



PERGAMON

Computers in Human Behavior 17 (2001) 373–391

www.elsevier.com/locate/comphumbeh

Computers in
Human Behavior

Just-in-time information presentation and the acquisition of complex cognitive skills

Liesbeth Kester ^{a,*}, Paul A. Kirschner ^a, Jeroen J.G. van Merriënboer ^a, Anita Baumer ^{b,1}

^a*Educational Technology Expertise Center, Open University of the Netherlands, PO Box 2960, 6401 DL Heerlen, The Netherlands*

^b*Department of Instructional Technology, University of Twente, The Netherlands*

Abstract

This paper describes a model for just-in-time (JIT) presentation of information. Learners receive the information needed to carry out a task precisely at the time it is needed. The model is twofold: *supportive* information is best presented before practising task clusters while *pre-requisite* information is best presented during practice on learning tasks. JIT supportive information presentation promotes schema construction through meaningful learning or elaboration. JIT prerequisite information presentation promotes schema automation through proceduralization or restricted encoding while learning the recurrent aspects of a task. It leads to a reduction in extraneous cognitive load because temporal split attention is avoided. This frees up cognitive capacity for learning the non-recurrent aspects of a complex cognitive skill and so enhances transfer performance. An exploratory, empirical study is described that supports our model. Implications for teaching and teacher training are discussed. © 2001 Published by Elsevier Science Ltd. All rights reserved.

In the late 1970s and early 1980s a new concept in inventory management revolutionized Japanese and American manufacturing companies. The traditional just-in-case (JIC) inventory systems based on long production runs, stockpiled inventories and uninterrupted production needed to be replaced by more flexible systems in order to meet new competitive and economic challenges. A just-in-time (JIT) inventory system was introduced to replace the traditional JIC inventory systems (Hoyt, 1996). JIT inventory management is a result of demand-pull production. This concept “involves the use of demand for a given product to signal when production should occur. The use of demand-pull allows a company to produce only

* Corresponding author.

¹ Anita Baumer no longer works at the University of Twente.

what is required in the appropriate quantity and at the right time” (Cheng & Podolsky, 1993, pp. 9–10). In this way, the application of demand-pull can prevent unnecessary inventory, which would otherwise stockpile following the traditional JIC approach to manufacturing.

Hoyt (1996) argues that a lot of similarities can be found between production systems and educational systems. He draws perfect matching parallels between the traditional JIC inventory systems and business education. For example, business education is characterized by a slow response to customer needs, just as JIC inventory systems. There is often a time lag between the skill or knowledge a business requires and the education provided. Regularly, graduate students find that newly learned skills do not meet the demands of the workplace. Hoyt (1996) proposes to apply JIT inventory system principles to business education in order to solve this problem. In a JIT educational system, specific business courses should be provided just before students need the skills at work. Applying the earlier definition of demand-pull to educational systems, the demand for given knowledge and skills is used to signal when training should occur. The use of demand-pull allows a student to learn what is required in an appropriate course and at the right time. JIT education principles are regularly applied in on-the-job training settings. On the Internet, numerous JIT training packages can be found. Unfortunately, little can be said about the effectiveness of those JIT courses because of a lack of appropriate research.

In this article, JIT inventory management principles, the demand-pull principle specifically, are applied to the instructional design of learning tasks that aim at mastering a complex cognitive skill. Within instructional design the demand for given information is used to signal when presentation of that information should occur. The use of demand-pull in instructional design allows a student to have required information available in an appropriate learning task, at the right time. In our Four-Component Instructional Design model (4C/ID model; Van Merriënboer, 1997), guidelines are given for JIT information presentation in training programs for complex cognitive skills. In this article these guidelines are outlined in more detail in our JIT information presentation model.

In the following sections the JIT information presentation model is described. Demand pull implications are different for the supportive and the prerequisite information a task requires. Therefore, the difference between supportive and prerequisite information is explained first. Second, the consequences of this difference with regard to JIT presentation is discussed. Third, an exploratory study is presented which supports the assumptions made by the JIT information presentation model. Finally, implications for future research, teaching and teacher training are discussed.

1. JIT information presentation model

The JIT information presentation model presented here is a specification of the 4C/ID model. Our specification concerns the instructional design of information related to whole-task practice and is twofold: *supportive* information is best presented *before* practicing task clusters (i.e. categories of learning tasks) and *pre-*

requisite information is best presented *during* practice on learning tasks. In Fig. 1, a schematized overview is given of JIC information presentation (Fig. 1A), JIT supportive information presentation (Fig. 1B), and complete JIT information presentation (Fig. 1C).

