Just-in-time, Schematic Supportive Information Presentation During Cognitive Skill Acquisition.

Open Universiteit

www.ou.nl

Running head: JUST-IN-TIME, SCHEMATIC SUPPORTIVE INFORMATION

PRESENTATION

Just-in-time, Schematic Supportive Information Presentation During Cognitive Skill Acquisition

Liesbeth Kester[1]

Educational Technology Expertise Center,

Open University of the Netherlands, Heerlen, the Netherlands

Chris Lehnen

Faculty of Psychology

University of Maastricht, the Netherlands

Pascal W.M. Van Gerven

Department of Neurocognition

University of Maastricht, the Netherlands

Paul A. Kirschner

Educational Technology Expertise Center,

Open University of the Netherlands, Heerlen, the Netherlands

[1] Correspondence concerning this article should be addressed to Liesbeth Kester, Open

University of the Netherlands, Educational Technology Expertise Center, P.O. Box 2960, 6401

DL Heerlen, The Netherlands. Telephone: +31 (0)45-5762428. Fax: +31 (0)455762802. Email:

liesbeth.kester@ou.nl

Abstract

Cognitive load theory states that well-designed learning material minimizes extraneous cognitive load and optimizes germane cognitive load within the thresholds of available cognitive resources. In this study, the extraneous cognitive load is minimized by avoiding temporal split attention with regard to supportive information (i.e., conceptual models or 'theory') and the germane cognitive load is optimized by using schematic representations of this information to direct learner's attention to concepts relevant for learning. A 2x2 between-groups design with the factors supportive information (before or during practice) and schematic representation (before or during practice) was used to investigate whether this balance between extraneous and germane load leads to more effective and efficient learning. It was found that the 'supportive during, schema before' format indeed yielded a higher learning efficiency than the 'supportive before, schema before' and the supportive during, schema during' format but no differences were found for learning effectiveness (i.e., test performance).

Just-in-time, Schematic Supportive Information Presentation during Cognitive Skill Acquisition

Practical educational approaches, such as project-based education, the case method, problem-based learning, and competency-based learning usually focus on (realistic and authentic) whole tasks or meaningful problems (Merrill, 2002; Reigeluth, 1999; van Merriënboer & Kirschner, 2001). Such tasks and problems help learners master the necessary cognitive skills and transfer what is learned in school to their daily life or work settings.

A potential pitfall of using such tasks is that, because of their complexity, they demand too much of the learner's cognitive system. Working memory capacity is limited and exceeding this capacity hampers learning (Baddeley, 1992; Miller, 1956; Sweller, 1988). The load imposed on working memory can be influenced in several ways. First, working on realistic tasks can be supported by sequencing tasks during practice from simple to complex (Reigeluth, 1983; 1999) or by using low-load task formats, such as worked-out examples (Paas & van Merriënboer, 1994; Van Gerven, Paas, van Merriënboer, & Schmidt, 2002) or completion tasks (van Merriënboer, Schuurman, de Croock, & Paas, 2002; van Merriënboer & de Croock, 1992). Second, information necessary to carry out the tasks can be presented just-in-time; in other words, precisely when the learner needs it for practice (van Merriënboer, Kirschner, & Kester, 2003; Kester, Kirschner, & van Merriënboer, in press). These instructional interventions aim at optimally utilize working memory capacity to facilitate cognitive skill acquisition. The latter approach is the focus of the current study.

<div align="center">Cognitive Load</div>

The cognitive load that a learner experiences during learning and practice is caused by a combination of the complexity of the problem to be solved and the design of the instructional materials. Solving the problem yields intrinsic cognitive load which is determined by the number of elements in the instructional material (e.g., practice problems, subject-relevant information) and the degree of element interactivity. Low element interactivity is associated with low intrinsic cognitive load because it allows the learner to serially process a small number of elements at a

time (Sweller, van Merriënboer, & Paas, 1998). For example, it is easy for a learner to serially process the meaning of the symbols used in a formula that is used to calculate the weighed input in a simple neural network. High element interactivity, on the other hand, is associated with a high intrinsic cognitive load because it requires the learner to process several elements and their relationships in working memory simultaneously in order to solve the problem. In this study, the learner has to simultaneously process several features of simple neural networks such as input nodes and output nodes to train such a system in categorizing input patterns.

The design of the instructional material yields both extraneous cognitive load and germane cognitive load. Extraneous cognitive load is caused by those processes a learner engages in that are not directly beneficial to learning, for instance, searching for relevant information sources, combining different information sources in order to understand the learning material, or weak-method problem solving. Extraneous cognitive load uses up cognitive resources at the cost of learning processes. Germane cognitive load includes all cognitive load associated with processes a learner engages in that are beneficial for learning, for instance, attending to important features of the problem (van Merriënboer, Schuurman, de Croock, & Paas, 2002), or abstracting from a variety of practice problems so as to construct more general cognitive schemata (Spiro, Coulson, Feltovich, & Anderson, 1988; Sweller et al., 1998). With a given intrinsic cognitive load, well-designed learning material minimizes extraneous cognitive load and optimizes germane cognitive load within the thresholds of totally available cognitive resources.

