Pergamon

Learning and Instruction

Learning and Instruction 12 (2002) 11–37

www.elsevier.com/locate/learninstruc

Redirecting learners' attention during training: effects on cognitive load, transfer test performance and training efficiency

J.J.G. van Merriënboer [a,*], J.G. Schuurman [b], M.B.M. de Croock [1,c], F.G.W.C. Paas [a]

[a] *Open University of the Netherlands, OTEC, P.O. Box 2960, 6401 DL Heerlen, The Netherlands*
[b] *Telematics Institute, Department of Educational Science and Technology, 7500 AE Enschede, The Netherlands*
[c] *University of Twente, P.O. Box 589, 7500 AE Enschede, The Netherlands*

Abstract

Cognitive load theory provides guidelines for improving the training of complex cognitive skills and their transfer to new situations. One guideline states that *extraneous* cognitive load that is irrelevant to the construction of cognitive schemata should be minimised. Experiment 1 (N=26) compares completion problems, conventional problems, and a learner-controlled condition in which learners may choose between problem formats. Completion problems decrease cognitive load during training and have a zero or positive effect on transfer performance. A second guideline states that *germane* cognitive load that is directly relevant to schema construction should be optimised. In Experiment 2 (N=69) practice schedules of either high or low contextual interference are compared (HCI and LCI). HCI increases cognitive load during training and shows a trend towards higher transfer performance. Experiment 3 (N=87) combines both guidelines in a factorial experiment with the factors problem format (completion vs. conventional) and contextual interference (HCI vs. LCI). It is hypothesised that *redirecting attention* from extraneous to germane processes will improve training efficiency, i.e. positively affect the balance between cognitive load during training and transfer test performance. In support of this hypothesis, it is found that the completion-HCI group shows highest training efficiency. But transfer test performance for this group is disappointing. The results are discussed in relation to the operationalisation of HCI in combination with completion problems.

* Corresponding author.
E-mail address: jeroen.vanmerrienboer@ou.nl (J.J.G. van Merriënboer).
[1] Marcel de Croock is now at the Open University of the Netherlands.

0959-4752/01/$ - see front matter
PII: S0959-4752(01)00020-2

Keywords: Cognitive load theory; Completion problems; Contextual interference; Transfer; Training efficiency

1. Introduction

The acquisition of complex cognitive skills is heavily constrained by the limited processing capacity of the human mind. Cognitive load theory (CLT; Sweller, 1988; Sweller, van Merriënboer and Paas, 1998) provides guidelines to circumvent those limitations in training situations. It provides guidelines to: (1) prevent cognitive overload (see van Merriënboer, 1997); (2) decrease extraneous cognitive load which is not relevant to learning; and (3) increase, within the limits of total available cognitive capacity, germane cognitive load which is directly relevant to learning. This article focuses on the decrease of extraneous cognitive load (Experiment 1), the increase of germane cognitive load (Experiment 2), and the combination of both methods called "redirecting attention" (Experiment 3).

Originally, CLT focused on instructional techniques for decreasing *extraneous* cognitive load, that is, load imposed by cognitive processes not directly relevant for learning. It identified several processes yielding a high extraneous load, such as applying means–ends-analysis in problem solving; mentally integrating physically separated sources of information, and dealing with redundant information. Research on alternative problem formats, such as goal-free problems and worked examples showed that such formats yielded less extraneous cognitive load and improved learning. In addition, the design of these problems proved to be critical. For instance, they should be designed in such a way that learners do not have to split their attention between different sources of information or deal with redundant information (see Sweller, van Merriënboer, & Paas, 1998 for an overview of instructional techniques). In the current article, the focus is on the use of *completion problems* as an instructional method to decrease extraneous cognitive load. Completion problems provide both a problem statement and a partial solution to the problem, which must be completed by the learners.

More recently, CLT has also been used to explain the effects of instructional techniques that increase *germane* cognitive load, that is, load imposed by cognitive processes directly relevant for learning. According to CLT, these are mainly processes that have to do with the construction of cognitive schemata. They are effort demanding, inductive processes that typically require "mindful abstraction" from the learners. There are indications that variability over problem situations or other task dimensions, such as the manner in which the task is presented, the saliency of defining characteristics, or the context in which the task is performed, may increase germane cognitive load and improve schema construction (e.g. Paas & van Merriënboer, 1994; Quilici & Mayer, 1996). The availability of generalised, more abstract cognitive schemata is then revealed by a higher performance on transfer tasks, which require the performance of the complex cognitive skill in new situations. In the current article, the focus is on the use of high *contextual interference* as an instructional method to increase germane cognitive load. High contextual interference may be seen as a particular type of variability that results from intertask interference.

The structure of this section is as follows. First, research on completion problems and their effects on transfer test performance and extraneous cognitive load are reviewed. Second, research on contextual interference and its effects on transfer test performance and germane cognitive load are reviewed. Third, the simultaneous use of completion problems and high contextual interference is discussed as a promising instructional method to redirect the learner's attention during training. In addition, training efficiency is introduced as a necessary measure to be able to compare instructional methods on their combined effects on cognitive load during training and resulting transfer test performance. Fourth, the main hypotheses are presented for the three experiments reported in this article.

1.1. The completion effect and extraneous cognitive load

van Merriënboer and Krammer (1987, 1990) first suggested the use of completion problems to increase the transfer of computer programming skills. Completion problems are problems for which a given state, a goal state, and a partial solution (e.g. an incomplete computer program) are provided to learners who must complete the partial solution. They provide a bridge between worked examples and conventional problems. Worked examples can be seen as completion problems with a full solution, and conventional problems can be seen as completion problems with no solution. Particular training strategies may start with completion problems that provide almost complete solutions, and gradually work to completion problems for which all or most of the solution must be generated by the learners. Such a strategy is known as the "completion strategy".

The completion effect indicates that solving completion problems yields higher transfer of acquired skills than conventional problem solving. An explanation for this effect is that learners who work on conventional problems apply weak-method problem solving methods, such as means–ends analysis. This process is expensive in terms of — extraneous — cognitive load (Sweller, 1988) and bears little relation to schema construction processes that are concerned with learning to recognise problem states and associated solution steps. In contrast, learners who work on completion problems can focus their attention on problem states and associated solution steps, enabling them to induce generalised solutions or cognitive schemata. Precisely those cognitive schemata allow for transfer of acquired skills.

The studies by van Merriënboer (1990) and van Merriënboer and de Croock (1992) on computer programming gave strong support to the completion effect, but unfortunately they did not gather data on cognitive load. In the domain of statistical problem solving, Paas (1992) first compared the effects of completion problems, worked examples, and conventional problems on cognitive load during transfer test performance and training performance. Learners rated their perceived mental effort for each problem or example on a 9-point symmetrical rating scale. Perceived mental effort is a valid measure of cognitive load (Paas, van Merriënboer, & Adam, 1994). As predicted on the basis of CLT, it was found that training with completion problems or worked examples required the same amount of mental effort and led to higher transfer test performance, combined with lower cognitive load during the test than training with conventional problems. This indicates that superior cognitive schemata were constructed in the completion and worked example conditions. The absence of significant differences in experienced cognitive load during training between the three conditions may — partly — be a bottom effect. In the present article, Experiment 1 reports a new study in the computer programming domain in which cognitive load during training and transfer test performance are compared for completion problems, conventional problems, and a learner-controlled condition in which the learners may choose between completion and conventional problems.

1.2. The contextual interference effect and germane cognitive load

Contextual interference refers to training conditions in which certain contextual factors prohibit a quick and smooth mastery of skills being trained. There are different types of sources for contextual interference (for an overview, see Magill & Hall, 1991). The type studied in this article can best be characterised as intertask interference, in which the amount of contextual interference is varied by manipulating the schedules by which the skills, necessary for solving a particular type of problem, are practised. Low contextual interference may be produced by a blocked practice schedule, in which the skills necessary for solving one type of problem are practised before continuing to a next type of problem (e.g. B-B-B, A-A-A, C-C-C). High contextual interference may be produced by a random practice schedule, in which different problems are sequenced in a random order (e.g. C-A-B, B-C-A, B-A-C). For instance, suppose that a troubleshooting task is practised where three types of malfunctions can each occur in four different components of a system. Then, a low contextual interference condition would practice troubleshooting *one* type of malfunction, occurring in each of the four components, before the other two types of malfunctions. In contrast, a high contextual interference condition would sequence malfunction-component pairs in a random order.

