British Journal of Educational Psychology (2005), 75, 71–85

The British Psychological Society

www.bpsjournals.co.uk

The management of cognitive load during complex cognitive skill acquisition by means of computer-simulated problem solving

Liesbeth Kester*, Paul A. Kirschner and Jeroen J. G. van Merriënboer

Educational Technology Expertise Centre, Open University of the Netherlands, Heerlen, The Netherlands

This study compared the effects of two information presentation formats on learning to solve problems in electrical circuits. In one condition, the split-source format, information relating to procedural aspects of the functioning of an electrical circuit was not integrated in a circuit diagram, while information in the integrated format condition was integrated in the circuit diagram. It was hypothesized that learners in the integrated format would achieve better test results than the learners in the split-source format. Equivalent-test problem and transfer-test problem performance were studied. Transfer-test scores confirmed the hypothesis, though no differences were found on the equivalent-test scores.

Since the late 1980s, many instructional design guidelines have been generated from cognitive load theory (CLT; Sweller, 1988). One of the major pillars of this theory is that working memory is limited (Baddeley, 1992; Miller, 1956). Effective and efficient instructional material avoids overloading working memory, so as not to hamper the learning process, attempting to redirect learners' attention to relevant learning processes.

CLT distinguishes three types of cognitive load: intrinsic, extraneous, and germane. According to Sweller, van Merriënboer, and Paas (1998), intrinsic cognitive load is inherent to the learning task itself, and, '. . . cannot be directly influenced by instructional designers . . .' (p. 262). Extraneous and germane load refer to the cognitive load that arises when learners interact with the instructional material, and can be influenced by instructional designers. Extraneous cognitive load includes all cognitive load associated with the processes a learner engages in, while interacting with the instructional material that is not beneficial for learning. Examples of activities which induce extraneous cognitive load are mentally integrating different sources of

*Correspondence should be addressed to Dr Liesbeth Kester, Educational Technology Expertise Centre, Open University of The Netherlands, P.O. Box 2960, 6401 DL Heerlen, The Netherlands (e-mail: liesbeth.kester@ou.nl).

DOI:10.1348/000709904X19254

information, for example, separate information in a figure and text, or searching for relevant information in order to understand the subject matter. Germane cognitive load includes all cognitive load associated with processes a learner engages in while interacting with the instructional material that is beneficial for learning. Variability of learning tasks, for example, may stimulate learners to construct better cognitive schemata (Spiro, Coulson, Feltovich, & Anderson, 1988; Sweller *et al.*, 1998).

The focus of the study reported here is on decreasing the extraneous cognitive load of the instructional material by avoiding the split attention which arises when a learner has to mentally integrate two sources of information in order to understand the learning material, for example a picture and its explanatory text. Chandler and Sweller (1992) called this the split attention effect (SAE). The idea is that spatially integrating two sources of information in the learning material relieves the learners of mentally integrating these information sources themselves before learning from the material can take place.

The notion of the SAE originates from research on worked examples, which is also an effective way to reduce extraneous cognitive load. For novice learners, studying worked examples facilitates learning, as compared to solving conventional problems (Atkinson, Derry, Renkl, & Wortham, 2000; Cooper & Sweller, 1987; Sweller & Cooper, 1985). According to Tarmizi and Sweller (1988), conventional problems necessitate a means–ends search, which is a weak problem-solving activity with a considerable claim on cognitive resources, while worked examples enable the learner to concentrate on the problem states and their solution steps, leading to better schema acquisition. In their research on worked examples, they were the first to show that in some cases the beneficial effects of worked examples on schema acquisition failed to occur. A closer examination of these worked examples showed that they all contained separate sources of information that needed to be mentally integrated in order to understand the worked example; for example, a geometry example that made use of a diagram and explanatory text. It was hypothesized that the mental integration of these separate sources of information, that is, the diagram and the explanatory text, used up available cognitive resources at a cost to the learning process, and that the integrated presentation of these separate sources of information would enlarge the cognitive resources available for learning, and would thus be beneficial for schema acquisition. Tarmizi and Sweller (1988), Ward and Sweller (1990), and Sweller, Chandler, Tierney, and Cooper (1990) all showed that worked examples without split-source information led to better learning results than split-source worked examples and conventional problems.