Within the broad range of different strategies available for managing the various forms of cognitive load (Chandler & Sweller, 1991; Sweller, 1988; Sweller et al., 1998), the present study focuses on the following strategies: (1) managing intrinsic cognitive load by sequentially presenting different training elements (e.g., information and practice tasks),
(2) minimizing extraneous cognitive load by just-in-time and just-in-place information presentation to avoid split attention, and (3) optimizing germane cognitive load by directing

learners' attention to concepts relevant for learning by providing schematic representations of information necessary to carry out the tasks.

Managing intrinsic cognitive load

In this study, a simulation is used to teach learners the functioning of a simple, two-layer neural network. Training such a simulated neural network to categorize input is the cognitive skill that learners have to master. Such cognitive skills are made up of different constituent skills (Fisk & Gallini, 1989; van Merriënboer, 1997; van Merriënboer et al., 2003). Non-recurrent constituent skills require interpretation of cognitive schemata, and the performance of such skills varies from problem situation to problem situation. In other words, it is the situation-specific use of the same, general knowledge. In the present study, an example of a non-recurrent skill is being able to reason about the working of a simple neural network in the case of an AND-classification-problem. Recurrent constituent skills require applying cognitive rules or automated schemata and the performance of these skills is the same in every problem situation. In other words, it entails the situation-independent use of the same, situation-specific knowledge. An example of a recurrent skill in this study is calculating the total input of a simple neural network by means of the appropriate formula. Non-recurrent and recurrent skills constitute the total cognitive skill that allows learners to solve problems in the domain.

In order to master the non-recurrent constituent skills, supportive information such as a conceptual model of how a learning domain is organized, is necessary. Learning supportive information is, in general, a task with high element interactivity because to-be-constructed mental models contain many interrelated elements. This has consequences for the proper management of cognitive load during practice (van Merriënboer et al., 2003). The foregoing discussion of cognitive load yields the advice to present supportive information before the practice problems since simultaneously solving the practice problems and processing the highly interactive supportive information is likely to increase the risk of overloading working memory. When supportive information is presented before practice, all cognitive capacity is available to

process this information, which allows learners to construct rich and adequate cognitive schemata. As a result, cognitive schemata that can easily be activated in working memory during the solution of practice problems will be available in long-term memory.

On the other hand, mastery of recurrent constituent skills requires the availability of procedural information; task-specific rules along with the facts, principles, or concepts needed to correctly apply these rules. Learning procedural information is, in general, a task with low element interactivity because task-specific rules always contain a limited number of related elements (i.e., some conditions and one action). Presenting procedural information during practice is thought to be warranted, because the risk of overloading working memory when simultaneously processing procedural information and solving practice problems is low due to the low element interactivity of the procedural information.

In sum, to properly manage intrinsic cognitive load it is important to sequentially present the supportive information and the practice tasks while simultaneously presenting the procedural information and the practice tasks.

Minimizing extraneous cognitive load

Extensive research has been carried out regarding the split attention effect (for an overview see Sweller et al., 1998; Kester, Kirschner, & van Merriënboer, in press). This research provides guidelines on when to (i.e., avoidance of temporal split attention), and how to (i.e., avoidance of spatial split attention) present the procedural and supportive information.

Mayer and his colleagues (Mayer & Anderson, 1991; Mayer & Moreno, 1998; Mayer & Sims, 1994) showed that temporal integration of narration and animation results in lower extraneous cognitive load because learners do not have to 'search-and-match' and mentally integrate the (narrated or animated) information that is presented. Translated to the present study, this could mean that although the supportive information and the practice task have to be sequentially presented in order to avoid cognitive overload, the time lapse between relevant supportive information and the practice task should be minimal to avoid temporal split-attention.

With regard to the procedural information, Sweller and his colleagues (Chandler & Sweller, 1991; Sweller, 1994) showed that extraneous cognitive load is reduced significantly by integrating explanatory text (i.e., here procedural information) in a diagram, instead of separating the diagram from the explanatory text spatially. Spatially integrating text and diagram frees learners from the need to mentally integrate the different information sources themselves which, in turn, lowers extraneous load. In this study the procedural information, that is the information necessary to carry out a specific task, accompanies graphical representations of neural networks. To avoid spatial split-attention and, thus, to minimize the extraneous cognitive load in the instructional material, it appears to be advisable to present procedural information during practice, integrated in the graphical representations of the simple neural networks (Kester, Kirschner, van Merriënboer, & Bäumer, 2001; van Merriënboer et al., 2003). Moreover, the simultaneous, integrated presentation of procedural information and practice tasks helps learners to automate schemata that they apply to familiar aspects of novel problem situations, because the necessary procedural information is made directly available in working memory during practice and so facilitates a process known as knowledge compilation (Anderson, 1996), by which the relevant knowledge is compiled in automated schemata.

Thus, in order to minimize extraneous cognitive load by avoiding temporal and spatial split-attention, it is important to make the time lapse between the supportive information and the practice task not too long and present the procedural information in an integrated format during the practice tasks (i.e., just-in-time and just-in-place).

