still pictures can be revisited any required number of times.
However, as the level of learner expertise in a domain increases, more domain-specific knowledge structures are acquired in long-term memory. These structures allow chunking of many elements of information into familiar units that are treated as single elements in working memory. As a result, more working memory resources become available for constructing and running dynamic mental representations. Static graphics may be less beneficial for more experienced learners than for novices because their available dynamic knowledge structures would need to be integrated and reconciled with redundant details displayed in graphics. Additional cognitive resources may be required for such processes, increasing working memory demands and reducing relative learning effects. Thus, static instructions which are more effective than animations for novices may become relatively ineffective with increased levels of learner expertise in a domain (Ayres, Kalyuga, Marcus, & Sweller, 2005). This hypothesis is in accordance with the general established role of prior knowledge as an important factor contributing to individual differences in the effect of instruction based on text and visual displays (Schnotz, 2002).

The experimental study described in this paper was designed to investigate the relationship between instructional effectiveness of animated vs static diagrams and levels of learner expertise in the task domain of transforming graphs of simple linear and quadratic functions. The tasks involved transforming graphs of the basic lines $y = x$ and $y = x^2$ into graphs of more complex lines, for example, $y = -3x + 2$ and $y = 2(x - 4)^2$ correspondingly. The tasks required application of several of the following operations: flipping a graph because of the minus sign in front of $x$ or $x^2$ (the negative slope in the case of linear function); squeezing (expanding) a graph toward (from) the $y$-axis according to the value of a coefficient in front of $x$ or $x^2$ (more or less than 1); and horizontal/vertical shifting. For evaluating levels of learner proficiency in this task domain, a novel rapid diagnostic approach was applied.

1. Rapid verification method for evaluating levels of expertise

To reliably diagnose individual levels of expertise in a specific domain, structured interviews and think-aloud protocols could be used. However, these laboratory tools are not quite suitable for online applications (partly because of technical implementation difficulties). On the other hand, simple to implement traditional multiple-choice items may have limited diagnostic capabilities (e.g., they do not provide information about actual solution steps). This study used a rapid online diagnostic technique, the rapid verification method, developed within a general rapid diagnostic approach that is based on contemporary knowledge of human cognitive architecture, especially interactions between working memory and long-term memory structures during complex knowledge-based performances (Kalyuga, 2006b, 2006c; Kalyuga & Sweller, 2004).

As was mentioned previously, knowledge structures in long-term memory define the characteristics of working memory: its content, capacity, and duration. Accordingly, a diagnostic approach based on evaluation of the extent to which working memory limits have been extended by knowledge structures held in long-term memory during problem solving could be used for evaluating levels of expertise in a domain. A possible rapid diagnostic method was suggested and validated by Kalyuga and Sweller (2004, 2005). The idea was to present learners with a problem for a limited time and to ask them to rapidly indicate their first step toward solution of the task. More knowledgeable learners presumably
should be better able to use appropriate higher-level knowledge structures for solution moves than less knowledgeable learners. Experts’ knowledge base would allow them to rapidly generate advance stages of the solution or even final answers to the tasks and skip many intermediate operations. Different first steps would indicate different levels of acquisition of corresponding knowledge structures in the learner’s long-term memory.

In an alternative approach, that is more suitable for computer-based online applications, learners could be presented with a series of possible (correct and incorrect) steps corresponding to various stages of the solution procedure, and asked to rapidly verify the suggested steps, for example, by clicking on screen buttons or pressing specified keys on the computer keyboard (e.g., correct, incorrect, or do not know). This method has been recently validated by demonstrating a high level of correlation (.75) with results of in-depth cognitive diagnosis of expertise using laboratory-based observations of the same students’ problem solving steps using video-recordings and concurrent verbal reports, with diagnostic time reduced by a factor of 3.5 (Kalyuga, submitted for publication). The method was also used for individualizing instructional procedures and optimizing levels of instructional guidance in an adaptive computer-based tutor for vector addition motion problems in kinematics (Kalyuga, 2006a).