The findings from worked-example research resulted in extensive investigation concentrating on this SAE (Cerpa, Chandler, & Sweller, 1996; Chandler & Sweller, 1991, 1992, 1996; Kalyuga, Chandler, & Sweller, 1999; Mayer & Anderson, 1992; Mayer & Moreno, 1998; Mayer & Sims, 1994; Sweller & Chandler, 1994). In this research, the focus shifted from split-source information in worked examples, to split-source information in more general learning material, such as pictures and text in an instructional textbook. In domains such as electrical training programmes (Chandler & Sweller, 1991) and computer or machine programming (Cerpa *et al.*, 1996; Chandler & Sweller, 1992, 1996; Sweller & Chandler, 1994), conventional learning material was compared to integrated learning material. Both learning materials contained identical information. In the conventional material, mutually referring sources of information were presented separately, for example, diagrams and explanatory text, while in the integrated material all mutually referring sources of information were integrated; again, for example, diagrams and explanatory text. After studying the learning material, the participants received both theoretical questions and a practical test. The test results favoured the integrated learning material, provided that the mutually referring information sources were unintelligible without mental integration. Instead of integrating two sources of visual material, some researchers choose to integrate animation and narration (Mayer & Anderson, 1992; Mayer & Moreno, 1998; Mayer & Sims, 1994), or auditory information and diagrams (Kalyuga *et al.*, 1999). This dual-mode of presenting mutually referring material also resulted in enhanced test performance.

In the study presented here, a new instructional approach was chosen to examine the SAE. The learning period in this study consisted of a series of conventional practice problems, such as computer-simulated malfunctioning electrical circuits, accompanied by just-in-time presented information. According to the just-in-time information presentation model (Kester, Kirschner, van Merriënboer, & Bäumer, 2001; van Merriënboer, Kirschner, & Kester, 2003), this means that the conventional problems are preceded by supportive information and accompanied by procedural information, both necessary to solve the problem. In this setting, split attention can be avoided by integrating the procedural information and the conventional practice problems. The question is whether avoiding split attention in this situation also leads to better test results, compared with avoiding split attention in worked examples or instructional textbooks.

The supportive information presented before practice was useful, explanatory information. For example, information was presented explaining what an electrical circuit is, and how electrons flow through a circuit, using a central heating system as a metaphor. This information enables the learners to construct general, abstract schemata of electrical circuits through elaboration. In other words, the supportive information is gradually coupled to already existing, relevant cognitive schemata in long-term memory. During practice, these cognitive schemata are modified and refined, resulting in more appropriate schemata given the experiences. This process is called induction (Proctor & Reeve, 1988; Holland, Holyoak, Nisbett, & Thagard, 1986). Elaboration and induction of supportive information yields cognitive schemata that contain general heuristic knowledge, rather than domain-specific algorithmic knowledge, that is particularly useful when learners have to deal with unfamiliar problem aspects.

The procedural information enables the learners to recognize elements to perform actions, and directly refers to the circuit. For example, ‘this is a voltmeter’ (next to the symbol of a voltmeter), or ‘a voltmeter has to be connected in parallel because current cannot flow through this meter’. This information enables the learners to form domain-specific automated schemata or production rules, through knowledge compilation, the compilation of declarative knowledge either from long-term memory or from the outside world, into procedural knowledge. During practice these domain-specific automated schemata gain strength every time they are applied successfully. This process is called strengthening (Anderson, 1982, 1996). Knowledge compilation, and strengthening of procedural information, yields automated schemata that contain domain-specific algorithmic knowledge, which is particularly useful when learners have to deal with familiar problem aspects.

Prior research has shown that the presentation of supportive information before practice, and procedural information during practice, substantially decreased searching behaviour irrelevant for learning (Kester, Kirschner, & van Merriënboer, 2004). The research reported here focuses on whether the instructional material could be further improved, by integrating the procedural information in the malfunctioning circuits, so that the learner’s necessity to mentally integrate split-source material is reduced. It is assumed that the presentation of supportive information before practice, and integrated procedural information during practice, frees up cognitive capacity that can be allocated to relevant learning processes such as elaboration, induction, and knowledge compilation. This will benefit the construction and refinement of general schemata and the forming of domain-specific automated schemata.

The effectiveness of the learning material is measured by test performance. Test problems can be distinguished into two types. Equivalent-test problems, that are similar to the practice problems, consisted of malfunctioning circuits containing the same elements that were used in the practice problems, but in new configurations. Transfer-test problems, that are more or less dissimilar to the practice problems, consisted of malfunctioning circuits containing both the same, and novel, elements in new configurations. In equivalent-test problems, the content learned during practice has to be applied in similar problem situations during the test, while in transfer-test problems the content learned during practice has to be applied in different problem situations during the test. It is assumed that solving the equivalent-test problems relies more on the domain-specific automated schemata than on the general schemata, because these problems only have familiar aspects. Solving the transfer-test problems relies more on the general schemata, than on the domain-specific automated schemata, because these problems introduce unfamiliar aspects.