Optimizing germane cognitive load

It is not, however enough to only minimize extraneous cognitive load. Instructional materials must also optimize germane cognitive load. To support learners in constructing rich and adequate schemata during task practice, schematic representations of the supportive information can be used. The idea that schematic representations direct a learner's attention to concepts relevant for learning and thus, facilitate learning originates from Ausubel's

Subsumption Theory (1963). According to Ausubel, meaningful learning only occurs when new material can be appropriately subsumed under a relevant, existing concept. In order to facilitate this process, advance organizers (e.g., in the form of schematic representations of learning material) can be applied in instructional design. These organizers provide advance ideational scaffolding: when the learner is confronted with unfamiliar material, a cognitive structure is available for incorporating the new concepts (Ausubel, 1963; Jonassen, 1982). In a meta-analysis of 132 studies, advance organizers appeared to facilitate learning and retrieval (Jonassen, 1982; Mayer, 1979). However, they have no effect when relevant cognitive structures are already available and are appropriately activated by either the learner or the material presented. Nowadays, the advance organizers are considered to be a type of graphic organizer (i.e., figural organizations of text information, Robinson & Kiewra, 1995; Robinson, 1998) used to promote meaningful learning. Graphic organizers can take many forms (e.g., matrices or tree diagrams) but an important aspect they all have in common is that they provide hierarchical and coordinate relations (Robinson & Kiewra, 1995; Robinson, Katayama, Dubois, & Devaney, 1998).

Thus, to optimize germane cognitive load, schematic representations (or graphic organizers) of the supportive information can be used to direct learners' attention to concepts relevant for schema construction, provided that the supportive information is subsequently available in the instructional material.

The focus of this study is on the presentation of supportive information with aim of shedding some light on minimizing extraneous cognitive load by avoiding temporal split-attention with regard to the supportive information. Moreover, a closer look is taken at the role of using schematic representations of the supportive information to optimize germane cognitive load. To investigate these two aspects, the timing of the presentation of the supportive information (before or during practice) and the schematic representation (before or during practice) is varied. The procedural information is presented during practice in all conditions. Practice consists of the presentation and solution of a series of practice tasks dealing with neural

networks. Presentation before practice means that the supportive information and/or the
schematic representation are available only before the practice task. Presentation during practice
means that the supportive information is available during each practice task and/or the schematic
representation is available during each practice task.

With respect to learning effectiveness, it is hypothesized that making supportive
information available during practice and the schematic representation before practice will yield
higher test performance than the other conditions, since minimizing extraneous cognitive load
and optimizing germane cognitive load during learning facilitates the acquisition of rich,
adequate schemata. Furthermore, learning will be more efficient when supportive information is
presented during practice and the schematic representation is given in advance. In other words, a
relatively high test performance will be obtained with relatively low cognitive load during
practice. Optimally balancing extraneous and germane cognitive load enables learners to use a
maximum of their cognitive capacity for processes relevant for learning because little cognitive
load is caused by activities irrelevant to learning.

Method

Participants

Forty-eight second- and third-year students (26 male and 22 female; mean age = 23.5
years, *SD* = 3.1) of the Faculty of Psychology at the Maastricht University voluntarily
participated in this study. During a practice session they had to learn to train a simple neural
network, using a computer-based simulation. All participants followed a comparable practical at
least one year earlier. Before the experiment the participants received a 12-item, multiple-choice,
factual knowledge test to assess their prior-knowledge and they had to fill in the grade they
received the first time they carried out the practical. An ANOVA showed no significant
differences between the experimental groups on the pre-experimental factual knowledge test,
$F(3, 48) = .31$; $p = .82$. Moreover, the participants achieved low scores on this test ($M = 2.62$; *SD*

= 1.86; max. = 12) which indicated that they al had low prior knowledge about the problem domain. A Kruskal-Wallis test on the ordinal grade data showed no significant differences between the experimental groups, $\chi^2(3, 48) = 1,48; p = .69$.

Materials

    Adaline®. A computer-based simulation program Adaline® was used to teach students the functioning of a simple, two-layer neural network. Adaline® stands for 'ADAptive LINear Element'. The network contains *nodes*, which are simplified representations of neurons, and *weights*, which are simplified representations of synaptic connections in the brain. The function of this simulated neural network is to categorize a set of nodes that form a specific input pattern. This input pattern is represented by an input vector in combination with a weight vector (i.e., the strength of synaptic connections) in a coordinate system. Adaline® has to be 'trained to recognize' input vectors as belonging to one of two categories, A or B. This is done by adjusting the weights of the network through vector processing. After proper training, the network is able to solve AND/OR classification problems. Adaline® is part of the Psychology curriculum at Maastricht University and is used as a practical for second-year students in an introductory course on cognitive science. Here, Adaline® is adapted to the experimental purposes and it starts with an introduction in which vector processing is explained. The introduction is followed by 11 practice tasks.