The contextual interference effect indicates that high contextual interference yields better retention and higher transfer of acquired skills than low contextual interference, but that learners who practice under interfering conditions will typically need more time and invest more mental effort to master the skills. This effect is believed to be the result of learning processes in which available information is more deeply elaborated and in which general applicable knowledge is induced from the solutions to the different problems that were presented. When practising under high contextual interference, each successive problem requires learners to apply different knowledge and skills in order to be able to understand and solve that particular problem. Therefore, they are stimulated to compare the solutions to each problem and to mindfully abstract differences and similarities in the solutions. As a result, more general versions of cognitive schemata are created that contain knowledge that allows partial or complete solving of large classes of problems, or, more specific versions of schemata are created that allow solving of specific groups of problems. Mindful abstraction is a controlled, effort demanding process and will therefore increase cognitive load. However, because the increased cognitive load is directly relevant for learning (i.e. to the construction of cognitive schemata relevant for the performance of the task) it is considered a useful, *germane* cognitive load.

Jelsma and Bijlstra (1988) studied contextual interference for the training of complex cognitive skills, in particular, procedural training of troubleshooting skills. They reported delayed acquisition of troubleshooting skill but superior transfer performance for high contextual interference. De Croock, van Merriënboer, and Paas (1998) extended Jelsma and Bijlstra's (1988) study by also measuring the effects of contextual interference on cognitive load. They predicted that high contextual interference would induce higher — germane — cognitive load while solving practice problems, but eventually would lead to better cognitive schemata enabling higher transfer test performance. In their study learners first received information describing what kind of symptoms were caused by different types of system failures (failure principles) and then practised procedures for solving system failures under either high or low contextual interference. After practice they made a transfer test which consisted of problems in which combinations of failures were presented and thus procedures had to be combined. De Croock, van Merriënboer, and Paas (1998) found the same results as Jelsma and Bijlstra (1988): learners who practised troubleshooting procedures under high contextual interference showed delayed acquisition of the skill during practice but superior transfer performance. However, despite a trend in the expected direction, no concluding evidence was found for the hypothesis that learners who practised under high contextual interference had to invest more mental effort during practice, probably because the task they had to perform was too easy.

In the present article, Experiment 2 reports a new study in which a more difficult troubleshooting task is used. First, the system the learners have to work with is highly complex. Second, the nature of the troubleshooting task is such that learners do not learn ready-made procedures but instead are stimulated to develop device knowledge (i.e. knowledge about the principles of system functioning). They are trained to use these principles to explain and predict system behaviour so that they can diagnose new failures on the basis of their device knowledge, that is, by interpreting symptoms of system failures in terms of violations of system principles. Because this process requires more effort from the learner, it is predicted that the overall investment of mental effort will be higher than that observed by de Croock et al. (1998).

1.3. Redirecting attention and training efficiency

In the previous sections, it was hypothesised that some instructional methods may decrease extraneous cognitive load during training without negatively affecting transfer performance, because less attention is directed to cognitive processes not directly relevant for learning (e.g. weak method problem solving, integrating different sources of information). In a first experiment, this hypothesis will be tested for completion problems. Furthermore, it was hypothesised that some instructional methods may increase germane cognitive load during training and improve transfer performance because more attention is directed to cognitive processes directly relevant for learning (e.g. schema construction and mindful abstraction). In a second experiment, this hypothesis will be tested for high contextual interference.

An obvious next step is to combine these assumptions in the hypothesis of "redirecting attention", stating that instructional methods that combine a decrease in extraneous cognitive load with an increase in germane cognitive load will show superior transfer performance. For instance, instructional designers and teachers typically confront learners with conventional problems that are blocked in a simple-to-complex order, yielding a low contextual interference. According to the hypothesis of redirecting attention, higher transfer performance may be predicted for a sequence of completion problems with high contextual interference, because such a sequence both decreases extraneous cognitive load and increases germane cognitive load.

However, a major problem with the hypothesis of redirecting attention is that it becomes impossible to make clear predictions on cognitive load during training. To the authors' knowledge, there are currently no measurement instruments available that are capable of making a distinction between extraneous cognitive load and germane cognitive load. Learners can only report perceived mental effort for their task performance, including *both* aspects of cognitive load. For instance, it was argued in the example above that replacing a sequence of conventional problems with low contextual interference by a sequence of completion problems with high contextual interference will probably have a positive effect on transfer performance. But with regard to cognitive load, replacing the conventional problems with completion problems is expected to decrease — extraneous — cognitive load with an unknown quantity; and replacing low contextual interference with high contextual interference is expected to increase — germane — cognitive load with another unknown quantity. Consequently, the total effect on cognitive load during practice is unpredictable.

In an attempt to solve this problem, Paas and van Merriënboer (1993) introduced a computational approach to *training efficiency*, assuming that the learners' behaviour will primarily be more or less efficient under particular experimental conditions. Using this approach, Marcus, Cooper, and Sweller (1996) state that training is: (a) more efficient if the learners' transfer test performance after training is higher than might be expected on the basis of their invested mental effort during training; and (b) equivalent if their invested mental effort during training is lower than might be expected on the basis of their transfer test performance. The calculation of training efficiency is based on the standardisation of raw mental effort scores (i.e. estimates of cognitive load) and performance scores (e.g. accuracy or speed) across conditions to z-scores. The means of these standardised scores for each experimental condition are plotted as dots in a co-ordinate system. Thus, each dot refers to the mean mental effort z-score and related mean performance z-score of a particular experimental condition. Shifts to the upper left in this co-ordinate system indicate an increase in efficiency (i.e. higher performance in relation to less invested mental effort), and shifts to the lower right indicate a decrease in efficiency (i.e. lower performance in relation to more invested mental effort). Now, the training efficiency is calculated as the perpendicular distance from a certain point in the co-ordinate system to the hypothetical "baseline" condition in which each unit of invested mental effort equals out one unit of performance (i.e. the diagonal from the lower left to the upper right of the co-ordinate system, indicating a zero-efficiency). This distance can be calculated by the formula:

$$E = \frac{P - M}{\sqrt{2}}.$$

According to this expression, E is positive if $P > M$, negative if $P < M$, and zero if $P = M$. Thus, the more efficient the experimental condition, the higher the value of E.

The method can be illustrated on the basis of the data presented by Paas and van Merriënboer (1994). In this study, students had to learn a number of procedures in the domain of geometry. They had to solve conventional problems that imposed a high extraneous cognitive load under either high variability and thus high germane cognitive load conditions, or under low variability and thus low germane cognitive load conditions. Alternatively, they had to study worked examples that imposed a low extraneous cognitive load, again combined with either high variability (high germane cognitive load) or low variability (low germane cognitive load) conditions. As indicated in Fig. 1, strong effects on training efficiency were found. The traditional instructional method, consisting of conventional problem solving with low variability among problems (CONV-LVAR), yielded together with the conventional problem solving/high variability condition (CONV-HVAR) the lowest training efficiency (i.e. relatively high cognitive load during training combined with low transfer test performance). Redirecting the learners' attention, by replacing conven-

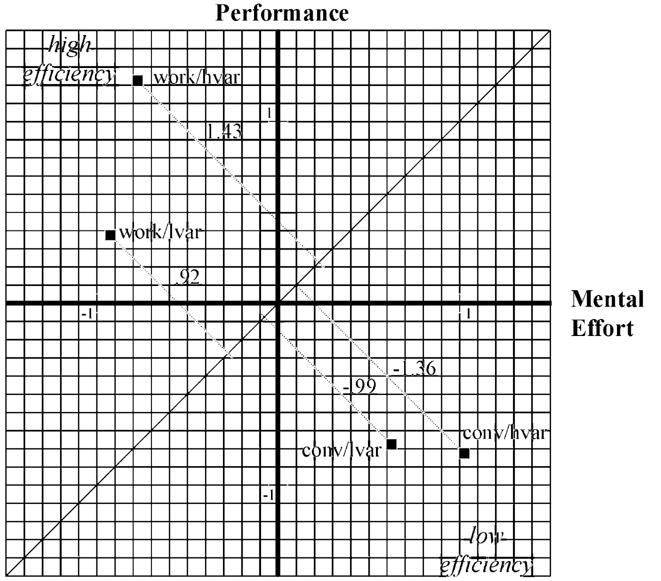

Fig. 1. Training efficiencies for the Paas and van Merriënboer (1994) study.

tional problems with worked examples and low variability with high variability (WORK-HVAR), yielded the highest training efficiency (i.e. relatively low cognitive load during training combined with the highest transfer test performance). In addition, Fig. 1 shows a clear superiority of worked examples above conventional problems.