1.1. Prerequisite vs. supportive knowledge

A complex cognitive skill consists of a number of highly interrelated constituent skills that show qualitative differences in performance; at least some of the constituent skills require conscious processing and they all exhibit goal-directed behaviors. With regard to desired exit behavior, a difference can be made between

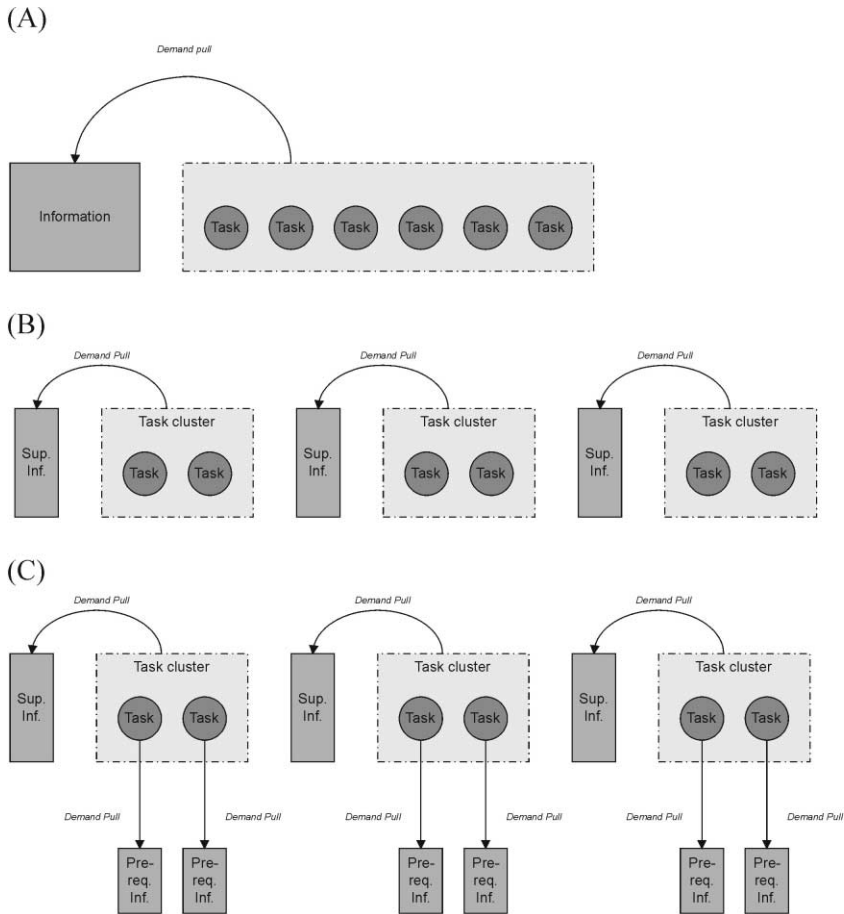


Fig. 1. Three models for information presentation: traditional JIC information presentation (A), JIT supportive information presentation in an elaborative sequence (B), and full, JIT supportive and prerequisite information presentation (C).

non-recurrent and recurrent constituent skills (Fig. 2). The application of non-recurrent constituent skills varies from task to task while recurrent constituent skills can be applied in the same manner in different tasks. For example, for the complex cognitive skill “troubleshooting in a alcohol–water distillery”, skills related to reasoning about the alcohol–water distillery using the principles underlying this functioning are considered non-recurrent skills. Recurrent skills, for instance, concern the procedures for operating the system and safety procedures. Mastering non-recurrent skills requires the construction of complex cognitive schemata that may guide subsequent problem solving behavior. Schema construction is a process of mindful abstractions from concrete experiences. Mastering recurrent skills requires the automation of schemata that is, highly domain-specific rules or procedures. Schema automation is mainly a function of the amount and quality of practice.

In order to design high quality instruction, non-recurrent and recurrent aspects of the complex cognitive skill have to be identified and analyzed. Moreover, an analysis