    Information presentation. The practice tasks, consisting of simulated simple neural networks, were accompanied by supportive information, a schematic representation of this information, and procedural information. The procedural information was presented stepwise during practice for all experimental groups. The supportive information and the schematic representation were made available either before practice, during practice or before *and* during practice. Four information-presentation formats were distinguished in a 2 x 2 factorial design with the factors and timing of supportive information, before or during practice, and timing of schematic representation presentation presentation, before or during practice. In the 'supportive

before, schema before' format *both* the supportive information and the schematic representation were presented *before* practice. The participants assigned to the 'supportive during, schema during' format received *both* the supportive information and the schematic representation *during* practice. In the 'supportive before, schema during' format the supportive information was presented *before* practice while the schematic representation was presented *during* practice. Finally, in the 'supportive during, schema before' format the supportive information was presented *during* practice while the schematic representation was presented *before* practice. For an example of a practice problem see Figure 1.

*** Insert Figure 1 about here***

In the first two practice tasks in every format, the participants had to read through the supportive information and study the schematic representation before they could fill in the answers to the practice tasks. In the remaining nine practice tasks the supportive information and the schematic representation were available on demand by clicking a 'more information' button or a 'schema' button.

Practice tasks. Adaline® contains 11 practice tasks that the participants have to carry out in a fixed order. In these practice tasks, participants have to perform a series of calculations (i.e., vector processing) to categorize the input pattern. First, they have to calculate the weighed input which results in a weighed input vector. Second, they have to define the output. That is, if the weighed input > 0, then the output is 1 (i.e., category A); if the weighed input is ≤ 1, then the output is −1 (i.e., category B). Third, they have to calculate the error, that is, they have to check whether the expected output (i.e., this is given) is the same as the real output defined in step 2. If this is not the case, then they have to perform a fourth step in which they have to adjust the weighed vector by applying the delta rule (i.e., a mathematical procedure that calculates the minimal error) until it fits the input as much as possible. Eight tasks only required the first three steps and three tasks also required the fourth step.

Factual knowledge test. Before and after the experimental session in Adaline® (i.e., the 11 practice tasks) the same multiple-choice test was administered to the participants. Before the experimental session this test was used to measure participants' prior factual knowledge while after the experiment it was used to measure factual learning or retention. This test contained 12 items and the participants could choose their answer from four alternatives.

Comprehension test. After the experimental session and the factual knowledge test a comprehension test was administered to the participants. This test contained four open questions and was meant to test whether the participants had gained insight in the working of simple, two-layer neural networks. They could receive one point for each correctly answered open question, so the maximum score on the comprehension test was four. To determine the inter-rater reliability, the comprehension test scores of ten participants were determined by two raters. The inter-rater reliability for the comprehension test scores of the two raters was .93 (Intraclass Correlation Coefficient).

Workload measurement. The NASA-TLX tool was used to measure workload (1) after the practice tasks, and (2) after the factual knowledge test (post experiment) and the comprehension test. The NASA-TLX is a subjective workload assessment tool that involves a multi-dimensional rating procedure, from which an overall workload score based on a weighted average of ratings on six subscales (i.e., mental demand, physical demand, temporal demand, performance, effort and frustration) is derived. Except for the performance scale, which ranged from 'good' to 'poor', all scales ranged from 'low' to 'high' and were divided in twenty sections (see Figure 2a). The participants had to click on one of the twenty sections that best represented their experience for each aspect. To establish the weighing, 15 possible pair-wise comparisons of the six scales were presented to the participants (see Figure 2b). The participants had to select the member of each pair that contributed more to the workload of that task than the other member (e.g., temporal demand vs. frustration or performance vs. mental effort). Based on the scale-scores and the

weighing, the total workload was calculated (minimum workload = 5; maximum workload = 100).

Insert Figure 2a, b about here

Learning and performance efficiency. The procedure of Paas and van Merriënboer (1993; see also Paas et al., 2003) was used to calculate efficiency. First, the performance measures (i.e., the factual knowledge test and the comprehension test) and the workload measures for each participant are transformed into z-scores, using the grand mean across conditions. Then, the mean z-scores for every condition are represented in a Cartesian coordinate system with Workload (W) on the horizontal axis and Performance (P) on the vertical axis. The line P = W through the origin of the system is assumed to indicate a neutral efficiency (slope = 45°). The efficiency, E, is calculated as the perpendicular distance from a data point in the coordinate system to the line $P = W$ (Paas & van Merriënboer, 1993). The formula for calculating this distance is:

$$E = \frac{\text{Performance - Workload}}{\sqrt{2}}$$

Equal performance (P) and workload (W) z-scores yield an instructional efficiency of zero, a neutral score. If $P > W$, the instructional material is more efficient, indicated by a positive value because the performance is higher than might be expected on the basis of workload. If $P < W$, the material is less efficient, indicated by a negative value because the performance is lower than might be expected on the basis of workload. *Learning* efficiency is calculated based on workload during practice and test performance on both tests separately while *performance* efficiency is calculated based on the workload during the tests and test performance.

Logging. Adaline® makes use of a built-in log tool which registers the number of times participants click the 'more information' button and the number of times they click the 'schema'

button (i.e., repeats). Moreover, it registers the time participants spend on the practice tasks and the test tasks (i.e., time on task).