In this study, two information presentation formats are compared. The supportive information is presented before practice and the procedural information during practice, in both formats. In the split-source format, the procedural information is not integrated in the malfunctioning circuit, while in the integrated format the procedural information is integrated in the circuit. It is hypothesized that learners in the integrated format achieve higher test scores than learners in the split-source format, because the SAE is avoided in the integrated format. As a result, in the integrated format more cognitive capacity is available during practice for the learning processes, induction, and knowledge compilation. This could facilitate both the refinement of general schemata needed to solve the transfer-test problems, as well as the forming of automated schemata, needed to solve the equivalent-test problems.

Method

Participants

Twenty-five fourth-year high school students at Sintermeerten College in Heerlen, The Netherlands (16 male, 9 female; mean age $= 16.2$ years, $SD = 0.72$) participated in this study. All of the participants spoke Dutch as their first language, the language in which the instruction was given. They voluntarily participated in a physics simulation on electrical circuits. No specific grade was given for participation. All participants followed the same physics curriculum beginning in their third year. They were all equally familiar with the topic of the physics simulation because they had all received the theory a year before. They received €10 for their participation.

Materials

Physics simulation

Crocodile Physics®, a simulation program for secondary school science classes, was used to develop the physics simulation for this experiment. The computer-based simulation contained an introduction and nine practice troubleshooting problems for faulty electrical circuits, and was followed by 10 test problems. In the introduction the participants received information on: (1) *what to expect*: the number of problems, available time and how to switch the circuit on and off; (2) *how to navigate within the simulation*: left and right arrows were used to go back or forth in the simulation, clicking on different icons allowed participants to jump to a practice problem or a test problem; and (3) *the experimental rules*: taking notes or changing the computer's configuration (e.g. changing the full screen presentation to part screen, making changes in the menu of Crocodile Physics®) was not allowed, and the work had to be done individually and independently.

The troubleshooting problems consisted of malfunctioning electrical circuits. Practice problems were preceded by supportive information on the working of electrical circuits, and accompanied by procedural information. The test problems were presented without this information. Inherent to a malfunctioning circuit is that elements (e.g. lamps) become irreversibly damaged after one try (i.e. they explode). To allow the participants a good look at what happens in a circuit when the switch is turned on, a 'repeat-button' was available to allow the participants to review the events in the circuit after turning the switch on. The use of the 'repeat-button' was unlimited. After manipulating the malfunctioning circuit, the participants had to provide a description of the problem, a diagnosis of the cause of the problem, and a solution for the problem.

Information presentation

Two information presentation formats were distinguished. The participants were randomly assigned to one of the two formats. The computers in the computer room were configured with one of the two formats so that both formats were divided per row, over the available computers. The monitors were shut down and the participants blindly choose a computer to work on.

The participants in both formats received the supportive information before practice, and the procedural information during practice. In the split-source format ($N = 13$) the procedural information was not integrated in the malfunctioning circuit. The circuit was presented on the left side of the screen, and the procedural information on the right side (see Fig. 1). In the integrated format ($N = 12$) the procedural information was integrated in the circuit (see Fig. 1). The amount and the content of the information in the integrated format was the same as in the split-source format, but it was cut up in chunks that referred to specific elements or features of the circuit. These chunks were positioned as close as possible to the elements or features of the circuit they were referring to in order to avoid split attention. The circuits and the accompanying information in the split-source format took up all the space on the computer screen, so for comparability, the circuits in the integrated format were enlarged to cover the screen, which allowed for convenient arrangement of the referring information chunks in the circuit.

Practice problems

Nine practice problems were administered to the participants. The malfunctioning circuits in the practice problems made use of six different elements: a toggle switch, a lamp, a battery, a resistor, a voltmeter; and an ammeter. All practice circuits contained one or more of each of these six elements in different configurations. Generally, three