The main advantage of this method is that it allows for predictions on a *combination* of changes in cognitive load and transfer performance. For example, the "extraneous load hypothesis" might predict that instructional methods decreasing extraneous load yield lower total cognitive load during training and at least an equal transfer test performance. But the same hypothesis might also predict that those methods do not effect total cognitive load during training, because learners now have the opportunity to invest extra effort in genuine learning, and yield a higher transfer test performance. It is important to note that both predictions can be tested as an increase in training efficiency. As another example, the "attention redirecting hypothesis" might predict that instructional methods that redirect attention yield lower total cognitive load during training and equal transfer test performance, equal load and higher performance, or even higher load and disproportionate higher performance. As argued above, it is impossible to predict which changes in total cognitive load will occur. But again, all three predictions will manifest themselves as an increase in training efficiency. In the present article, Experiment 3 reports a study in which training efficiency is studied in relation to completion problems and contextual interference by means of a 2×2 factorial experiment in the computer programming domain.

1.4. Hypotheses

Three experiments will be discussed. The focus of the first experiment is on decreasing extraneous cognitive load. Novice learners are trained in the generation of new computer programs according to three strategies. The completion strategy stresses the completion of partially given programs; the conventional strategy stresses the design and coding of new programs, and the learner-controlled strategy allows learners to switch between problem formats. The main hypothesis of this study is that perceived cognitive load during training is higher for the conventional group than for the other two groups. In addition, at least equal transfer test performance is predicted for the completion and learner-controlled groups compared to the conventional group. Such a finding would strongly support the extraneous load hypothesis, because it excludes the explanation that superior transfer test performance for learners who work with completion problems only results from changes in cognitive processing (e.g. replacing means–ends-analysis by mindful abstraction) and is unrelated to cognitive load as such.

The focus of the second experiment is on increasing germane cognitive load. Learners first study device knowledge and then practice applying this knowledge to diagnose system failures under either high or low contextual interference. It is hypothesised that the high contextual interference group will perceive higher cognitive load during training than the low contextual interference group. In addition, higher transfer test performance is predicted for the high contextual interference group. This may seem a trivial prediction, but it has not been fully supported in previous experiments concerned with contextual interference. In addition, together with the findings from Experiment 1 the results may indicate the viability of the distinction made in CLT between extraneous and germane cognitive load.

The focus of the third experiment is on redirecting the learners' attention by the simultaneous use of completion problems and high contextual interference in the computer programming domain. In line with Experiments 1 and 2, highest perceived cognitive load is expected for the group working with conventional problems under high contextual interference and lowest perceived cognitive load is expected for the group working with completion problems under low contextual interference. But more importantly, it is no longer possible to make predictions on cognitive load during training for learners in the group aiming at a redirection of attention (i.e. completion problems under high contextual interference). Therefore, the main hypothesis of this study is that redirecting the learners' attention will increase training efficiency. Training efficiency is predicted to be higher for the group with completion problems under high contextual interference, compared to the most traditional group working with conventional problems under low contextual interference. Furthermore, higher training efficiencies are predicted for the completion groups than for the conventional groups.

2. Experiment 1

In the programming domain, novice learners were trained in the design and coding of new computer programs according to three instructional strategies. The first strategy used conventional problems. Basically, learners had to write programs which conformed to a given problem specification, starting from a blank editor screen. The second strategy used completion problems. Learners had to finish partially completed programs present in the editor screen, which again had to conform to the given problem specification. As learners gradually became more proficient, they received increasingly incomplete programs until the problems became indistinguishable from conventional problems. The third strategy offered full learner control over problem formats. Learners could easily switch between conventional problems and completion problems according to their own preference.

Cognitive load was measured during training by a rating scale for perceived mental effort. After the training, learners received a transfer test for which they had to find plan errors in a computer program. The errors concerned violations of stereotyped patterns of programming code. From other research (e.g. Pennington, Nicolich, & Rahm, 1995), it is known that transfer from the generation to the interpretation of computer programs is a far from trivial task. Highest cognitive load during training combined with low transfer test performance is predicted for the conventional group.

2.1. Method

2.1.1. Participants

In the experiment, 26 first year Communication Science students, aged 19–26 years, were randomly assigned to a completion group (COMP; n=10), a conventional group (CONV; n=8), and a learner-controlled group (LC; n=8). Learners had some computer experience but no programming experience.

2.1.2. Materials

Participants received an introductory computer programming course for the language Comal (Christensen, 1982). The computer-based learning environment Completion ASsignment COnstructor (CASCO) was used; van Merriënboer, Krammer & Maaswinkel, 1994; van Merriënboer, Luursema, Kingma, Houweling, & de Vries, 1995; van Merriënboer & Luursema, 1995). CASCO presents the learner with a sequence of assignments, which may be either completion problems or conventional problems. Each next assignment is dynamically selected and constructed from a large database of problems, based on the performance of the individual student. Assignments consist of a problem statement: (1) *explanations*, concerning new programming concepts that are necessary for solving the problem; (2) specific *tasks* that may help to solve the problem; and (3) *questions* that are relevant for the problem at hand. It was not possible to evade questions and tasks, because CASCO would prompt the learners with respect to unfinished or incomplete actions. For completion problems, a partial, to-be-completed program is presented in a full-fledged editor window; for conventional problems, this editor window is empty. Learners were free to compile their code and test and debug their programs at any time.

The CASCO interface is presented in Fig. 2. The large upper window right below the pull-down menu is the editor. Explanations, tasks, and questions are presented in the bottom left window; if they pertain to particular pieces of the programming code, this code is highlighted in the editor screen. The learners give answers to questions in the bottom right window. For the learner-controlled condition, learners can switch between conventional problems and completion problems as indicated by the dialogue box in the centre of the screen. The interface of the conventional and the completion condition was different from the learner-controlled condition only with respect to the missing “switch” option in the pull-down menu.

The amount of invested mental effort was estimated with a 9-point symmetrical rating scale. This rating scale was a modified version of the scale presented by Bratfisch, Borg, and Dornic (1972) for measuring perceived task difficulty. The scale has numerical values and labels assigned to categories, ranging from very, very low mental effort (1) to very, very high mental effort (9). In earlier research, this rating scale proved to be highly reliable (for an overview, see Paas, van Merriënboer, & Adam, 1994).

The transfer test consisted of a troubleshooting task. The learners were presented with a problem statement and a program containing 10 plan errors in stereotyped patterns of programming code (e.g. plans for looping structures, initialisation, conditional actions, etc.). They were informed about the number of errors in the program.

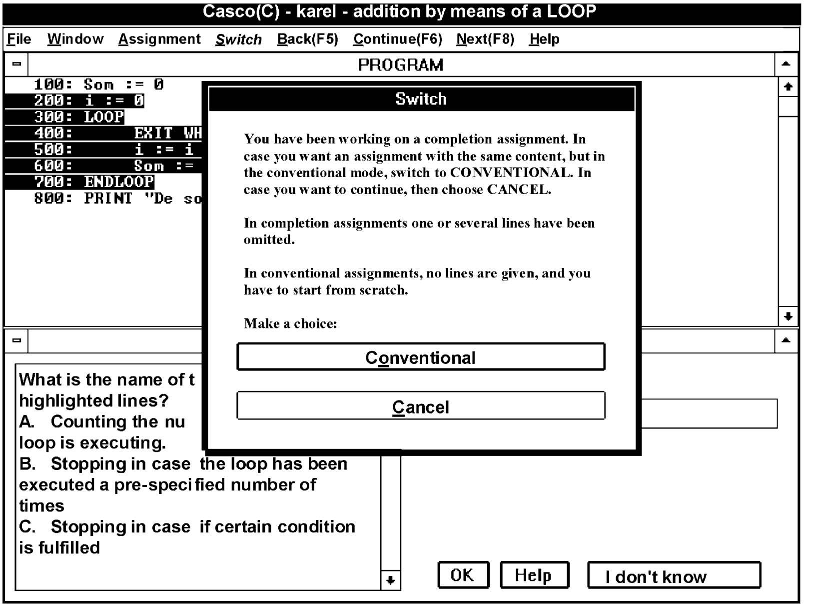

Fig. 2. The CASCO interface.