Procedure

Before starting the experiment, the participants individually received oral instructions on how to behave during the experiment. They were instructed to work in a concentrated manner and at a steady pace. No time limit was imposed. After the oral instruction, they read the paper-based instruction in which the sequence of the experiment was outlined (i.e., factual knowledge test, Adaline®, factual knowledge test, comprehension test), the workload measure was introduced, and their practical grade was asked. During this instruction, the participants had the opportunity to ask questions about the experimental procedure. After filling in their answers on the 12 item factual knowledge test, the participants started with Adaline®. First, they had to read an introduction to simulating simple neural networks. Then the practice tasks started. During the first practice task, the experimenter explained how to fill in the solutions to the practice problems (i.e., always fill in two digits after the decimal point in your answer). No instructions about content were given. After the 11 practice tasks, the participants were required to read an instruction about how to fill in the NASA-TLX and then they filled in the scale for the practice tasks. Then the same factual knowledge test was administered, followed by the comprehension test and the NASA-TLX for the factual knowledge test and the comprehension test.

Results

A 2 x 2 between-groups analysis of variance (ANOVA) was performed on all dependent measures. A repeated-measures ANOVA was performed on pre- and post-test performance. In all analyses, a significance level of .05 is maintained for all statistical analyses. For an overview of the results see Table 1.

Performance

A repeated-measures ANOVA showed a main effect for factual knowledge test moment (i.e., before and after the experiment), $F(1, 48) = 442.54$; $MSE = 988.17$; $p < .001$; $\eta^2 = .91$, the

participants all performed significantly better on the factual knowledge test after the experiment ($M = 9.04$; $SD = 1.86$) than before ($M = 2.62$; $SD = 1.86$). No interaction effect between factual knowledge, test moment, and condition were found, $F(3, 48) = .39$; $MSE = .86$; $p = .76$. So, participants all improved equally well.

The between-groups ANOVA showed no significant differences between the experimental groups for the scores on the factual knowledge test after the experiment, $F(1, 48) = .37$; $MSE = 1.33$; $p = .55$. Moreover, the ANOVA showed no differences between the experimental groups for the scores on the comprehension test, $F(1, 48) = .14$; $MSE = .09$; $p = .71$.

Workload

For total practice workload, the ANOVA revealed an interaction between the timing of the schematic representation and the timing of supportive information, $F(1,48) = 6.23$; $MSE = 271.55$; $p < .05$; $\eta^2 = 1.24$ (see Figure 3). If the schematic representation was given before practice, the presentation of supportive information during practice led to a lower workload ($M = 52.67$; $SD = 19.43$) than when supportive information was also presented before practice ($M = 62.19$; $SD = 10.95$). Reversely, when the schematic representation was presented during practice, the presentation of supportive information before practice led to a lower workload ($M = 54.78$; $SD = 16.16$) than when supportive information was also presented during practice ($M = 69.00$; $SD = 18.10$). Tukey's HSD showed no significant differences between the experimental conditions.

No significant differences were found for the total test workload, $F(1,48) = 3.23$; $MSE = 939.28$; $p = .08$.

Insert Figure 3 about here

Efficiency

For the factual knowledge test after the experiment, the ANOVA revealed an interaction for *learning efficiency* between the timing of the supportive information and the timing of the schematic representation, $F(1,48) = 7.92$; $MSE = 4.16$; $p = .047$; $\eta^2 = 0.09$. If the supportive

information was presented before practice, it was more efficient to present the schematic

representation during practice (M = .25; SD = 1.08) than to present it before practice (M = - .23;

SD = .90). If the supportive information was presented during practice, it was more efficient to

present the schematic representation before practice (M = .37; SD = 1.9) than to present it during

practice (M = - .40; SD = .97) (see Figure 4a). Tukey's HSD showed no significant differences

between the experimental conditions.

A same significant interaction effect for *learning efficiency* was found for the

comprehension test, $F(1, 48) = 4.26$; $MSE = 3.85$; $p = .045$; $\eta^2 = 0.09$. If the supportive

information was presented before practice, it was more efficient to present the schematic

representation during practice (M = .22; SD = 1.16) than to present it before practice (M = - .17;

SD = .93). If the supportive information was presented during practice, it was more efficient to

present the schematic representation before practice (M = .34; SD = .83) than to present it during

practice (M = - .39; SD = .85) (see Figure 4b). Tukey's HSD showed no significant differences

between the experimental conditions.

No significant results were found for *performance efficiency* on both the factual

knowledge test after the experiment, $F(1,48) = 2.45$; $MSE = 2.87$; $p = .13$, and the comprehension

test, $F(1,48) = 2.54$; $MSE = 2.34$; $p = .12$.

Insert Figure 4a,b about here

Logging

Repeats. No significant results were found for the number of times the participants either

clicked the schematic representation $F(1, 48) = .18$; $MSE = .75$; $p = .67$ or the supportive

information, $F(1, 48) = .84$; $MSE = 1.69$; $p = .36$.