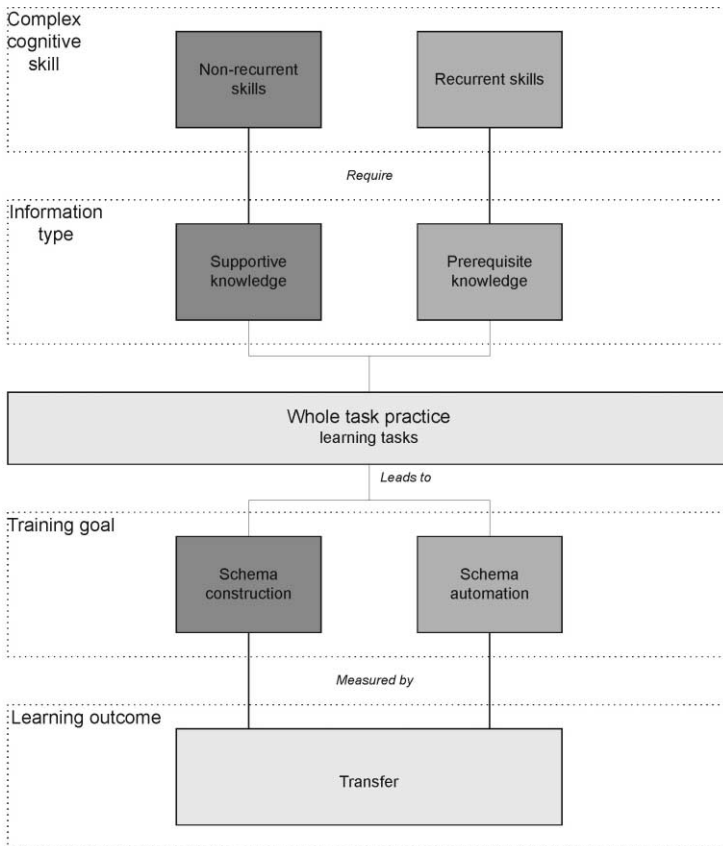


Fig. 2. The distinction between supportive and prerequisite information in the JIT information presentation model.

has to be made regarding the information that is *supportive* to the performance of the non-recurrent aspects of the complex skill (e.g. mental, causal and conceptual models) and information that is *prerequisite* to carry out the recurrent aspects of the skill (e.g. facts, concepts, etc.). In our example, information about the working of a valve or about the consequences of temperature fluctuations for the system is considered supportive while information about the names of the different components of the system or about specific operations like switching the alarm off is considered prerequisite. Once the constituent skills and the appropriate knowledge have been identified, the skills are practiced in learning tasks that promote schema construction for non-recurrent aspects, and schema automation for recurrent aspects. Learning tasks typically provide whole-task practice, that means, most or all of the constituent skills are trained simultaneously. Whole-task practice can be seen as a sequence of simple-to-complex task clusters in which concrete learning tasks include all aspects of the complex cognitive skill. The twofold purpose of whole-task practice is the construction of schemata that allow learners to perform unfamiliar task aspects (schema-based behavior) and the automation of schemata that allow learners to effortlessly perform familiar task aspects (rule-based behavior, Fig. 2).

In order to make assumptions about the effectiveness of JIT supportive and prerequisite information presentation, a measurement of desired outcome behavior has to be made. For the mix of schema-based and rule-based behaviors, effectiveness can be measured by the ability to solve transfer tasks. Two mechanisms to explain transfer are distinguished. The first mechanism is schema-based transfer in which the *same* declarative knowledge is used in a *different* way in solving transfer tasks. The second mechanism is rule-based transfer in which the *same* schemata, automated during practice, are used in the *same* way in solving transfer tasks. To elaborate on this, a continuum exists from near transfer tasks at the one end to far transfer tasks at the other end. Near transfer tasks are similar to the trained tasks while far transfer tasks are more and more different from the trained tasks. Both mechanisms are used simultaneously in solving transfer tasks. At the near transfer end of the continuum performance relies more on rule-based transfer than on schema-based transfer and at the far transfer end it is the other way round.

2. Just-in-time supportive information presentation

In our model supportive information is sequentially coupled to a range of task clusters (Fig. 1B). First, the training program for the complex cognitive skill is divided in task clusters in which simple-to-complex categories of learning tasks are identified. Second, for each task cluster concrete learning tasks are formulated. Before each task cluster, appropriate supportive information is given in order to promote meaningful learning or elaboration, that is, learning processes in which the learner connects new information to already existing, relevant cognitive structures. This elaborated supportive information is helpful in performing the non-recurrent aspects of the learning tasks in this particular cluster. The new information becomes embedded in these structures and because of this anchoring the new information no

longer depends on working memory for its existence (Ausubel, 1963). The use of *simple-to-complex sequences* for supportive information and task clusters enhances the elaboration process which leads to better schema construction. The idea that the use of simple-to-complex sequences of task clusters and related information has enhancing effects on learning stems from theories about prose learning which originated in the 1960s. Three of the most relevant theories in this respect are, the Subsumption Theory (Ausubel, 1963), the Assimilation Encoding Theory (Mayer, 1979) and the Elaboration Theory (Reigeluth & Stein, 1983) which are discussed here.

According to Ausubel's Subsumption Theory (1963), our cognitive architecture is hierarchically organized in terms of representations of past experiences. These representations are arranged from greater to lesser inclusiveness; every higher step in the hierarchy is linked to the former through a process of subsumption. Subsumption means the hierarchical incorporation of new material into an existing cognitive structure. Subsumption Theory thus assumes that meaningful learning only occurs when new material can be appropriately subsumed under a relevant, existing concept. The new material is "anchored" and becomes more stable by subsuming it under existing cognitive structures (Ausubel, 1963). In order to facilitate the subsumption process, advance organizers can be applied in instructional design. These organizers are introduced in advance of the learning material and distinguish themselves from summaries and overviews by a higher level of abstraction, generality, and inclusiveness. Moreover, because the content of a given organizer or series of organizers (cf. the proposed elaborative sequence of supportive information and task clusters) is selected on the basis of their suitability for explaining, integrating, and interrelating the material they precede, the organizational strength of the cognitive structure is enhanced. These organizers provide advance ideational scaffolding, that is, when the learner is confronted with unfamiliar material a cognitive structure is available for incorporating the new concepts (Ausubel, 1963; Jonassen, 1982).