To detect the errors, they were obliged to study the program and relate it to the corresponding specification. It was not possible to detect errors by compiling the code.

2.1.3. Procedure

All learners were informed about the experimental procedure and received a demonstration and written guidelines on working with CASCO. Learners worked with CASCO for 180 min in either the completion, conventional or learner-controlled condition. An assistant was present during the whole experiment. While working with CASCO, a window with the 9-point rating scale popped up with intervals of 20 min. Learners had to rate the amount of perceived mental effort on this scale. The transfer test was presented one hour after practice had finished. Thirty minutes were available for completing the test.

2.2. Results

The results are summarised in Table 1. Practice data show that learners in the completion condition finished the largest number of problems in the 3-h practice phase (M=28.1), compared to the conventional condition (M=8.3) and the learner control condition (M=21.3). ANOVA confirmed the significance of the differences in numbers of completed assignments, $F(2,23)=13.7$, MSE=66.4, $p<0.001$. In post-hoc tests, using Tukey's HSD, it was found that both the conventional and the completion group ($p<0.001$) and the conventional and the learner control group ($p<$

Table 1
Summary of results for experiment 1

	Conventional (n=8)		Completion (n=10)		Learner control (n=8)	
	M	SD	*M*	SD	*M*	SD
Practice behaviour						
# Completed assignments	8.3	4.8	28.1	11.1	21.3	5.0
Cognitive load (1–9)[a]	6.3	1.6	4.7	1.0	5.5	1.1
Transfer test						
Proportion of errors found	0.33	0.17	0.39	0.21	0.55	0.19

[a] Measured by a 9-point mental effort rating scale that was presented each 20 min. Score: 4 — “rather low mental effort”; 5 — “neither low nor high mental effort”; and 6 — “rather high mental effort”.

0.01) differed significantly; there was no significant difference between the completion and learner control group. In the learner control group, learners switched between problem formats (i.e. between completion and conventional problems) on average 5.5 times (SD=4.3). They spent a mean of 76% of their total practice time working on completion problems (SD=21.2), showing a clear preference for completion problems above conventional problems.

For the perceived amount of mental effort related to the practice problems, ANOVA indicated a significant effect, $F(2,23)=4.7$, MSE=1.7, $p<0.05$. As predicted, the highest cognitive load was reported by the conventional group (M=6.3), followed by the learner control group (M=5.5) and the completion group (M=4.7). Post-hoc tests, using Tukey’s HSD, indicated that only the conventional and completion groups differed significantly, $p<0.05$.

Finally, the proportion of correctly identified plan errors in the transfer test was 0.33 for the conventional group, 0.39 for the completion group, and 0.55 for the learner control group. ANOVA indicated a significant difference between conditions, $F(2,23)=3.6$, MSE=0.0318, $p<0.05$. As predicted, the completion and learner-controlled groups outperformed the conventional group, but this difference only appeared to be significant for the learner control group.

2.3. Discussion

The results clearly support the extraneous load hypothesis. As predicted, cognitive load during training, measured as perceived mental effort, was higher in the conventional group than in the completion group. Nonetheless, the completion group showed equal transfer test performance; although not statistically significant, there was even a tendency for the completion group to outperform the conventional group. This is in contrast with the results of Paas (1992), who found no difference in cognitive load during training between a conventional and completion group. This may be related to the fact that participants in his study rated perceived mental effort for the training problems as (very) low, which may have caused a bottom effect. In our study, training problems were experienced as fairly difficult, with perceived mental effort ratings between 4.7 (completion) and 6.3 (conventional), where 5 indicates a "neither low nor high mental effort" and 6 a "rather high mental effort".

It seems that learners in the completion group did not fully use the freed-up processing capacity for processes directly relevant to learning (i.e. increasing germane cognitive load). Thus, completion problems do not seem to have a direct effect on germane cognitive load. However, another pattern is found for the learner-controlled group. Here, the reported mental effort during training is not significantly lower than for the conventional group, but superior transfer test performance is reached. It is difficult to explain this finding as an effect of instructional methods only, as both the completion group and the learner-controlled group spent most of their time working on completion problems. As a possible explanation, giving learners control over the type of problems might increase their task involvement, so that they are more inclined to invest germane cognitive load in learning. As a result, the learner-controlled group does not show a significant difference in perceived mental effort during training compared to the conventional group, because they re-invest the freed-up processing resources in genuine learning, but it does show superior transfer test performance.

To summarise this section, clear support is found for the extraneous load hypothesis. It is possible to reduce extraneous cognitive load during training, indicated by a lower perceived mental effort, and still maintain at least an equal transfer test performance. In addition, learner control may give learners the opportunity to optimise their learning. It may stimulate them to re-invest freed-up processing resources in schema construction, that is, increase their germane cognitive load and so improve transfer test performance. Instructional manipulations explicitly aimed at an increase of germane cognitive load are studied in the next experiment.

3. Experiment 2

In this experiment, contextual interference is studied using a training approach that aims for far transfer of troubleshooting skills. Learners are stimulated to develop device knowledge, that is, knowledge about the principles of system functioning and knowledge about how these principles can be used to explain and predict system states. On the basis of this knowledge diagnosis of system failures can then be performed by identifying the components in the system that do not function according to the principles that they were designed for. Learners first study system principles and then practice troubleshooting skills by solving a series of problems in which they have to diagnose a system failure. A high contextual interference group and a low contextual interference group were created that differ with respect to the order in which practice problems are presented. In the high contextual interference group each successive practice problem requires different causal reasoning with system principles. In the low contextual interference group, blocks of practice problems are presented where in each block the causal reasoning with system principles, required to diagnose the system failure, is identical.

It is predicted that the participants in the high contextual interference group will acquire better cognitive schemata for troubleshooting, that is, better organisation of knowledge about classes of system failures and their relation to device knowledge. They will be able to use this knowledge more flexibly and therefore perform better in transfer situations. However, the inductive processes that are promoted under high contextual interference require mindful engagement of the learners which increases cognitive load. We therefore predict that during practice, learners in the high contextual interference group will need more time to finish the practice tasks, will show a lower overall practice performance, and will invest more mental effort compared to the low contextual interference group.

3.1. Method

3.1.1. Participants

In the experiment, 69 first year Engineering students, aged 18–22 years, were randomly assigned to a low contextual interference group (LCI; *n*=34) and a high contextual interference group (HCI; *n*=35). Most learners had some knowledge of process technology but no experience with troubleshooting dynamic chemical systems.

3.1.2. Materials

The learners practised troubleshooting skills with DISTILLER, a computer-based simulation of a water–alcohol distillery plant (de Croock, 1999). The simulator consists of a control room computer-display system for process control (see Fig. 3), that looks and works just like actual display systems in distillery plant control rooms. In addition to the display system, DISTILLER contains a software distillery plant-model that causes the display system to operate and respond just like a display system

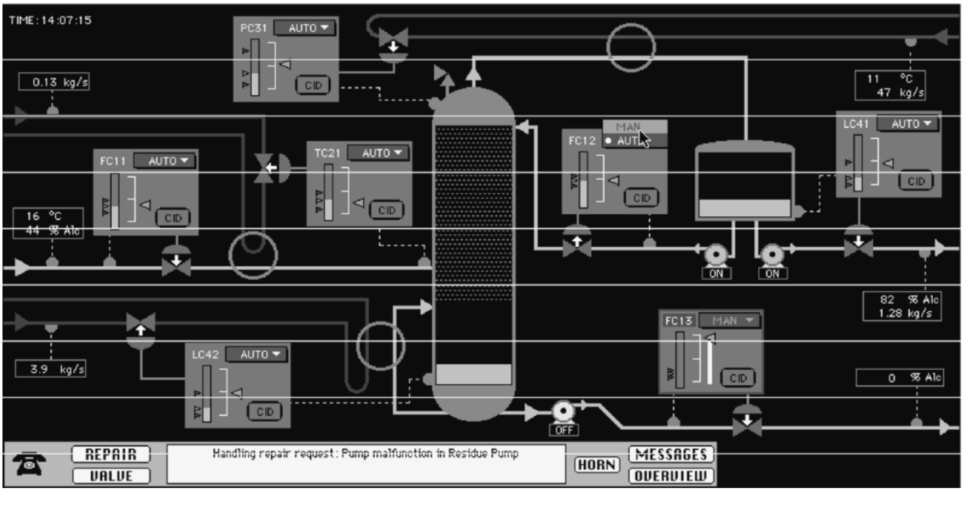

Fig. 3. The DISTILLER interface.

operating and responding to a real plant would do. In the plant-model system failures can be simulated, which cause disturbances and alarm situations on the control display. It is the learner's task to identify the component in the plant and the type of failure that caused the system failure.