Time on task. No significant results were found for introduction time on task, $F(1,$

$48) = 1.95$; $MSE = 33.33$; $p = .17$, and for practice time on task, $F(1, 48) = 3.15$; $MSE = 609.19$;

$p = .08$.

Discussion

The question this study attempted to answer was: How can intrinsic, extraneous and germane cognitive load be optimally balanced during complex cognitive skill acquisition? It was investigated whether (1) managing intrinsic cognitive load by sequentially presenting the supportive information and the practice tasks, (2) minimizing extraneous cognitive load by avoiding *temporal* split attention through presenting supportive information during practice and (3) optimizing germane cognitive load by directing attention to concepts relevant for learning using a schematic representation of the supportive information before practice led to more effective and efficient learning. It was found that the 'supportive during, schema before' format indeed yielded a higher *learning efficiency* than the 'supportive before, schema before' format and the 'supportive during, schema during' format. Participants in the 'supportive during, schema before' format performed better while experiencing a lower workload. However, the same was true for the 'supportive before, schema during' format. Furthermore, every participant performed much better on the post-experimental factual knowledge test than on the pre-experimental factual knowledge test but no differences between formats were found for learning *effectiveness* (i.e., post-experimental factual knowledge test and the comprehension test).

The higher learning efficiency of the formats that presented the supportive information and the schematic representation separately is due to a difference between *workload* during practice between these two formats and the formats that presented this information and the schematic representation simultaneously. The workload during practice in the 'supportive before, schema during' format and the 'supportive during, schema before' format was lower than the workload during practice in the other formats. So, participants in the former formats reached the *same* performance while perceiving a *lower* workload.

From these results could be concluded that (1) managing intrinsic cognitive load by sequentially presenting the supportive information and the practice tasks indeed reduced intrinsic cognitive load and led to more efficient learning. The workload was lower and the learning efficiency was higher for participants who received the supportive information *or* the schematic

representation of this information separately before or during practice. These results are comparable to findings from Kester, Kirschner and Van Merriënboer (2004b) and Kester, Kirschner, and Van Merriënboer (submitted) who varied the timing of supportive information (i.e., before or during practice) and procedural information (i.e., before or during practice) and found that a 'supportive before, procedural during' information presentation format and a 'supportive during, procedural before format' led to more efficient learning than a 'both before' format and a 'both during' format. In these studies the simultaneous presentation of supportive information and procedural information interferes with learning just as the simultaneous presentation of the supportive information and the schematic representation. It has to be noted that in the present study the simultaneous presentation of supportive information and procedural information does not hamper learning because the procedural information used here is much more concise than the procedural information used in Kester et al. (2004b) and Kester et al. (submitted).

Furthermore, the results show that (2) minimizing extraneous cognitive load by avoiding *temporal* split attention through presenting supportive information during practice did not lead to a reduction in extraneous cognitive load. No differences between formats for learning *effectiveness* were found. The participants in the different experimental groups all reached the same performance on the post-experimental factual knowledge test and the comprehension test. Moreover, they all benefited substantially from practice shown by the difference in performance on the pre-experimental factual knowledge test (i.e., low) and the post-experimental factual knowledge test (i.e., high). This means that the differences in workload between the formats that presented the supportive information and the schematic representation separately and the formats that presented this information and the schematic representation simultaneously are not related to extraneous cognitive load. Although the instructional material in the 'supportive before, schema before' and the 'supportive during, schema during' formats yielded a higher workload this did not

hamper learning as would have been the case if learner had to allocate cognitive capacity to processes irrelevant for learning.

These results could be explained by looking more close at the split attention research that did show the deteriorating effects on learning of learning material that causes split attention. Research that investigated the *spatial* split attention effect showed that integrated learning material was more effective than conventional learning material but only when the mutually referring information sources were *unintelligible* without mental integration. (Chandler & Sweller, 1991; Cerpa, Chandler & Sweller, 1996; Chandler & Sweller, 1992; Chandler & Sweller, 1996; Sweller & Chandler, 1994). Research pertaining to investigating *temporal* split attention showed that *mutually referring* auditory information and visual information resulted in enhanced test performance (Kalyuga, Chandler & Sweller, 1999; Mayer & Anderson, 1991,1992; Mayer & Moreno, 1998; Mayer & Sims, 1994). In the present study, the supportive information and the practice tasks are intelligible by themselves. Although the supportive information is needed to understand the practice tasks, the calculations could be carried out based on the procedural information that was integrated into the tasks. This makes it less relevant to avoid temporal split attention because the necessity to mentally integrate the supportive information and the practice tasks is low. Therefore, it might have been not suitable to apply the split attention guidelines to the instructional material that was used in this study.

From the results can also be concluded that (3) optimizing germane cognitive load by directing attention to concepts relevant for learning using a schematic representation of the supportive information before practice did not lead to an increase in germane cognitive load. For all conditions the mean workload varied between 52.67 and 69.00 which means that none of the participants experienced a workload that came close to or exceeded their cognitive capacity. This means that although participants had cognitive capacity left to allocate to processes relevant for learning, the 'supportive during, schema before' format, that aims at redirecting learners' attention to concepts relevant for schema construction, did not elicited them to do so. Since the

participants in this study all followed a similar course in Adaline® at least one year earlier it is possible that they already had relevant cognitive structures available and as Jonassen (1982) and Mayer (1979) showed advance organizers do not have a surplus value when relevant cognitive structures already exist.