However, studies concerning the effectiveness of advance organizers show inconsistent results (Jonassen, 1982). Mayer (1979) blamed this on the basic assumptions of the Subsumption Theory itself and therefore proposed a different theory, the Assimilation Encoding Theory (AET). The latter assumes that human cognitive structure is heterarchically organized and not hierarchically as Ausubel posits. AET distinguishes three stages of encoding: reception (the receipt of information in working memory), availability (the accessibility of anchoring knowledge in long-term memory), and activation (the appropriate transfer of anchoring knowledge from long-term memory to working memory so that it can be actively integrated with the received information and transferred back again to long-term memory). Meaningful learning will take place only when these three stages are passed through. Advance organizers only contribute to the learning process when they help the learner to complete all three stages. They have no effect when anchoring knowledge is already available in and appropriately activated by either the learner or the material presented. Moreover, when the material is badly structured and consists of isolated facts, an advance organizer cannot even be constructed. Mayer found in a number of studies that the predictions made by the AET regarding advance organi-

zers were fairly supported. In a meta-analysis of 132 studies advance organizers appeared to facilitate learning and retrieval (Jonassen, 1982; Mayer, 1979).

Elaboration Theory (ET; Reigeluth & Stein, 1983) provides us with another example of simple-to-complex sequences in instructional design, the elaborative sequence. Such a sequence starts with an epitome which entails the most inclusive and most general principles of a learning task, and gradually progresses to less inclusive, more detailed and more precise principles. The ET is explained by an analogy with a zoom lens that operates in steps. A person starts with a wide-angle view of a picture and sees the major parts and the relationships between those parts, but no detail. By zooming in the person sees more about a certain subpart (in our model, a task cluster and related supportive information) and learns more about its interrelationships. After viewing all the relevant features of a certain subpart the person can zoom out to the wide-angle view again to review the other subparts (i.e. the next task cluster and related supportive information) of the whole picture and the context of the subpart in the whole picture. This process of zooming in and zooming out can continue until all subparts are viewed in detail. Similarly, the ET prescribes that the epitome needs to be reviewed, thereby showing relationships between the most recent ideas and the ideas presented earlier. This pattern of elaboration followed by summary and synthesis is continued until the desired level of complexity is reached for all the aspects of the epitome (Reigeluth & Stein, 1983).

The three theories discussed above all expect beneficial effects from sequencing instructional tasks and related information in a simple-to-complex order. However, the effectiveness of sequencing depends on two major factors: the strength of the relationship between task clusters and the amount of time involved in mastering the whole complex cognitive skill. Sequencing becomes more important when the relationship between the constituent skills is stronger and when mastering a complex cognitive skill takes more than a couple of hours. When a cognitive skill is mastered in less time the learner can compensate for weaknesses in sequence (Reigeluth, 1999). These two basic rules should be kept in mind while designing instruction according to the JIT information presentation model.

2.1. JIT prerequisite information presentation

While supportive information is best presented before a new task cluster, prerequisite information is best presented during practice on the learning tasks (Fig. 1C). By JIT prerequisite information presentation, the proceduralization process associated with this information is facilitated because the proper information is active in working memory when the skill is practiced. Because the prerequisite information is directly available during practice, it may be embedded in the automated schemata or rules which constitute the recurrent skill. Moreover, by establishing schema automation for recurrent aspects of whole-task practice, more working memory capacity becomes available which can be allocated to learning the non-recurrent aspects of the complex cognitive skill. Thus, JIT information presentation aimed at schema automation is also beneficial to schema construction. The

temporal aspects of information presentation in instructional design, especially those regarding prerequisite information, have consequences for cognitive load.