Participants were introduced to their task by a computer-assisted instructional program (CAI program) providing general information on troubleshooting and explaining the distillation system. In order, the following topics were presented by the CAI-program: (a) the topology of the distillation system and its main functional parts; (b) the working of the components of the system (system principles); (c) failures in the system; (d) troubleshooting the system; and (e) operating the system. Note that for topic c only an explanation was given of what is meant with the terms on the malfunction list from which the learners have to choose when they report a failure. No information was given about the effects of the malfunctions on system functioning and about the symptoms they cause.

Following the information presentation stage participants practised a series of troubleshooting problems under either a high or a low contextual interference condition. Each troubleshooting problem required learners to study a disturbed production process and to identify the component that caused the disturbance. A problem is solved as soon as the correct malfunctioning component is identified. Practice for the LCI group consisted of four blocks of five problems. In each block a different type of malfunction had to be diagnosed (i.e. a PID controller malfunction, a leakage in a conduit, a sensor malfunction, or a valve energy failure) and for each of the five problems within a block, the malfunctions occurred in different components of the system. The design was counterbalanced for presentation order of blocks. In the HCI group, the 20 problems were presented in a randomised order.

At the end of each problem, learners were asked to rate the amount of invested mental effort on a rating scale before the next problem was presented. The rating scale was identical to the scale used in Experiment 1. Immediately after practice a transfer test was administered which consisted of eight transfer problems, each requiring the participants to diagnose a *new* type of failure within 6 min.

3.1.3. Procedure

An assistant was present during the whole experiment. At the onset of the experiment, participants were first introduced to their task by the CAI program. Following this information presentation stage they practised 20 troubleshooting problems under either a high or a low contextual interference condition. For each problem, 6 min were available to diagnose the malfunction, except if the particular malfunction was presented for the first time, in which case 10 min were available. At the end of each problem, participants were asked to rate the amount of invested mental effort on the rating scale before the next problem was presented. Immediately after practice the transfer test was administered. The complete experimental procedure took about 4 h.

3.2. Results

The results are summarised in Table 2. Practice data show that the HCI group needed more time (M=80.6) to complete the practice problems than the LCI group

Table 2
Summary of results for experiment 2

	Low contextual interference (n=34)		High contextual interference (n=35)	
	M	SD	*M*	SD
Practice behaviour				
Time (min.)	66.1	29.1	80.6	11.0
# Incorrect diagnoses	1.15	.85	1.55	1.8
Cognitive load (1–9)[a]	4.85	.95	5.45	0.65
Transfer test				
Time (min.)	27.6	5.6	27.7	4.0
# Incorrect diagnoses	1.6	1.0	1.2	0.75

[a] Measured by a 9-point mental effort rating scale that was presented after each problem. Score: 4 — "rather low mental effort"; 5 — "neither low nor high mental effort"; and 6 — "rather high mental effort".

(M=66.1); $F(1,67)=21.47$; MSE=168.27, $p<0.01$. There was also a marginally significant difference on the mean number of incorrect diagnoses, $F(1,67)=3.92$; $MSE=0.88$; $p=0.052$. The HCI group made more incorrect diagnoses (M=1.55) than the LCI group (M=1.15). Furthermore, a significant effect was found for perceived mental effort, $F(1,67)=7.53$; MSE=0.69; $p<0.01$. As predicted, the HCI group reported a higher cognitive load (M=5.45) than the LCI group (M=4.85).

The time needed to complete the transfer test did not differ between groups. The HCI group needed 27.7 min and the LCI group 27.6 min for completing the test. For transfer test performance, there was a trend in the expected direction: the HCI group made less incorrect diagnoses (M=1.2) than the LCI group (M=1.6); $F(1,67)=3.26$; MSE=0.81; $p<0.10$.

3.3. Discussion

Regarding the effects of contextual interference on performance, both during practice and on the transfer test, the data of this experiment are in agreement with the results found in the study by de Croock, van Merriënboer, and Paas (1998). In addition, evidence is found for the hypothesised effects of contextual interference on cognitive load. The practice data show that, as expected, the high contextual interference group had to work harder than the low contextual interference group. Learners in the high group needed more time to complete the practice cases, they made more incorrect diagnoses and they reported higher investment of mental effort in trying to solve the practice cases. Most importantly this extra effort seemed to pay off. There was a trend indicating that they performed better on the transfer test by making fewer errors.

In contrast with the study by de Croock et al. (1998), learners in this study did not have to diagnose new combinations of known failures for which they had already learned a procedure. Instead they had to diagnose completely new failures, and thus had to construct a new procedure based on the knowledge about the system and system failures they had developed so far. The data on the transfer test indicate that the high interference group probably had better knowledge for constructing new troubleshooting procedures than the low interference group. So it is therefore justified to reason that the higher cognitive load induced in the high interference group, was indeed a germane cognitive load that stimulated learning processes that resulted in cognitive schemata relevant for transfer of troubleshooting skill.

It is further interesting to note that the mental effort data showed that learners perceived the task in the current experiment as more difficult than the task in the de Croock et al. (1998) study. The investment of mental effort during practice for this study ranged between 4.8 (LCI) and 5.4 (HCI) whereas for the de Croock et al. (1998) study it ranged between 3.3 (LCI) and 3.8 (HCI). Possibly the task has to be sufficiently difficult for difference in investment of mental effort between high and low contextual interference training formats to occur. Experiment 3 will study combined effects of high contextual interference and completion problems.

4. Experiment 3

In the programming domain, novice learners were trained in the design and coding of new computer programs according to four instructional strategies. The experiment used a 2×2 factorial design, with the factors Problem Format (completion, conventional) and contextual interference (low, high). The practice for the completion groups and conventional groups was nearly identical to the practice described in Experiment 1. However, a distinction was made between low and high contextual interference. For low contextual interference, each new problem that was presented to the learner assumed the use of relatively many programming constructs also practised in the previous problem and introduced only a few new programming constructs. For high contextual interference, each new problem required the use of relatively few already known programming constructs and thus introduced relatively many new programming constructs.

Cognitive load was measured for each problem during training by a rating scale for perceived mental effort. After the training, learners received a transfer test for which they had to design and code computer programs in Comal (near transfer), Pascal (intermediate transfer), and C (far transfer). In line with Experiments 1 and 2, highest cognitive load during training is predicted for the conventional-high interference group. More importantly, low training efficiency is predicted for the conventional-low interference group (i.e. the “traditional” instructional strategy) and superior training efficiency is predicted for the completion-high interference group. In this last group, the learners’ attention is redirected from cognitive processes not relevant for learning to processes that are directly relevant for the construction of cognitive schemata.

4.1. Method

4.1.1. Participants

In the experiment, 87 first year students in Educational Science and Communication Science, aged 18–23 years, were randomly assigned to a completion-low contextual interference group (COMP-LCI, n=22), a completion-high contextual interference group (COMP-HCI, n=25), a conventional-low contextual interference group (CONV-LCI, n=20), and a conventional-high contextual interference group (CONV-HCI, n=20). Learners had some computer experience but no programming experience.