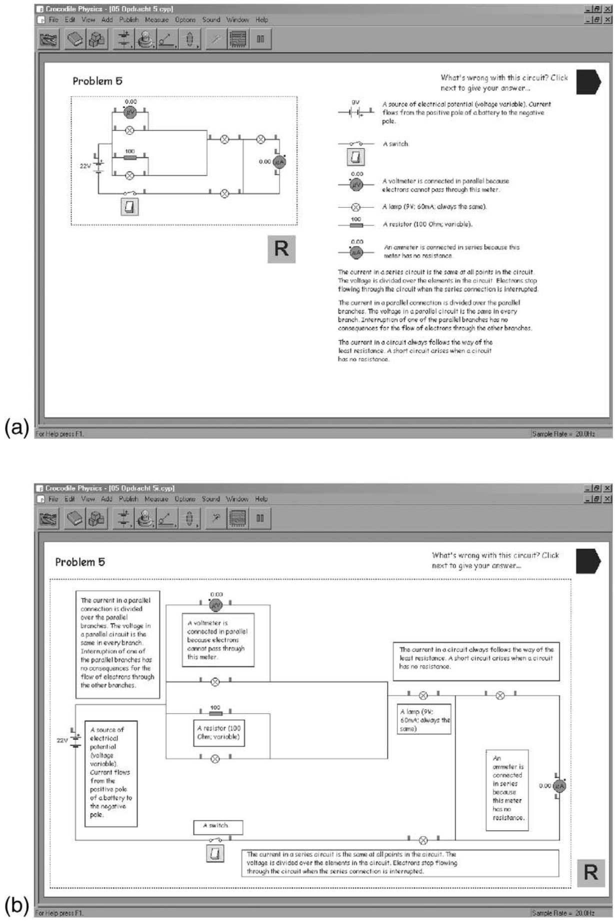

Figure 1. An example of a practice problem in the split-source (a) and the integrated format (b).

problems could occur in these circuits, either alone or in combination, namely (1) too high voltage, (2) wrongly connected elements, and/or (3) a short circuit.

During practice, the participants could obtain a maximum of 27 points by giving a problem description, a problem cause, and a problem solution for the malfunctioning circuits. Participants received 1 point for a correct problem description, 1 point for a correct diagnosis of the problem cause, and 1 point for a correct problem solution. For example, in one of the practice problems (see Fig. 2), the inserted battery is too powerful for the elements in the circuit, and a voltmeter is incorrectly connected. In this problem the following correct responses could be made: (1) *the problem description*: the lamp in series and the bottom lamp in parallel explode ($=\frac{1}{2}$ point), the top lamp in parallel does not work at all ($=\frac{1}{2}$ point; see Fig. 2); (2) *the problem cause*: the power supply (i.e. the battery) is too strong for both the lamp in series and for the bottom lamp in parallel ($=\frac{1}{2}$ point), the top lamp in parallel does not work because the voltmeter is connected in series to this lamp ($=\frac{1}{2}$ point); and (3) *the problem solution*: insert a weaker battery or add elements in series ($=\frac{1}{2}$ point; note that equivalent answers do not yield extra points), connect the voltmeter in parallel ($=\frac{1}{2}$ point).

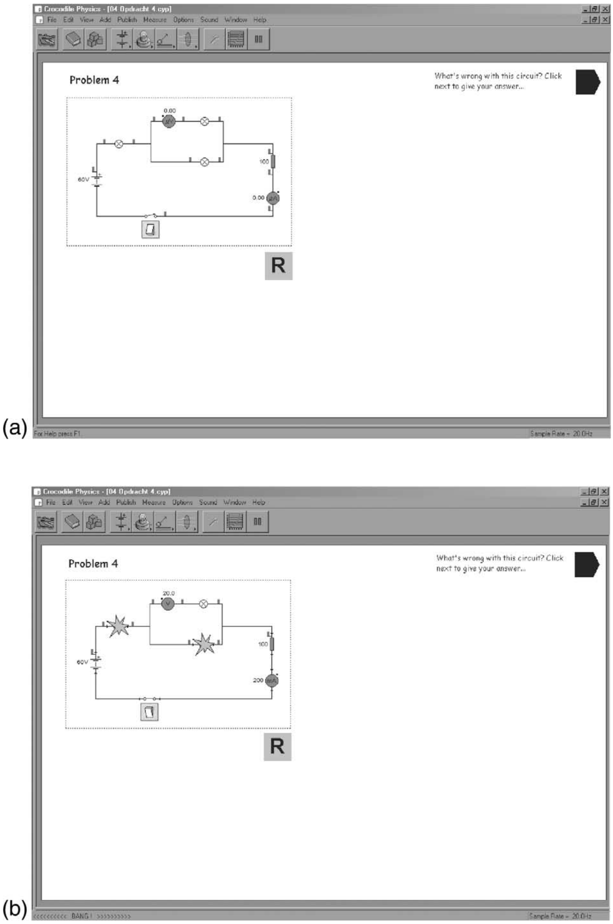

Figure 2. An example of a practice problem before (a) and after (b) closing the toggle switch.