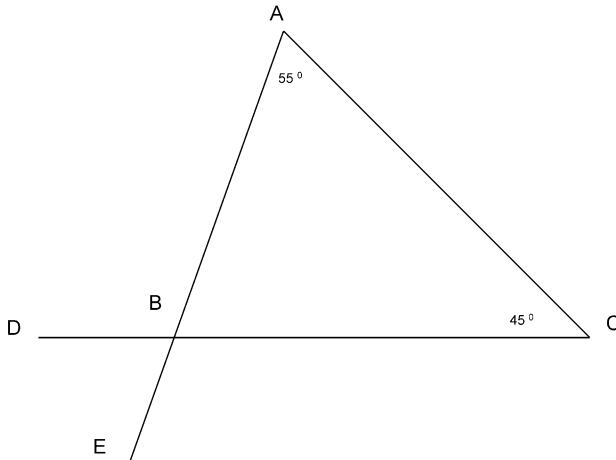
Cognitive Load Theory (Chandler & Sweller, 1991; Sweller, 1988; Sweller, Van Merriënboer & Paas, 1998) generates guidelines for the development of instructional material, starting from the assumption that working memory is severely limited. CLT distinguishes three forms of cognitive load that, together, determine the total cognitive load: intrinsic cognitive load, extraneous cognitive load and germane cognitive load. Intrinsic cognitive load is determined by the task being learned and the expertise of the learner (e.g. it is high for tasks with high element interactivity and/or learners with low expertise). Extraneous cognitive load is caused by the instructional material itself and refers to all the processes a learner engages in during a task which are not beneficial to learning (e.g. searching for relevant information sources, combining information sources). Germane cognitive load is a result of the processes directly relevant for learning and therefore appropriate within a learning context (e.g. conscious, mindful abstraction of generalized knowledge). With a given intrinsic cognitive load, the extraneous cognitive load should be minimized and the germane cognitive load should be optimized by instructional design (Cooper, 1998; Sweller et al., 1998). Of course, the total cognitive load should always remain within the boundaries of working memory capacity.

From the perspective of CLT, JIT prerequisite information presentation can be considered as a means to avoid *temporal split attention*. Split attention arises when multiple sources of information have to be mentally integrated in order to understand the complete picture. For instance, in geometry a diagram is often explained by accompanying text. In order to understand the diagram the text has to be mentally integrated with the diagram. This causes an increase in extraneous cognitive load because of visual search activities. In trying to understand different parts of the picture, the learner continuously has to search for the matching written information. This phenomenon is called the split attention effect (Chandler & Sweller, 1991; Sweller, 1994). Integrating the multiple conditional sources of information reduces extraneous cognitive load (Fig. 3). By integrating the multiple sources of information the need to search for relevant referents and mentally integrating them is diminished, resulting in a decrease of extraneous cognitive load and higher transfer test performance.

Avoidance of temporal split attention in instruction by JIT prerequisite information presentation should also reduce extraneous cognitive load. Analogous to integrating pictures and text, information separated in time is now integrated resulting in a lower extraneous cognitive load because the learners do not have to “search-and-match” and mentally integrate the information presented. Mayer carried out several experiments which provide a demonstration of the temporal split attention effect (Mayer & Anderson, 1991, 1992; Mayer & Sims, 1994).

In the first two experiments considered here, in an instructive animation on the working of a bicycle tire pump or an automobile braking system, integrated narration and animation was compared to successive presentation of narration and animation, narration only, animation only and no instruction at all. It was found that integrated instruction led to better performances on relevant creative problem

(1)



In the above Figure, find a value for Angle DBE.

Solution:

$$\begin{aligned} \text{Angle ABC} &= 180^{\circ} - \text{Angle BAC} - \text{Angle BCA (Internal angles of a triangle sum to } 180^{\circ}) \\ &= 180^{\circ} - 55^{\circ} - 45^{\circ} \\ &= 80^{\circ} \end{aligned}$$

$$\begin{aligned} \text{Angle DBE} &= \text{Angle ABC (vertically opposite angles are equal)} \\ &= 80^{\circ} \end{aligned}$$

(2)

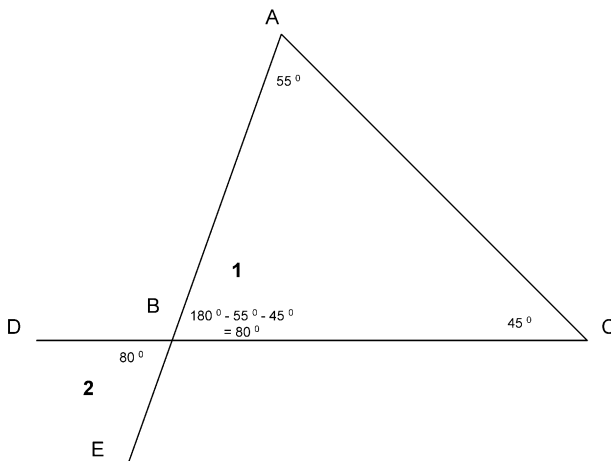


Fig. 3. Example demonstrating split attention (1). Integrated example with no split attention (2) (Sweller, Van Merriënboer, & Paas, 1998).

solving tests (Mayer & Anderson, 1991, 1992). Similar results were found in a third experiment where an instructive animation was given on the working of a bicycle tire pump or the human respiratory system. Integrated narration and animation was compared to successive presentation of narration and animation or no instruction at

all. It appeared that learners in the integrated condition performed better on the problem solving test than the learners in the other conditions (Mayer & Sims, 1994). According to these results, eliminating temporal split attention leads to more effective learning. From the perspective of CLT, the reduction of extraneous cognitive load by JIT prerequisite information presentation (i.e. avoidance of temporal split attention) facilitates the mastering of recurrent constituent skills. This, in turn, further decreases extraneous cognitive load in the acquisition of non-recurrent skills, thereby facilitating schema construction.