The rapid verification test that was used in this study for evaluating levels of learner prior experience in the domain included four tasks. Two tasks asked students to transform a provided graph of the basic line \( y = x \) into graphs of more complex lines, \( y = -2x + 3 \) and \( y = 1/3x - 2 \). The following two tasks asked students to transform a provided graph of the basic line \( y = x^2 \) into graphs of more complex quadratic functions, \( y = -1/3x^2 \) and \( y = 2(x - 2)^2 \) (see Fig. 1 for an example of a task’s statement). The tasks required application of two or three of the previously listed operations.

![Graph of the line](image)

This is a graph of the line \( y = x^2 \)

Transform it into a graph of the line \( y = 2(x - 2)^2 \)

Fig. 1. Snapshot of the statement for a graph transformation task.
Each task statement was presented for 10 s (this time was established in pre-experimental trials as sufficient for comprehending task statements) and was followed by four suggested solution steps for rapid verification. Students were instructed that most of the suggested steps were supposed to be intermediate steps on a way to the solution, but some suggested steps could indicate final answers. Before the rapid test, the participants were ‘coached’ in responding sufficiently rapidly using exercises with tasks from a different area. During those exercises, the students had got a sense of what was considered a rapid response (if a student did not respond within a set short time interval, she or he was asked to respond faster next time).

Fig. 2 shows an example of an incorrect intermediate step for the task represented in Fig. 1. Some verification subtasks indicated results of the application of only one operation (e.g., expanding in Fig. 2), while other subtasks indicated results of the application of several operations. A score 1 was allocated for each correctly verified step. Null scores were allocated for incorrect responses and “Do not know” entries. The maximum possible score in the rapid test was 16. Students’ performance scores were automatically recorded by the software.

2. Experiment

In this study, the above described diagnostic tool was used for investigating the interaction between levels of learner expertise and effectiveness of animated vs static instructions.
Learners were subdivided into groups of relative “experts” and “novices” based on results of the initial rapid verification test. Students in each of these two groups were allocated to different treatments (animated vs. static diagrams). The rapid verification test was used again to measure levels of learner performance after the instructional session. The relative knowledge gains were calculated, compared, and analyzed for an expertise reversal effect.

3. Method

3.1. Participants

33 university students (18 females and 15 males, aged from 18 to 25) participated in this experiment. They were in different years of study, with most students enrolled in undergraduate courses. Participants also represented variety of subject areas, from education to mechanical engineering (most studied arts and social sciences). All students had been taught elementary mathematics courses in high school. Such courses always include graphs of linear and quadratic functions. Most of participants (especially those enrolled in arts and social sciences courses), however, did not deal regularly with such problems, and all participants had not previously encountered tasks formulated using the current format.

3.2. Materials and procedure

Based on the scores obtained in the initial rapid test, participants were allocated to one of the two groups: more knowledgeable learners (upper median group) and less knowledgeable learners (lower median group). Students in each of these two groups were randomly assigned to either a continuous animation instructional format or a static instructional format. Thus, there were four experimental groups: (1) high knowledge/animated instruction (8 students), (2) high knowledge/static instruction (8 students), (3) low knowledge/animated instruction (8 students), and (4) low knowledge/static instruction (9 students).

Participants in the animated instruction groups were presented two sequential animated instructional segments, each around one minute long. The first segment showed a worked example on how to transform a given graph of the line $y = x^2$ into a graph of the line $y = 2(x - 1)^2 - 3$. The second segment showed a worked example on how to transform a graph of the line $y = x^2$ into a graph of the line $y = -1/2(x + 3)^2 + 1$. For each transformation step, the components of the functional expressions (signs, coefficients) that determined the corresponding step were highlighted in red color. The directions of movements were indicated by highlighted arrows. The distances for shifting transformations (in numbers of linear units) were indicated next to the corresponding arrows (see Fig. 3 for a snapshot of a frame showing the expanding transformation for the second task).