Moreover, meta-analyses of Moore and Readance (1980, 1984; as cited in Griffin and Tulbert, 1995) revealed that post-reading learner-constructed graphic organizers produced more learning than pre-reading teacher-constructed graphic organizers. In the present study a teacher-constructed schematic representation was used that was presented both pre-reading and post-reading. It is possible that this teacher-constructed characteristic of the schematic representation accounts for the lack of beneficial effects on learning of this representation. Learner-constructed graphic organizers urge learners to actively participate in the learning process while teacher-constructed graphic organizers probably do not.

To conclude, the present study succeeded in properly managing intrinsic cognitive load during complex cognitive skill acquisition by presenting relevant instructional elements (i.e., the supportive information, the schematic representation and the practice task with the integrated procedural information) separately instead of simultaneously. This led to a higher learning efficiency and less workload. Measures to enhance the learning material by minimizing extraneous cognitive load (i.e., avoiding temporal split attention) and optimizing germane cognitive load (i.e., directing attention to relevant concepts) did not result in more effective or efficient learning. This was probably due to (1) the intelligibility of the supportive information and the practice tasks in isolation which made avoiding split attention less important, (2) the level of prior-knowledge of the participants which was probably to high to experience beneficial effects on learning of the schematic representation and (3) the use of a teacher-constructed schematic representation that probably prevented the learners to become actively involved with the learning material.

In a setting where learning tasks involve supportive information, procedural information and realistic practice tasks, future research should aim for optimizing germane cognitive load. This could be done by giving learners more control over their own learning, for example by urging them to create their own schematic representation of the supportive information after reading it, by giving them control over task selection which seems promising (Camp, Paas, Rikers, Van Merriënboer, 2001) or by giving them control over information selection (Kester, Kirschner, Corbalán Pérez, submitted). Although, stimulating learners to become actively involved with the learning material probably increases cognitive load this also has beneficial effects on motivation (Bandura, 1997) and learning outcomes (Gray, 1987).

<div align="center">References</div>

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

Ausubel, D. P. (1963). *The psychology of meaningful verbal learning.* New York: Grune & Stratton.

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

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: W. H. Freeman.

Camp, G., Paas, F., Rikers, R., & Van Merriënboer, J. J. G. (2001). Dynamic problem selection in air traffic control training: A comparison between performance, mental effort and mental efficiency. *Computers in Human Behavior, 17*, 575-595.

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*, 345-367.

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

Gray, S. J. (1987). The effects of sequence control on computer learning. *Journal of Computer-based Instruction. 14*(2), 54-56.

Fisk, A. D., & Gallini, J. K. (1989). Training consistent components of tasks: Developing an instructional system based on automatic-controlled processing principles. *Human Factors, 31*, 453-463.

Jonassen, D. H. (1982). *The technology of text. Principles for structuring, designing, and displaying text.* New Jersey: Englewood Cliffs.

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., Corbalán Pérez, G. (submitted). Learner control over information presentation in powerful electronic learning environments. Leuven University Press.

Kester, L., Kirschner, P. A., & van Merriënboer, J. J. G. (submitted). Just-in-time information presentation: Improving transfer performance and learning efficiency of a complex troubleshooting skill.

Kester, L., Kirschner, P. A., & van Merriënboer, J. J. G. (in press). The split attention effect in computer simulated troubleshooting of electrical circuits. *British Journal of Educational Psychology.*

Kester, L., Kirschner, P. A. & Van Merriënboer, J. J. G. (2004a). The optimal timing of information presentation during mastering a complex skill in science. *International Journal of Science Education, 26*, 239-256.

Kester, L., Kirschner, P. A., van Merriënboer, J. J. G. (2004b). Just in time presentation of different types of information in learning statistics. *Instructional Science, 32, 233-252.*

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

Mayer, R. E. (1979). Can advance organizers influence meaningful learning? *Review of educational research, 49*, 371-383.

Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: an experimental test of a dual-coding hypothesis. *Journal of Educational Psychology, 83*, 484-490.

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*, 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*, 389-401.

Merrill, M. D. (2002). First principles of instruction. *Educational Technology, Research and Development, 50*(3), 43-59.

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*, 429-434.

Paas, G. W. C., & van Merriënboer, J. J. G. (1993). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human Factors, 35*, 737-743.

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., & Adam, J. J. (1994). Measurement of cognitive load in instructional research. *Perceptual and Motor Skills, 79*, 419-430.

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

Reigeluth, C. M. (1983). *Instructional design theories and models: An overview of their current status.* Hillsdale, NJ:: Lawrence Erlbaum..