3. JIT information presentation: an exploratory study

The objective of this exploratory study is to investigate the relationship between our JIT information presentation model and learners' natural behavior. The learners were expected to get all the information they needed for performing a specific task cluster of a complex cognitive task, namely troubleshooting a water–alcohol distillery plant, by asking questions to the instructor. The study aimed at determining both *which* type of information they requested (either supportive or prerequisite) and *when* they requested it (either before task clusters or during learning tasks). The three main questions were the following. First, do learners ask prerequisite information mainly during practice of learning tasks? Second, do learners ask supportive information mainly before they start a new task cluster? And finally, do learners who show behavior that is consistent with our theoretical model, that is, mainly ask supportive knowledge before practicing task clusters and prerequisite information during practice of learning tasks, achieve higher transfer test performance than other learners? The following study was designed in order to answer those questions.

4. Method

4.1. Participants

Eight engineering science students (seven male, one female) from the University of Twente in the Netherlands participated in this experiment. They were selected at random from the population of second year students in this discipline and participated voluntarily. They were paid 30 Dutch guilders (approximately 15 USD) for their participation.

4.2. Materials

4.2.1. Distiller

The computer-based simulation program Distiller (De Croock, 1999) was used to practice and test troubleshooting skills. Distiller is a dynamic simulation of a water-alcohol distillery plant. Fig. 4 shows the Overview screen of Distiller, displaying a schematic representation of the distillery system.

The continuous production process is automated by means of six Proportional Integrative Differential (PID) controllers that measure process values, that is, flow or temperature of liquid mixtures running through pipes or the level of mixtures in columns in the system. On the basis of these measurements each PID controller opens or closes associated valves in such a way that the process value best approaches a pre-set optimal value, the so-called set point. Under normal conditions, the simulated system is stable and produces a liquid mixture close to 85% alcohol from a liquid mixture of approximately 40% alcohol.

To practice and test troubleshooting skills, learning tasks are defined in which at a specified time after the start of the simulation (ranging from 10 to 30 s) one component of the system starts to malfunction. As a result of this, the production process gets disturbed. Malfunctions manifest themselves in deviations of process values from their set points. In a controller, large deviations are indicated by means of a visual and acoustic alarm. After detection of an out-of-bounds situation in a particular controller, the behavior of that controller can be examined in more detail by selecting an associated Controller Information Display (CID). In a CID, the process value and the valve position over the last 10 min as well as the set point and alarm limits for that particular controller are displayed. By manipulating the controls of the system, that is, by switching the controller mode from AUTO to MANUAL and vice versa, by manually changing the valve positions, and by studying the system's behavior, the type of malfunction can be inferred. Once a diagnosis has been made, it can be reported to an imaginary dispatch crew. After about 12 s the dispatch crew reports back whether or not the diagnosis was correct and the faulty component repaired.

The following four system failures were used in this experiment: (1) a valve malfunction in which the valve cannot be controlled anymore and is stuck in a certain position (“valve no-energy”); (2) a failure in the PID controller (“PID controller malfunction”); (3) a leakage, in which fluid leaks out of a pipe (“leakage”) and (4) a sensor malfunction in which a sensor provides false information to the PID controller, on the basis of which the valve will be steered incorrectly (“PID sensor-malfunction”).



Fig. 4. Overview screen of Distiller.

4.2.2. Paper-based instruction

The participants received a short, one-page instruction in which they were told, in general terms, what their task would be: to learn to troubleshoot in Distiller. The introduction also briefly described the procedure of the experiment. They should simply start with the first learning task and ask the instructor anything they would like to know. The instruction included some examples of questions that could be asked to the instructor (see also the section “available information”).

4.2.3. Learning tasks

Learning tasks were developed to practice the following failures: (1) valve malfunction, (2) PID controller malfunction, (3) sensor malfunction, and (4) leakage. Twenty-two tasks were available for practice: six for the first hour of the experiment and 16 for the second hour. The order in which the learning tasks were presented was the same for all learners. A task was completed as soon as the correct diagnosis was made, the malfunctioning component was repaired, and the production process was stabilized.

4.2.4. Test tasks

A test was developed which consisted of five learning tasks. The same malfunctions as for practice were used, but the malfunctions occurred in other, new components. To assess test performance the following measures were computed: (1) the total time necessary for completing the test, (2) the mean time before a correct diagnosis was made, (3) the number of learning tasks solved, (4) the number of learning tasks solved at first attempt (i.e. no incorrect diagnoses are made), and (5) the mean number of wrong diagnoses.