The practice performance scores of 10 participants were determined by two raters. The inter-rater reliability for practice performance was .74 (intra-class correlation coefficient). The internal consistency of the practice items was analysed and resulted in a reliability of .78 (Cronbach's α).

Test problems

After the nine troubleshooting practice problems, 10 troubleshooting test problems were administered to the participants. The test problems also consisted of malfunctioning electrical circuits designed in Crocodile Physics®, but without the accompanying information. Five equivalent-test problems consisted of malfunctioning circuits that were novel, yet equivalent, to the circuits that were used in the practice problems. These equivalent-test problems made use of the same elements that were used during practice: a toggle switch; a lamp; a battery; a resistor; a voltmeter; and an ammeter, but in different configurations. Five transfer-test problems consisted of malfunctioning circuits that contained, in addition to these elements, one or two new elements, namely: a variable resistor; a fuse, a LED, a buzzer, and push-to-make switch or a motor and gears, in different configurations.

Again, the participants had to give a problem description, a problem cause, and a problem solution for each test problem. The equivalent-test problems were meant to determine primarily whether the participants had formed adequate domain-specific automated schemata that allowed them to solve the familiar aspects of these equivalent problems. The transfer-test problems were meant to measure whether the participants had constructed general schemata that allowed them to solve the unfamiliar aspects of these transfer problems. The participants could obtain a maximum of 15 points for the five equivalent problems and 15 points for the five transfer problems. The scoring procedure was equivalent to the practice problems. The total test performance scores of 10 participants were determined by two raters. The inter-rater reliability for the total test performance was .86 (intra-class correlation coefficient). The internal consistencies of the equivalent problems and the transfer problems were analysed separately and resulted in a reliability of .72 and .70 (Cronbach's α), respectively.

Mental effort measurement

Mental effort refers to the cognitive capacity actually allocated to meet the problem requirements. The amount of mental effort invested by participants is considered to be the essence of cognitive load; therefore, mental effort is used as an index for cognitive load (Paas, 1992). In instructional research, and especially in research based on CLT, a 9-point rating scale is used to measure mental effort (Paas, 1992; Paas, van Merriënboer, & Adam, 1994). Here, for reasons of comparability, the same scale was used. The mental effort measures ranged from very low, very high, to very very high, mental effort. The rating scale was administered during practice and during the test, directly after each problem. The following question was posed to the participants after each practice and test troubleshooting problem: 'How much mental effort did it cost you to solve the problem(s) in the preceding circuit?' Moreover, a separate mental effort measurement was taken with regard to the subject matter after the nine practice problems. The participants had to answer the question: 'How much mental effort did it cost you to understand all subject matter?' The internal consistency of the mental effort measures was analysed and resulted in a reliability of .71 (Cronbach's α) for the nine practice problems, .70 for the five equivalent-test problems, and .76 for the five transfer-test problems.

Log tool

A logging program was developed especially for this experiment to keep track of the time on task, and of the navigation of the participants through the physics simulation. This program generates a text file with a list of window headers coupled to a timestamp. The route the participants followed through the simulation, and the time it took them to complete it, could be determined.

Procedure

Participants received an oral instruction which stressed that they had to work independently, adhere to the time limit, work seriously, carefully study the supportive information on the working of electrical circuits, and not ask questions during the experiment. It was emphasized that they were not allowed to skip any part of the answer (description, cause or solution), even if they did not know the answer. In that case they were required to give the answer 'no answer', or 'do not know'. It was made clear that, after finishing the physics simulation, their responses would be checked with respect to omissions. They were told that the aim of the experiment was to find out whether it is useful to integrate this kind of simulation software in the curriculum, and, if this is the case, how this should be done. Before the participants could actually start with the physics simulation they were 'walked through' an example of a troubleshooting problem by the experimenter. Note that this example did not contain either supportive or procedural information, and was only meant to show the participants how to use the navigation buttons, and how to start the simulation, namely by pushing the switch. During this example they could ask questions. This 'walk through' functioned to assure that the procedure and working of the simulation program was clear to all of the participants before the actual experiment started.

All participants had 2 hours to work through the introduction in class and complete the simulation with the supportive information, the practice problems, and the test problems. After the plenary introduction the participants went through the supportive information, the practice, and the test at their own pace, but in fixed order. Participants could not go back to the supportive information once they started the practice problems, and could not go back to the practice problems once they started the test problems. They could move around unrestrictedly within the screens containing the supportive information, the practice problems, or the test problems. During the physics simulation the behaviour of the participants on the computer, and the time spent on each problem, was logged.

Results

Practice problems

For an overview of the practice results see Table 1.