Spiro, R. J., Coulson, R. L., Feltovich, P. J., & Anderson, D. K. (1988). Cognitive flexibility theory: Advanced knowledge acquisition in ill-structured domains. In V. Patel (ed.), *Proceedings of the 10th Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

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

Sweller, J. (1994). Cognitive load theory, learning difficulty and instructional design. *Learning and Instruction, 4*, 295-312.

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

Sweller, J., van Merriënboer, J. J. G., & Paas, G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251-296.

Van Gerven, P. W. M., Paas, G. W. C., van Merriënboer, J. J. G, & Schmidt, H. G. (2002). Cognitive load theory and aging: Effects of worked examples on training efficiency. *Learning and Instruction, 38*, 87-107.

Van Merriënboer, J. J. G. (1997). *Training complex cognitive skills: A four-component instructional design model for technical training.* Englewood Cliffs, NJ: Educational Technology Publications.

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. (2001). Three worlds of instructional design: State of the art and future directions. *Instructional Science, 29*, 429-441.

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.

Van Merriënboer, J. J. G. , Schuurman, J. G., de Croock, M. B. M., & Paas, G. W. C. (2002). Directing learners' attention during training: Effects on cognitive load, transfer test performance and training efficiency. *Learning and Instruction, 12*, 11-37.

Table 1

Overview of the results

	SupB-SchemaD[a]		SupB-SchemaB		SupD-SchemaD		SupD-SchemaB	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Performance								
Pre-factual knowledge test (max. 12)	2.75	1.49	2.83	1.64	2.17	2.17	2.75	2.22
Post-factual knowledge test (max. 12)	9.17	2.04	8.75	1.66	9.00	1.60	9.25	2.26
Comprehension test (max. 4)	2.04	1.01	1.94	.82	2.00	.60	1.94	.70
Workload								
Practice workload *	54.78	16.16	62.19	10.95	69.00	18.10	52.67	19.43
Test workload	58.41	19.32	68.64	16.25	71.72	16.13	64.25	16.34
Efficiency								
Learning efficiency post factual knowlegde test *	.25	1.08	-.22	.90	-.40	.97	.37	1.20
Learning efficiency comprehension test *	.22	1.16	-.17	.93	-.39	.84	.34	.83
Performance efficiency post factual knowledge test	.35	1.31	-.23	1.05	-.26	.87	.14	1.06
Performance efficiency comprehension test	.32	1.41	-.19	.83	-.26	.55	.11	.84
Logging								
# Repeats	1.75	2.34	1.75	1.87	.83	1.75	1.33	2.15
Introduction TOT[b] (mins.)	9.17	3.10	12.08	6.32	10.17	2.29	9.75	3.67
Practice TOT (mins.)	54.58	15.92	55.92	16.73	67.33	7.4	54.42	13.65

*p < .05

[a]Sup = supportive information; Schema = schematic representation; B = before, and D = during.

[b]TOT = Time on Task

Figure captions

Figure 1. Screendump of a practice problem in the 'schema during, supportive during' format

Figure 2a, b. Examples of the NASA-TLX tool.

Figure 3. Interaction effect of practice workload.

Figure 4a, b. The mean learning efficiency scores of the pre-experimental factual knowledge

test (a) and the comprehension test (b) drawn in a Cartesian coordinate system.
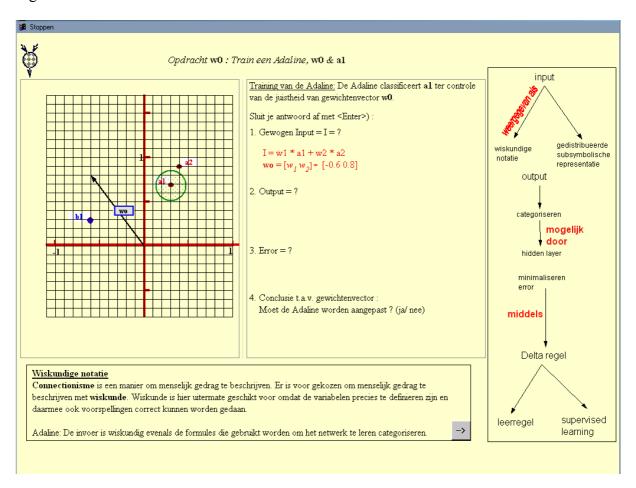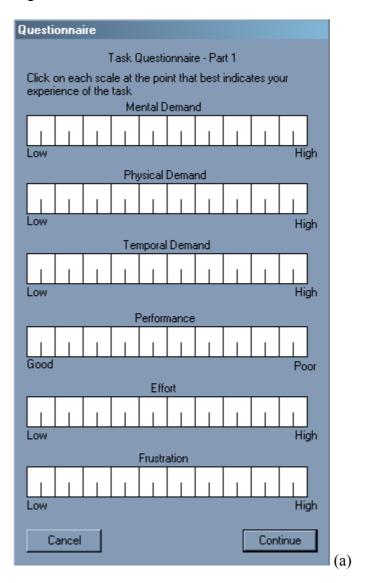
.

Figure 1

Figure 2

(a)

(b)

Figure 3

Figure 4

Learning efficiency

(a)

Learning efficiency

(b)