4.2.5. Available information

Information necessary to perform the task was available for the learners. However, it was not provided in any tangible form. Participants could ask the instructor anything they wanted. In order to make sure that all participants received the same information in response to their questions, all available information was written down on cards that were used by the instructor in answering learners' questions. If the instructor was not able to relate a question to a particular card, the learner was told that the required information was not available. Three kinds of cards were distinguished:

1. Supportive information. These cards described information concerning (a) the functioning of the system and (b) the principles underlying this functioning. Example questions concerning the functioning of the system are “How does a valve work” and “What is the function of the reboiler?” With regard to principles, a further distinction is made between physical principles and system principles². Example questions are “What is happening to the boiling tem-

² Questions concerning *failure principles*, which link a particular symptom to a particular cause, were not acceptable (e.g. “Does this process value indicate that this valve is stuck?”). This could easily lead to trial-and-error behavior and is not very useful in learning to understand the system.

perature if the amount of alcohol increases in a water-alcohol mixture?” (i.e. a physical principle not directly related to Distiller) and “What will happen to this PID controller if the feed temperature decreases?” (i.e. a system principle directly related to Distiller).

2. Prerequisite information. These cards described information concerning (a) the topology of the system (names and locations of components) and (b) procedures for operating the system. Example questions concerning the topology of the system are: “What is the name of this component?” and “Which parts are connected by this feed pipe?” Example questions concerning the procedures necessary for operating the system are “How do I switch off the alarm?” and “How do I change the controller mode from auto to manual?”.
3. Other information. These cards described information that was not directly relevant to the troubleshooting task or Distiller. They were mainly related to the experimental procedure. Example questions are “How much time do I have to solve a task?” and “How long will the test take?”.

4.2.6. Observation list

A simple observation list was developed for scoring the number of questions, the types of questions, and the time at which they were asked. All questions were classified as either supportive (function and principles), prerequisite (topology and procedures), or “other” information.

4.3. Design and procedure

An instructor and an observer ran the experiment, the first providing all requested information to the learner and the observer instructing the learners beforehand and observing the sessions by scoring all questions asked on the observation lists. The sessions were videotaped. A second observer scored the tapes in order to be able to compute the inter-observer reliability. The consecutive stages of the sessions included a brief introduction (welcome and reading the paper-based instruction), a first block of practice of 60 min during which six learning tasks were presented, a 15-min break, a second block of practice of 60 min during which 16 learning tasks were presented, again a 15-min break, and the test consisting of five new learning tasks. The whole session took about three and a half hours.

During the introduction, participants were invited to ask everything they wanted to know to the instructor, either before starting practice or during practice. Subsequently, the troubleshooting task was explained by handing out the paper-based instruction, including examples of the kind of questions they could ask. All information necessary to learn the task was available by means of the answers given by the instructor, but no information was explicitly presented. At the start of the practice session, the computer screen showed the overview of Distiller. As long as no practice case was started, learners could try out everything that could be manipulated (e.g. changing the controller mode from automatic to manual). However, the system was not running yet which means that the imaginary plant remained static. For instance, if learners closed the valve in a pipe it had no impact on the

system's behavior except for the valve-position. Once a learning task was started, a failure was introduced and the system responded to all actions. For the test, learners had 6 min to solve a particular task. If the task was not solved within this time limit, it was scored as not diagnosed.

5. Results

5.1. Questioning behavior

Table 1 provides an overview of the numbers of questions asked for prerequisite and supportive information. Learners asked a mean of 36.5 questions (S.D. = 9.4), with 15.6 questions asked before practice (S.D. = 10.1) and 20.9 questions asked during practice (S.D. = 8.7). As predicted, questions about prerequisite information were asked significantly more often during practice ($M = 11.4$, S.D. = 4.1) than before practice ($M = 3.6$, S.D. = 2.2; Wilcoxon signed rank = -2.24 , $P < 0.025$), meaning that 76% of all questions about prerequisite information were asked *during* practice.

The reverse pattern for question asking was expected for supportive information. Indeed, questions about supportive information were asked slightly more often before practice ($m = 12.0$, S.D. = 8.5) than during practice ($m = 9.5$, S.D. = 6.2), indicating that 56% of all questions about supportive information were asked *before* practice. However, this difference is not significant.

For all other types of information, learners asked a mean of 13.6 questions (S.D. = 7.3), with 1.1 questions asked before practice (S.D. = 1.1) and 12.5 questions asked during practice (S.D. = 6.3). Of all questions that were asked before practice, only three questions were repeated during practice: one learner repeated two questions and one learner repeated one question. Thus, six learners never repeated the same question which may be an effect of the short duration of this study. The inter-observer reliability for categorizing questions as “prerequisite”, “supportive”, and “other” was found to be high, kappa = 0.929.