Table 1. Overview of the mean practice results

	Information presentation format					
	Split-source			Integrated		
Variable	*M*	*SD*	*N*	*M*	*SD*	*N*
			Practice scores (max. = 30)			
Practice problems	12.35	3.98	13	12.63	2.53	12
			Mental effort			
Practice problems	4.97	0.93	13	4.94	0.83	12
Subject matter	5.00	1.22	13	4.58	1.38	12
			Time on task			
Supportive information	7.42*	4.81	13	4.19*	1.47	12
Practice problems	30.67	11.66	13	31.34	7.51	12
			Repeats			
Practice problems	36.69	11.97	13	28.5	14.54	12

*$p < .05$.

Practice scores

An α level of .05 was used for all statistical tests. A t test showed no significant differences between the split-source and integrated format on practice performance.

Mental effort

The mean mental effort during practice and the mental effort considering subject matter were considered separately. The *t* tests showed no significant differences between the experimental groups on these two measures.

Time on task

Two time measures were analysed, namely: time spent on the supportive information on the working of electrical circuits, and time spent on practice. A *t* test (equal variances not assumed) showed a significant difference between groups on time spent on the supportive information $t(14, 39) = 2.31$; $p < .05$; $d = -.93$. The participants in the integrated format ($M = 4.19$ minutes; $SD = 1.47$) spent less time on the supportive information than participants in the split-source format ($M = 7.42$ minutes; $SD = 4.81$). For an overview of the mean time on task see Table 1.

Repeats

The number of times the participants used the repeat-button was counted. A *t* test showed no significant differences between the experimental groups on use of the repeat-button during practice.

Test problems

For an overview of the test results see Table 2.

Table 2. Overview of the mean test results

	Information presentation format					
	Split-source			Integrated		
Variable	*M*	*SD*	*N*	*M*	*SD*	*N*
			Test scores (max. = 15)			
Equivalent-test problems	5.91	3.19	13	6.10	1.95	12
Transfer-test problems	4.19*	1.81	13	5.96*	2.64	12
			Mental effort			
Equivalent-test problems	5.25	1.08	13	4.72	0.99	12
Transfer-test problems	4.91	1.17	13	5.32	1.34	12
			Time on task			
Equivalent-test problems	8.64	2.36	13	10.91	2.89	12
Transfer-test problems	9.61*	2.76	13	11.46*	5.54	12
			Repeats			
Equivalent-test problems	14.46	4.03	13	12.92	5.13	12
Transfer-test problems	12.09	2.39	11	11.33	5.16	12

*$p < .05$.

Test scores

The equivalent-test problems and the transfer-test problems were considered separately. A *t* test (equal variances assumed) showed a significant difference between the experimental groups on the performance on the transfer-test problems, $t(23) = -1.94$;

$p < .05$; $d = .78$. Participants in the integrated format ($M = 5.96$; $SD = 2.64$) performed better on these test problems than participants in the split-source format ($M = 4.19$; $SD = 1.81$). No significant difference was found for the performance on the equivalent-test problems.

Mental effort

A *t* test showed no significant differences between the experimental groups on mental effort on the equivalent-test problems or on the transfer-test problems.

Time on task

A *t* test (equal variances assumed) revealed a significant difference between the experimental groups on the time the participants spent on the transfer-test problems $t(23) = -2.15$; $p < .05$; $d = -.86$. The participants in the split-source format ($M = 9.61$ minutes; $SD = 2.76$) spent less time on solving these test problems than the participants in the integrated format ($M = 11.46$ minutes; $SD = 5.54$). No significant differences were found between the experimental groups on the time the participants spent on the equivalent-test problems.

Repeats

The number of times the participants used the repeat-button was counted. A *t* test showed no significant differences for the equivalent and the transfer-test problems between the experimental groups on use of the repeat-button.

Discussion

The hypothesis that learning material presented in an integrated way leads to better transfer-test results than split-source presented learning material was supported in this study. Participants receiving the procedural information integrated in the malfunctioning electrical circuit performed better on the transfer problems, that is, problems that contained a new element, than participants receiving the procedural information not integrated in the electrical circuit.

No differences were found between the experimental groups on the equivalent-test problems. However, the relative decrease between equivalent-test scores and transfer-test scores for the split-source group (29%), is much larger than this decrease for the integrated group (2%). The split-source group performed worse on the transfer-test, while the integrated group performed almost equally well on both tests. Both groups seem to apply what they have learned equally well on familiar test problems, with participants in the integrated group more able to transfer their knowledge to new situations. Apparently, the participants in both groups were able to form adequate domain-specific automated schemata, but only the participants in the integrated format were able to construct the general schemata necessary to solve the transfer problems. These results can be explained by taking a closer look at the mental effort results.