The static instruction included two consecutive screens. The first screen contained four sequentially numbered (from 1 to 4) frames from the first animated segment depicting four major stages of the transformation procedure for the corresponding task. The second screen contained five numbered frames from the second animated segment depicting five major stages of the transformation procedure for the second task. Participants were requested to follow all steps in each example according to the numbered sequence. Both tasks in this condition were identical to the tasks used in the animated instruction condition.
Both conditions used only pictorial explanations and contained no on-screen or narrated text. This feature was important for the control of experimental factors, because it excluded confounding with other possible cognitive load effects associated with simultaneous processing of verbal and pictorial information in static and animated instructions, such as split-attention (spatial and/or temporal contiguity), redundancy, or modality effects (Mayer, 2005; Sweller, van Merriënboer, & Paas, 1998). Highlighted indicators of procedural steps (arrows) were embedded into the diagrams in close proximity to the corresponding diagrammatic elements to limit any unnecessary search processes that would require additional cognitive resources. Instruction times were equalized in both conditions. Static instructions were presented for exactly the same amounts of time as corresponding animated versions. The instruction phase of the experiment took 2 minutes for each condition.

Post-instruction performance levels were again measured using the rapid verification method. The procedure was similar to that used during the initial rapid test, however, the tasks in the post-instruction test were more difficult. All four tasks involved quadratic functions that required 3 (for the first two tasks) or 4 (for the remaining two tasks) transformation steps.

4. Results

A 2 (instructional format) × 2 (level of knowledge) ANOVA was conducted using experimental data. The dependent variables under analysis were knowledge gain scores as the difference between the test scores for the post-instruction test and scores for the initial test; independent variables were levels of knowledge (high/low) and format of instruction (animated/static worked examples). Means and standard deviations are given in Table 1.
An analysis of the low knowledge/high knowledge gain main effect produced a significant difference, $F(1, 31) = 4.23$, $MSe = 6.15$, $p < .05$. Low knowledge learners ($M = 1.59$, $SD = 2.15$) benefited more from the instructions than high knowledge learners ($M = -.19$, $SD = 2.79$). The relative decrease in performance in the latter case (as well as for another group below), was due to the increased complexity of the post-test tasks in comparison to the pre-test tasks. No main effect of experimental formats was found (for the animated instruction group, $M = 0.50$, $SD = 2.68$, and for the static instruction group, $M = .94$, $SD = 2.59$).

Since the main purpose of this experiment was to study the change in effectiveness of instructional formats with expertise, the interaction data of the $2 \times 2$ ANOVA between levels of expertise and instructional formats were of major interest in this study. It was expected that a difference in relative expertise in the domain would produce a change in the effectiveness of different methods of instruction. There was a significant level of interaction for the knowledge gains measured by the rapid verification testing method, $F(1, 31) = 4.25$, $MSe = 5.69$, $p < .05$, suggesting that the most effective format of instruction depends on the level of learners’ expertise.

Following the significant interaction, simple effect tests indicated that novice learners benefited more from static examples (mean knowledge gain $M = 2.56$, $SD = 1.24$) than from animated examples ($M = 0.50$, $SD = 2.51$), $F(1,15) = 4.77, MSe = 3.75, p < 0.05$. On the other hand, more knowledgeable learners benefited more from animated examples (mean gain $M = 0.50$, $SD = 3.02$) rather than from static examples (mean decrease $M = -.88$, $SD = 2.53$). Although this effect was not statistically significant, $F(1,14) = 0.97, MSE = 7.78$, Cohen’s $d$ index of 0.46 indicated a medium effect size.

The results of this experiment demonstrated an expertise reversal effect with levels of expertise determined by the rapid verification method. A theoretical explanation of the effect within the cognitive load framework takes into account duration and capacity limitations of working memory for novice learners studying animated explanations and a reduction of these limitations for more knowledgeable learners due to their available knowledge base in long-term memory. As the level of prior knowledge was raised, the performance of the animated examples group improved more than performance of the static examples group. Novice learners gained significantly more knowledge after studying static

---

Table 1
Means and standard deviations for test scores and knowledge gains by experimental conditions