4.1.2. Materials

Participants received an introductory computer-programming course for the language Comal. The same computer-based learning environment, CASCO, was used as in Experiment 1 (see Fig. 1). All groups received assignments consisting of a problem statement, followed by explanations, questions and tasks. As in Experiment 1, the completion groups also received partial, to-be-completed programs while the conventional groups had to design and code their programs from scratch. In addition, CASCO enabled easy modification of contextual interference by setting two parameters in its instructional model. The first parameter determines the degree of relatedness between previously presented problems and new problems. For high contextual interference, a next to-be-written program may contain many programming constructs that were not present in the previous problem; for low contextual interference, a next to-be-written program contains only few new programming constructs. Programming constructs are plans, or patterns of programming code for "initialising a variable", "asking for input", "printing results", "defining a function", "putting conditions to an action", and so forth. CASCO distinguished 36 constructs. The second parameter determines the number of candidate problems from which a new to-be-presented problem is selected. For high contextual interference, the number of candidate problems is high; for low contextual interference, this number is low (see Schuurman, 1999 for a detailed description of parameter settings).

For each finished assignment, learners had to rate their perceived amount of mental effort on a 9-point rating scale, which was identical to the scales used in Experiment 1 and 2.

The transfer test consisted of three problem statements or tasks for which a program had to be written. For the first task the program had to be written in Comal; for the second task in Pascal, and for the third task in C. Learners received a very short manual related to Pascal and C programming constructs and syntax which they could consult during the tests. The first task was considered to be a near transfer task; the second task an intermediate transfer task (Pascal is closely related to Comal), and the third task a far transfer task (C is more different from Comal). Since each programming task should not be repeated twice, three analogue programming problems were constructed. These were presented in random order and balanced over the experimental groups.

Written programs for the transfer test were scored with respect to a correct use of programming constructs. Two expert programmers scored a subset of the data (25 randomly chosen participants) independently from each other. It was up to the programmers to decide whether a particular programming construct was — correctly — used. The inter-observer reliability was ρ=0.72, which is rather low but acceptable.

4.1.3. Procedure

All learners were informed about the experimental procedure and received a demonstration and written guidelines on working with CASCO. Learners worked with CASCO for 180 min in one of the four conditions. An assistant was present during the whole experiment. After finishing each assignment, learners had to rate the amount of invested effort on the 9-point rating scale. The transfer test was presented one hour after practice had finished (there was a 1 h lunch break between the training and the test). There were 60 min available for completing the test: 20 min for each of the three tasks.

4.2. Results

The results are summarised in Table 3. Practice data show that the completion condition (M=15.88) finished more assignments than the conventional condition (M=4.71); $F(1,83)=85.87$; MSE=31.16; $p<0.001$. For the number of practised assignments, there was no effect for contextual interference and no interaction.

For the perceived amount of mental effort related to the practice problems, ANOVA indicated a significant effect for problem format, $F(1,83)=16.58$;

Table 3
Summary of results for experiment 3

	Conventional				Completion			
	Low CI (*n*=20)		High CI (*n*=20)		Low CI (*n*=22)		High CI (*n*=25)	
	M	SD	*M*	SD	*M*	SD	*M*	SD
Practice behaviour								
# Completed assignments	5.4	1.8	4.0	1.6	15.3	6.3	16.4	8.3
Cognitive load (1–9)[a]	5.4	1.1	6.1	1.1	5.5	0.7	4.6	0.4
Transfer test								
Comal schemata	0.28	0.20	0.37	0.09	0.39	0.15	0.34	0.17
Pascal schemata	0.19	0.18	0.35	0.11	0.32	0.15	0.29	0.18
C schemata	0.20	0.14	0.24	0.18	0.31	0.13	0.28	0.17
Training efficiency								
Based on Comal	−0.34	1.04	−0.45	0.68	0.12	0.71	0.52	0.95
Based on Pascal	−0.46	0.92	−0.28	0.91	0.07	0.73	0.53	0.89
Based on C	−0.32	1.04	−0.65	1.04	0.15	0.76	0.65	0.88

[a] Measured by a 9-point mental effort rating scale that was presented after each problem. Score 4 — "rather low mental effort"; 5 — "neither low nor high mental effort"; and 6 — "rather high mental effort".

MSE=0.70; $p<0.001$. As predicted, cognitive load in the conventional condition (M=5.8) was higher than in the completion condition (M=5.0). There was no main effect for contextual interference but an unexpected interaction occurred. There was no difference between the conventional group and the completion group for low contextual interference (in order, M=5.4 and M=5.5). However, perceived mental effort under high contextual interference increased for the conventional group (M=6.1) but decreased for the completion group (M=4.6; $F(1,83)=18.60$; MSE=0.70, $p<0.001$).

In order to make a comparison possible, Table 3 presents standardised scores on the interval [0–1] for the Comal, Pascal, and C transfer tests, indicating in order near, intermediate, and far transfer. Problem format yielded a significant effect on the C transfer test, $F(1,83)=5.87$; MSE=0.02; $p<0.025$. As expected, far transfer performance was higher for the completion condition (M=0.30) than for the conventional condition (M=0.22). The same tendency can be observed for the near (Comal) and intermediate (Pascal) transfer tests, but these differences did not reach significance. Contextual interference yielded no significant main effects on transfer test performance. There was only a marginally significant effect on the Pascal transfer test, $F(1,83)=3.39$; MSE=0.02; $p<0.1$, indicating higher performance for high contextual interference (M=0.32) than for low contextual interference (M=0.26). A significant interaction between problem format and contextual interference was found for the Comal and Pascal transfer tests with $F(1,83)=4.05$; MSE=0.02, $p<0.05$, and $F(1,83)=8.19$, MSE=0.02, $p<0.005$, respectively. High contextual interference positively affected the performance of the conventional groups. For Comal M=0.28 and M=0.37 for, in order, the conventional low interference and the conventional high interference group; for Pascal M=0.19 and M=0.35 for, in order, the conventional low interference and the conventional high interference group. In contrast, high contextual interference had no positive effect on the performance of the completion groups. For Comal M=0.39 and M=0.34 for, in order, the completion low interference and the completion high interference group; for Pascal M=0.32 and M=0.29 for, in order, the completion low interference and the completion high interference group. The same pattern can be observed for the C transfer test, but this interaction did not reach significance.

Training efficiencies were computed on the basis of perceived mental effort during training and performance on either the near, intermediate or far transfer test. As indicated in Fig. 4, the three training efficiencies based on the Comal, Pascal and C transfer tests showed almost the same pattern among experimental groups. Problem format showed a highly significant effect on all computed training efficiencies. The completion condition reached higher values than the conventional condition for the efficiency based on Comal performance ($F(1,83)=14.71$; MSE=0.74; $p<0.001$; M=0.33 for completion and $M=-0.39$ for conventional). For the efficiency based on Pascal performance ($F(1,83)=13.17$; MSE=0.75; $p<0.001$; M=0.32 for completion and $M=-0.37$ for conventional). And for the efficiency based on C performance ($F(1,83)=19.55$; MSE=0.87; $p<0.001$; M=0.41 for completion and $M=-0.49$ for conventional). Contextual interference yielded no significant main effects on training efficiencies. There was only a marginally significant effect for efficiency based on

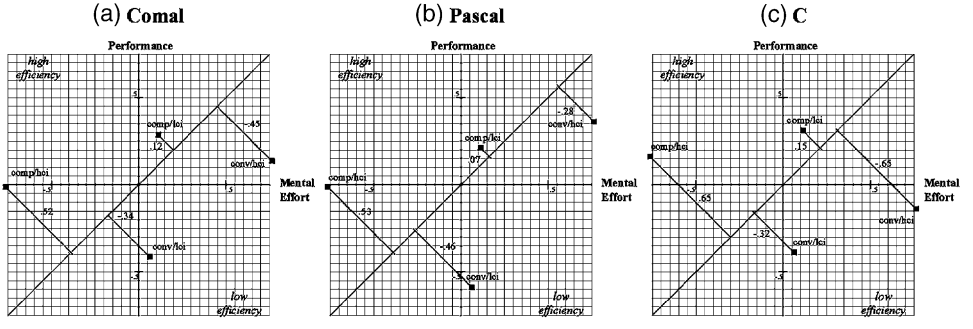

Fig. 4. Training efficiencies based on perceived mental effort during training and the (A) Comal; (B) Pascal; and (C) C transfer tests.