The mean mental effort reported by the participants during practice in both formats never exceeded the rating 'not low, not high'. This indicates that all participants had enough cognitive capacity available to learn from the practice problems, and that they all invested the same average amount of mental effort. What is important is how the participants in both formats allocated the mental effort they invested. Participants in the split-source format had to divide their cognitive capacity during practice over learning processes (i.e. induction and knowledge compilation), and mental integration processes (i.e. integrating text and diagram), while participants in the integrated format could allocate it all to the relevant learning processes. As expected, this led to better transfer-test scores. However, differences in equivalent-test scores failed to occur. It is assumed that since the practice problems were product oriented (i.e. participants had to give a description, cause and solution of the problem), the participants in both formats were forced to apply the procedural information during practice. If they did not, they could not come up with an answer. This allowed for knowledge compilation and formation of domain-specific automated schemata in both formats. In the split-source format, however, the remaining invested mental effort had to be divided over mental integration processes and induction, while the remaining cognitive capacity in the integrated format was all available for induction. Participants in the latter format, therefore, had more cognitive capacity left to refine their general schemata, than participants in the split-source format. This explains why participants of the split-source format and the integrated format did not differ on equivalent-test performance, but did differ on transfer-test performance.

The participants in the integrated group spent less time reading the supportive information, and more time on solving the transfer-test problems, than the participants in the split-source group. It remains unclear why the participants in the integrated group spent less time on studying the supportive information, because at that point in the experiment no manipulation had yet taken place. The supportive information for both groups is exactly the same, and is presented exactly at the same point in the simulation. Nevertheless, although participants in the integrated group spent less time on the supportive information presented before practice, this had no influence on practice performance and equivalent-test problem performance; both groups performed equally well on these measures. Moreover, participants in the integrated format performed better on the transfer-test problems than the other participants.

The difference in time on task spent on the transfer-test problems can be easily explained by the fact that it takes more time to fill in the answers for the problems presented, instead of simply stating ‘no answer’ or ‘do not know’. Yet, when this point is left aside and the relative increase in time of the integrated group is compared to the relative increase in transfer-test performance of this group, it is seen that a small increase in time investment leads to a large increase in performance. In other words, it is unlikely that the participants in the integrated group performed better on the transfer-test than the other participants only because they invested more time during the transfer-test.

Based on the results of this study, it seems that generalization of the SAE to the context of conventional problem solving in computer-based simulations is possible. Beneficial effects on learning, comparable to those shown in earlier split-attention research (Cerpa *et al.*, 1996; Chandler & Sweller, 1991, 1992, 1996; Kalyuga *et al.*, 1999; Mayer & Anderson, 1992; Mayer & Moreno, 1998; Mayer & Sims, 1994; Sweller & Chandler, 1994), are also found in this study. This study emphasizes the importance of integrating referring material in a context other than worked examples. The results of this study are particularly important in an era in which instructional theories focus on real-life problems as the driving force for learning a complex cognitive skill (Reigeluth, 1999). These real-life, conventional problems, are often too complex and too cognitively demanding for novice learners, so all measures that can be taken to properly manage the cognitive load during problem solving are beneficial for skill acquisition. Besides the importance of integrating referring material as shown here, there are several other ways to properly manage cognitive load. First, working on real-life tasks can be scaffolded by sequencing learning tasks during practice from simple to complex (Reigeluth, 1999; Reigeluth & Stein, 1983), or by using task formats that need less cognitive capacity, such as worked-out examples (Paas & van Merriënboer, 1994), or completion tasks (van Merriënboer & de Croock, 1992). Second, information necessary to carry out the real-life tasks can be presented just-in-time; in other words precisely when the learner needs it to meet the task requirements (Kester *et al.*, 2001). Future research should further study on how all these measures can complement each other in supporting learners during complex cognitive skill acquisition, by means of realistic problem solving. Moreover, since cognitive load is such a crucial factor for complex cognitive skill acquisition, special care should be taken to the measurement of cognitive load reflected by mental effort. Although the subjective measurements of mental effort often have a high reliability and a sufficient face validity, the actual validity remains the subject of discussion (Gopher & Donchin, 1986). Future research should examine other methods of measuring mental effort such as psychophysiological measures, for example, pupillary diameter, or specific measures (e.g. secondary tasks).