<table>
<thead>
<tr>
<th>Variable</th>
<th>Experimental conditions</th>
<th>Novices</th>
<th>Experts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Animated (8)</td>
<td>Static (9)</td>
</tr>
<tr>
<td><strong>Initial test scores (max. 16)</strong></td>
<td></td>
<td>\textbf{M}</td>
<td>\textbf{SD}</td>
</tr>
<tr>
<td>$M$</td>
<td>6.63</td>
<td>6.00</td>
<td>10.63</td>
</tr>
<tr>
<td>$SD$</td>
<td>1.51</td>
<td>2.00</td>
<td>2.00</td>
</tr>
<tr>
<td><strong>Final test scores (max.16)</strong></td>
<td></td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>$M$</td>
<td>7.13</td>
<td>8.56</td>
<td>11.13</td>
</tr>
<tr>
<td>$SD$</td>
<td>2.03</td>
<td>2.13</td>
<td>3.27</td>
</tr>
<tr>
<td><strong>Knowledge gains</strong></td>
<td></td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>$M$</td>
<td>0.50</td>
<td>2.56</td>
<td>0.50</td>
</tr>
<tr>
<td>$SD$</td>
<td>2.51</td>
<td>1.24</td>
<td>3.02</td>
</tr>
</tbody>
</table>

---

Please cite this article in press as: Kalyuga, S., Relative effectiveness of animated and static ..., Computers in Human Behavior (2007), doi:10.1016/j.chb.2007.02.018
worked examples. For more knowledgeable learners, there was an indication of more benefits from studying animated instructional explanations than from studying static examples.

5. Discussion

The experiment described in this paper was designed to investigate the relation between levels of learner expertise and effectiveness of animated and static instructions in the domain of graph transformation tasks. The rapid diagnostic procedure was used to establish which instructional procedure, animated or static diagrams, should be used for students with differing levels of prior knowledge. According to the cognitive load hypothesis, continuous animations may be too cognitively demanding for novice learners due to a high degree of transitivity of these visualizations, on the one hand, and limited capacity and duration of working memory, on the other hand.

Therefore, it was expected that less knowledgeable learners would benefit more from a set of equivalent static diagrams. However, animations could be relatively more beneficial for more knowledgeable learners who have already acquired a sufficient knowledge base for dealing with issues of transitivity and limited working memory capacity. Such an interaction would represent an example of the expertise reversal effect in relation to these instructional formats. The results confirmed the hypothesis by indicating a statistically significant interaction between levels of learner expertise and instructional formats. Novice learners benefited more from static instructions, while more knowledgeable learners benefited more from animated instructions.

An adaptive instructional procedure that allows learners to progress from static to animated diagrams as levels of their knowledge increase could create more efficient learning environments than using a single format. Rapid online tests of task-specific expertise could be important components of such environments. This experiment demonstrated that the rapid verification method was sufficiently precise to allow separating learners in groups according to levels of their prior knowledge and detecting an expertise reversal effect. Another study in the domain of kinematics (Kalyuga, 2006a) indicated that the method could also be applied for the dynamic selection of appropriate levels of instructional guidance and procedures that are optimal for learners with different levels of expertise. In that study, the method was used to build learner-adapted faded worked examples in a tutor based on online monitoring of current levels of learner task-specific proficiency. The adaptive approach proved to be superior to a traditional non-adapted instruction. This study was limited to a rather narrow task domain associated with well-structured procedural tasks, and involved relatively brief instructional sessions. In further research, the generality and limits of usability of the results, especially in poorly structured and more extensive task domains need to be established. More comprehensive studies should be conducted comparing different rapid diagnosis-based adaptive methodologies involving animated and static instructional procedures with different levels of interactivity and learner control over the pace of instruction.

Our current knowledge of human cognitive architecture, its limitations and strengths may allow us to better understand conditions under which animations and static graphics enhance learning outcomes. This paper argues that tailoring these instructional formats to learners with different levels of expertise should be an essential part of efficient learning environments.
References


Kalyuga, S. (submitted for publication). When less is more in cognitive diagnosis: A rapid online method for diagnosing learner task-specific expertise.