Pascal performance, $F(1,83)=2.93$; MSE=0.75; $p<0.1$, indicating a higher efficiency for the high contextual interference condition (M=0.17) than for the low contextual interference condition (M=−0.18). Finally, there was a significant interaction for the efficiency based on far transfer performance, $F(1,83)=4.22$; MSE=0.87; $p<0.05$. High contextual interference negatively affected the efficiency of the conventional groups (M=−0.32 and M=−0.65 for, in order, low and high contextual interference) but positively affected the efficiency of the completion groups (M=0.15 and M=0.65 for, in order, low and high contextual interference; see also Fig. 4c).

4.3. Discussion

With regard to the use of completion problems, the results clearly support the extraneous load hypothesis. As predicted, cognitive load during training was higher in the conventional condition than in the completion condition. Nonetheless, the completion condition showed equal transfer test performance for the near (Comal) and intermediate (Pascal) transfer test, and even higher performance for the far (C) transfer test. This latter finding provides additional support to the idea that decreasing extraneous cognitive load may be helpful to the construction of cognitive schemata, because such schemata are particularly important for reaching further transfer. The positive combined effect on cognitive load during training and transfer test performance is clearly reflected in the training efficiencies. All training efficiencies, computed for near, intermediate, and far transfer test performance are higher for the completion condition than for the conventional condition.

The results for contextual interference are less clear. There was no overall effect on cognitive load during training. As will be discussed below, this is mainly due to the fact that load in the conventional condition was higher for the high interference group than for the low interference group, but in the completion condition load was higher for the low interference group than for the high interference group. Nonetheless, tendencies on transfer test performance and training efficiencies were in the expected direction. For the intermediate transfer test (Pascal) and the training efficiency computed on the basis of this test, a marginal effect was found in favour of high contextual interference.

The main hypothesis of this experiment concerned the effect of redirecting attention on training efficiency. In support of this hypothesis, a significant interaction was found for the efficiency computed on the basis of the far transfer test and perceived mental effort during training. Replacing low contextual interference with high contextual interference had a positive effect on training efficiency for the completion condition, but a negative effect for the conventional condition. Consequently, the highest training efficiency is found for the completion-high interference group and the lowest efficiency is found for the conventional-high interference group (see Fig. 4c). Those training efficiencies are in full agreement with the results of Paas and van Merriënboer (1994); see also Fig. 1). While no significant interactions occurred on training efficiencies computed on the basis of near and intermediate transfer test performance, the results show a similar pattern. As shown in Fig. 4a and b, training efficiencies are again highest for the completion-high interference group, that is, for the group in which attention is redirected. Significant effects are found for efficiencies computed on the basis of far transfer test performance, but not for efficiencies computed on the basis of near and intermediate transfer test performance. This may possibly be explained by the finding that instructions based on CLT typically have the strongest effect on far transfer (see also Paas & van Merriënboer, 1994). An explanation for this is that better organised cognitive schemata are mainly useful when learners must deal with new situations.

The results of redirecting attention on training efficiency are promising, but at the same time its effects on near and intermediate transfer test performance are disappointing. The observed interactions for the Comal and Pascal transfer tests indicate a positive effect of high contextual interference for the conventional condition, but a zero or negative effect for the completion condition. Consequently, the “traditional”, conventional-low interference group consistently shows the lowest transfer test performance but the conventional-high interference group reached similar scores on the near and intermediate transfer tests as the completion groups. In order to explain this finding, it is necessary to take a closer look at the conventional and completion condition separately.

For the conventional condition, high contextual interference ensured that each new problem required the application of relatively many new programming constructs. Because learners only worked on a small set of problems (much smaller than the set of problems in the completion condition), each problem was probably experienced as “new” which may have increased the learners’ task involvement and willingness to invest effort in learning. This extra effort seems to pay off, yielding a higher transfer test performance for the high contextual interference group than the low contextual interference group.

For the completion condition, high contextual interference decreased the perceived mental effort during training and had a small but consistent negative effect on transfer test performance. This is a counter-intuitive finding. A possible explanation may be related to the nature of completion problems, which integrate the presentation of information (i.e. new programming constructs are illustrated in the partial program) and practice. On the one hand, high contextual interference ensured that each new partial program introduced the learners to relatively many new programming constructs that were illustrated in the partial program. But as a result of this, learners only had to complete relatively small program parts. While the interference in the presented information is high, interference between programming constructs that are simultaneously practised is low. Low contextual interference, on the other hand, ensured that each partial program presented to the learners introduced relatively few new programming constructs. As a result, learners had to complete relatively large program parts. While the interference in the presented information is low, interference between programming constructs that are actually practised is high. The point is that the *experienced* interference in the completion condition may have been higher for the low interference group than for the high interference group. This potential problem with making contextual interference operational will be further discussed in Section 5.

5. General discussion

This article described three experiments. The first experiment studied the decrease of extraneous cognitive load by the use of completion problems. The second experiment studied the increase of germane cognitive load by the use of high contextual interference. And the third experiment studied the redirection of attention by simultaneously decreasing extraneous cognitive load and increasing germane cognitive load. In this section we discuss, in order, the integrated findings of the three experiments with regard to problem formats, contextual interference, and redirecting attention. To conclude this article, directions for future research and practical implications are discussed.

With regard to problem formats, the results of Experiments 1 and 3 are fully in line with each other. In both experiments, completion problems yielded lower cognitive load during training than conventional problems. Nonetheless, training with completion problems yielded at least equal transfer test performance: tendencies were in favour of the completion problems and in Experiment 3 there was a positive effect of completion problems on generating computer programs in a different programming language (i.e. far transfer). The positive combined effects on cognitive load during training and transfer test performance are clearly reflected in computed training efficiencies, which are consistently higher for completion problems than for conventional problems.

These results provide strong support to the extraneous load hypothesis in relation to completion problems. The superiority of completion problems cannot solely be explained by a change in cognitive processing, such as replacing means–ends-analysis in problem solving by mindful abstraction from partial solutions, but also seems to be related to cognitive load as such. Thus, extraneous cognitive load proved to be a useful concept for instructional design. At the same time, the results of Experiments 1 and 3 show that the positive effects of completion problems on transfer performance are limited and mainly manifest themselves for far transfer. It may be necessary to give learners the opportunity to optimise their allocation of cognitive resources for learning in order to reach more stable effects on transfer performance. As indicated in Experiment 1, offering control to learners over problem formats may possibly help to do so. In the learner-controlled group, learners worked most of their time on completion problems but also spent some time working on conventional problems. As a result, cognitive load during training did no longer differ from the conventional condition but superior transfer test performance was reached. Thus, offering control to learners over problem formats may give them better opportunities to increase their germane cognitive load.

Contextual interference was studied in Experiments 2 and 3 as a more direct manipulation to increase germane cognitive load. In both experiments, there were trends on transfer test scores indicating higher performance for high contextual interference than low contextual interference. In Experiment 3, there was also a trend on training efficiency in favour of high contextual interference. In Experiment 2, learners working under high contextual interference reported higher perceived mental effort during practice, which supports the germane load hypothesis. However, in Experiment 3 no difference in cognitive load during training was found between high and low contextual interference. On closer inspection of the data, it turns out that in the conventional condition perceived mental effort during training was, as expected, higher for the high contextual interference group than the low contextual interference group; however, it showed the opposite effect for the completion condition. Furthermore, in the conventional condition the high contextual interference group outperformed the low contextual interference group on the transfer tests. Thus, the findings of Experiment 2 and the conventional condition of Experiment 3 are in full agreement with each other.

To summarise the findings on contextual interference, our results provide some preliminary support to the germane load hypothesis in relation to high contextual interference, but *only* when interference is applied in combination with conventional problems. For high contextual interference to be effective, it may be necessary that learners have to mentally integrate a relatively large set of newly acquired knowledge structures during actual practice. For the conventional problems as used in Experiments 2 and 3, the presentation of information (i.e. knowledge structures such as system principles or programming constructs) and practice (i.e. conventional problem solving) are clearly separated from each other. And for high contextual interference, the set of presented knowledge structures that must subsequently be applied and mentally integrated while solving the problem is relatively large compared to low contextual interference. We assume that the key element in contextual interference is indeed the size of the set of knowledge structures that must be mentally integrated during actual practice. Below, this assumption will be used to explain why our particular way to make high contextual interference operational may not have produced the expected results on transfer test performance for completion problems.