In conclusion, this study shows that in the context of conventional problem solving, the integrated presentation of procedural information and a conventional problem leads to better transfer-test performance than the non-integrated presentation of this information and a conventional problem. Apparently, the integrated presentation of information and problem is especially beneficial for general schema construction, provided that the problem does not exceed cognitive capacity, so that domain-specific schema automation can occur in both the integrated, and the not integrated instructional material.

Acknowledgements

The authors would like to thank Dr. I. Hamers, Sintermeertencollege, Heerlen, The Netherlands for his help in organizing and setting up this experiment.

References

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, *89*(4), 369–406.

Anderson, J. R. (1996). ACT: A simple theory of complex cognition. *American Psychologist*, *51*, 355–365.

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research*, *70*, 181–214.

Baddeley, A. D. (1992). Working memory. *Science*, *255*, 556–559.

Cerpa, N., Chandler, P., & Sweller, J. (1996). Some conditions under which integrated computer-based training software can facilitate learning. *Journal of Educational Computing Research*, *15*(4), 345–367.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, *8*(4), 293–332.

Chandler, P., & Sweller, J. (1992). The split attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, *62*, 233–246.

Chandler, P., & Sweller, J. (1996). Cognitive load while learning to use a computer program. *Applied Cognitive Psychology*, *10*, 151–170.

Cooper, G., & Sweller, J. (1987). The effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology*, *79*, 347–362.

Gopher, D., & Donchin, E. (1986). Workload: An examination of the concept. In K. R. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of perception and human performance, Vol. 2: Cognitive processes and performance* (pp. 41-1–41-49). New York: Wiley.

Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. R. (1986). *Induction: Processes of inference, learning, and discovery.* Cambridge, MA: MIT Press.

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology*, *13*(4), 351–371.

Kester, L., Kirschner, P. A., & van Merriënboer, J. J. G. (2004) Information presentation and troubleshooting in electrical circuits. *International Journal of Science Education*, *26*, 239–256.

Kester, L., Kirschner, P. A., van Merriënboer, J. J. G., & Baumer, A. (2001). Just-in-time information presentation and the acquisition of complex cognitive skills. *Computers in Human Behavior*, *17*, 373–391.

Mayer, R. E., & Anderson, R. B. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology*, *84*(4), 444–452.

Mayer, R. E., & Moreno, R. (1998). Split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology*, *90*(2), 312–320.

Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology*, *86*(3), 389–401.

Miller, G. A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.

Paas, G. W. C. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, *84*(4), 429–434.

Paas, G. W. C., & van Merriënboer, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive load approach. *Journal of Educational Psychology*, *86*, 122–133.

Paas, G. W. C., van Merriënboer, J. J. G., & Adams, J. J. (1994). Measurement of cognitive load in instructional research. *Perceptual and Motor Skills*, *79*, 419–430.

Proctor, R. W., & Reeve, T. G. (1988). The acquisition of task-specific productions and modification of declarative representations in spatial-precueing tasks. *Journal of Experimental Psychology: General*, *117*, 182–196.

C. M. Reigeluth (Ed.) (1999). *Instructional design theories and models: A new paradigm of instructional theory, Vol. 2*. Mahwah, NJ: Lawrence Erlbaum.

Reigeluth, C. M., & Stein, F. S. (1983). The elaboration theory of instruction. In C. M. Reigeluth (Ed.), *Instructional-design theories and models: An overview of their current status* (1st ed.) (pp. 335–381). Lawrence Erlbaum Mahawah, NJ.

Spiro, R. J., Coulson, R. L., Feltovich, P. J., & Anderson, D. K. (1988). *Cognitive flexibility theory: Advanced knowledge acquisition in ill-structured domains* (Technical Report No. 441). Champaign, IL: University of Illinois, Center for the Study of Reading.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*, 257–285.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction*, *12*(3), 185–233.

Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load as a factor in the structuring of technical material. *Journal of Experimental Psychology: General*, *119*(2), 176–192.

Sweller, J., & Cooper, G. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction*, *2*, 59–89.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*(3), 251–296.
Tarmizi, R. A., & Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Education Psychology*, *80*(4), 424–436.
van Merriënboer, J. J. G., & de Croock, M. B. M. (1992). Strategies for computer-based programming instruction: Program completion vs. program generation. *Journal of Educational Computing Research*, *8*, 365–394.
van Merriënboer, J. J. G., Kirschner, P. A., & Kester, L. (2003). Taking the load of a learner's mind: Instructional design for complex learning. *Educational Psychologist*, *38*(1), 5–13.
Ward, M., & Sweller, J. (1990). Structuring effective worked examples. *Cognition and Instruction*, *7*(1), 1–39.

Received 11 February 2003; revised version received 8 October 2003