The effects of redirecting attention on training efficiency were studied in Experiment 3. For reaching far transfer performance, training efficiency was highest for the completion group working under high contextual interference, that is, the group in which attention was redirected. For near and intermediate transfer performance the same tendency on training efficiencies could be observed. These results on training efficiencies are in full agreement with the findings of van Merriënboer and Paas (1994), who used worked examples with a high variability in presentation formats as a combined method to redirect the learners' attention. It may be concluded that redirecting attention is a promising approach to improve training efficiency.

A surprising result of Experiment 3 is that the higher training efficiency of the completion-high interference group is not reflected in higher transfer test performance. Above, it was assumed that the key element in contextual interference might be the size of the set of knowledge structures that must be mentally integrated during actual practice. This may explain the unexpected effect of contextual interference in combination with completion problems. For high contextual interference, relatively many new programming constructs were presented in each new problem and illustrated in the partial program. Because of this, learners only had to complete small parts of the program, that is, interference in the actual application of newly acquired knowledge structures is *low*! For low contextual interference, in contrast, relatively few new programming constructs were presented in each new problem and illustrated in the partial program. Consequently, learners had to complete larger parts of the program, that is, interference in the actual application of newly acquired knowledge structures is *high*! In summary, the particular way in which contextual interference was made operational may have yielded the expected effect for conventional problems but the opposite effect for completion problems.

Some support for this line of reasoning is provided by our results. First, it explains the surprisingly low perceived mental effort during training in the completion-high interference group. In this group, the completion problems required not much actual coding from the learners and most resembled fully worked examples. Other research has shown that worked examples impose a relatively light cognitive load on learners (Paas & van Merriënboer, 1994; van Gerven, Paas, van Merriënboer, & Schmidt, 2002). Second, it explains the good transfer performance of the completion-low interference group. Actually, it was this group that was both confronted with useful, partially worked examples *and* high interference in the application of newly learned programming constructs while completing those examples.

The studies reported in this article point out several directions for future research. First, there are some preliminary indications that giving learners control over problem formats may help them to optimise their allocation of cognitive resources for learning. We are not aware of studies that connect learner control to CLT, but a further exploration of this finding may provide a good link between working memory models, preferred styles of students who may choose between different problem formats, and theories of self-regulated learning. Second, research is needed to specify the conditions under which high contextual interference is effective and to find out how it can best be made operational in computer-based learning environments. As an example, one might study for completion problems the conditions that distinguish interference between partially presented programs from (mental) interference between newly acquired programming constructs while completing those programs. Finally, this study focused on completion problems and high contextual interference as two particular instructional methods that may be combined to redirect learners' attention. Future research on redirecting attention should also study other combinations of instructional methods.

The practical implications of our studies are quite straightforward. By now, there is strong empirical evidence for the superiority of completion problems over conventional problems when teaching design tasks at an introductory level. The only drawback of this approach is that good completion problems are more time-consuming to develop than conventional problems. For more advanced learners, who practice with conventional problems, high contextual interference must be preferred over low contextual interference. This is an important finding, because teachers and instructional designers typically confront learners with problems that are blocked in a simple-to-complex order, yielding low contextual interference. And finally, our results indicate that minimising extraneous cognitive load is only one side of the coin. If teachers or instructional designers succeed in this, they should also search for viable methods to increase the learners' germane cognitive load, the other side of the coin. As an alternative to contextual interference, variability in presentation formats, asking learners questions to increase their depth of processing, or provoking group discussions are methods that may help to reach this goal.

To conclude this paper, we think there is accumulating evidence that extraneous and germane cognitive load are useful concepts to take into account for instructional design. Redirecting the learners' attention, by minimising extraneous and simultaneously increasing germane cognitive load, is a very promising method to improve training efficiency and provides good opportunities to reach higher transfer performance. However, there are many instructional methods to minimise extraneous load and many methods to increase germane load, yielding a large matrix of combined methods for redirecting attention. This article made clear that not all combined methods are equally effective. The use of completion problems under high contextual interference, as made operational in this study, had a positive effect on training efficiency but a disappointing effect on transfer test performance. An important challenge for the future is to identify the most effective combined methods.

Acknowledgements

The authors thank the Dutch Organisation for Scientific Research (NWO) for funding part of the research reported in this article.

References

Bratfisch, O., Borg, G., & Dornic, S. (1972). *Perceived item-difficulty in three tests of intellectual performance capacity (Tech. Rep. No. 29)*. Stockholm, Sweden: Institute of Applied Psychology.

Christensen, B.R. (1982). *Beginning comal*. Chichester: Ellis Horwood.

de Croock, M.B.M. (1999). *The transfer paradox: training design for troubleshooting skills*. Unpublished PhD thesis. Enschede: University of Twente, The Netherlands.

de Croock, M. B. M., van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). High vs. low contextual interference in simulation-based training of troubleshooting skills: effects on transfer performance and invested mental effort. *Computers in Human Behaviour, 14*, 249–267.

van Gerven, P. W. M., Paas, F. 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, 12*, 87–105.

Jelsma, O., & Bijlstra, J.P. (1988). Training for transfer in learning to detect, diagnose, and compensate system failures. *Proceedings of the Seventh European Annual Conference on Human Decision Making and Manual Control* (pp. 256–262). Paris, France.

Magill, R.A., & Hall, K.G. (1991). A review of the contextual interference effect in motor skill acquisition. In R.B. Wilberg (Ed.), *The learning, memory, and perception of perceptual-motor skills* (pp. 35–83). Amsterdam, The Netherlands: Elsevier.

Marcus, N., Cooper, M., & Sweller, J. (1996). Understanding instructions. *Journal of Educational Psychology, 88*, 49–63.

van Merriënboer, J. J. G. (1990). Strategies for programming instruction in high school: program completion vs. program generation. *Journal of Educational Computing Research, 6*, 265–287.

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., & Krammer, H. P. M. (1987). Instructional strategies and tactics for the design of introductory computer programming courses in high school. *Instructional Science, 16*, 251–285.

van Merriënboer, J.J.G., & Krammer, H.P.M. (1990). The "completion strategy" in programming instruction: theoretical and empirical support. In S. Dijkstra, B.H.M. van Hout-Wolters, & P.C. van der Sijde (Eds.), *Research on instruction* (pp. 45–61). Englewood Cliffs, NJ: Educational Technology Publications.

van Merriënboer, J.J.G., & Luursema, J.J. (1995). Implementing instructional models in computer-based learning environments: a case study in problem selection. In T.T. Liao (Ed.), *Advanced educational technology: research issues and future potential* (pp. 184–206). Berlin: Springer.

van Merriënboer, J.J.G., Krammer, H.P.M., & Maaswinkel, R.M. (1994). Automating the planning and construction of programming assignments for teaching introductory computer programming. In R.D. Tennyson (Ed.), *Automating instructional design, development, and delivery* (pp. 61–77). Berlin: Springer.

van Merriënboer, J.J.G., Luursema, J.J., Kingma, H., Houweling, F., & de Vries, A.P. (1995). Fuzzy logic instructional models: the dynamic construction of programming assignments in CASCO. In R.D. Tennyson, & A.E.B. Barron (Eds.), *Automating instructional design: computer-based development and delivery tools* (pp. 265–302). Berlin: Springer.

Paas, F. 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, F. 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, F. 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, F. 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.

Pennington, N., Nicolich, R., & Rahm, J. (1995). Transfer of training between cognitive subskills: is knowledge use specific? *Journal of Educational Psychology, 28*, 175–224.

Quilici, J. L., & Mayer, R. E. (1996). Role of examples in how students learn to categorize statistics word problems. *Journal of Educational Psychology, 88*, 144–161.

Schuurman, J.G. (1999). *On the proper treatment of learning and transfer: a study on introductory computer programming*. Unpublished PhD thesis. Enschede: University of Twente, The Netherlands.

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

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