The data clearly support our expectation that prerequisite information is mainly asked during practice, probably because it is most useful when it is presented at exactly the moment that it is needed. With regard to supportive information, there is only a slight but non-significant tendency to ask most of this information before

Table 1
Mean numbers of questions asked for prerequisite and supportive information

	Prerequisite information		Supportive information		Total	
	<i>M</i>	S.D.	<i>M</i>	S.D.	<i>M</i>	S.D.
Before practice	3.6	2.2	12.0	8.5	15.6	10.1
During practice	11.4	4.1	9.5	6.2	20.9	8.7
Total	15.0	2.6	21.5	10.1	36.5	9.4

practice. Large differences between learners occur. While some ask for most of the supportive information before they start to practice others show a reverse pattern.

5.2. Test data

Learners needed a mean time of 17.9 min. (S.D. = 5.7) to complete the test tasks and a mean of 3.6 min. (S.D. = 1.1) for reaching a correct diagnosis on each of the five tasks. The mean number of tasks actually solved was 3.6 (S.D. = 1.2), of which 3.1 tasks (S.D. = 0.8) were correctly solved at the first attempt. The mean number of wrong conclusions for each task was 0.6 (S.D. = 0.4). These data indicate that the test was of a moderate difficulty and differentiated quite well between learners.

In Table 2, medians and ranges (min–max) are presented for so-called Matchers and Mismatchers. Matchers are those Learners who both (1) ask more questions about prerequisite information during practice than before practice ($n=7$) and (2) ask more questions about supportive information before practice than during practice ($n=4$). Thus, Matchers ($n=3$) show questioning behavior that is assumed to be optimal on the basis of our theoretical model and Mismatchers ($n=5$) are all other Learners. The Matchers show higher test scores than the Mismatchers on all test variables. Using the Wilcoxon signed rank test, this difference turns out to be significant for the accuracy measure “number of tasks solved correctly on the first attempt”; the same tendency is observed for the “mean number of wrong conclusions”. In addition, the Matchers needed significantly less time to reach a correct diagnosis, less time to complete the test, and solved more learning tasks.

The superior test performance of the Matchers cannot be explained by time on task, because the time available for practice was fixed for all Participants, or by the fact that they asked more task-related questions than the Mismatchers did. On the contrary, the Matchers asked a mean of 29.7 questions (S.D. = 4.5), with 15.7 questions pertaining to prerequisite information (S.D. = 4.2) and 14.0 questions pertaining to supportive information (S.D. = 3.0); the Mismatchers asked a mean of 40.6 questions (S.D. = 9.4), with 14.6 questions pertaining to prerequisite information (S.D. = 1.7) and 26.0 questions pertaining to supportive information (S.D. = 10.3). Thus, the Mismatchers asked more supportive information, but nonetheless performed worse

Table 2
Test results for matchers and mismatchers

	Matchers ($n=3$)		Mismatchers ($n=5$)		<i>W</i>
	MD	Range	MD	Range	
Learning tasks solved at first attempt	4	4–4	3	2–3	15.0**
Mean No. of wrong conclusions	0.4	0.2–0.6	0.8	0.2–1.4	10.0
Mean time until diagnosis (min)	2.4	1.9–2.6	4.4	3.7–4.7	6.0**
Total time for test (min)	11.8	9.5–12.9	21.9	18.6–23.6	6.0**
Learning tasks solved	5	4–5	3	2–4	16.0*

* $P < 0.05$.

** $P < 0.025$.

on the test. From the background information that was available on the Learners, there is, however, no indication that the Matchers had more prior knowledge about the working of distillation plants or the principles underlying distillation.

6. Discussion

This article described a model for JIT information presentation and an exploratory study. In this study an attempt was made to answer the following questions: “Do learners ask prerequisite information mainly during practice of learning tasks?”, “Do learners ask supportive information mainly before they start a new task cluster?” and “Do learners who show behavior that is consistent with our theoretical model, that is, mainly ask supportive knowledge before practicing task clusters and prerequisite information during practice of learning tasks, achieve higher transfer test performance than other learners?”. It was found that the learners indeed mainly asked for prerequisite information during practice of learning tasks while the findings for supportive information were less clear. Only a small tendency to ask for supportive information before practicing task clusters was found. Regarding our last research question, learners who show questioning behavior that is assumed to be optimal on the basis of our model (i.e. mainly ask supportive knowledge before practicing task clusters and prerequisite information during practice of learning tasks), do better on transfer tasks than learners who exhibit other questioning behavior. The preliminary data from the exploratory study support the practical model.

However, on the basis of this exploratory study, it is not possible to conclude that higher transfer test performance is indeed the result of the timing of information presentation. Due to the weak design of the present study, there are other possible explanations for the different results obtained by Matchers and Mismatchers, for example, differences in individual abilities or styles. Nevertheless, the results of the exploratory study discussed here seem promising enough to direct future research. To control for the individual differences that hamper the interpretation of the results obtained in the exploratory study, the information presentation should be detached from the learner’s questioning behavior. In future research, a task analysis should be made regarding the information that is supportive or prerequisite to carry out the task. In the task analysis the task itself and the desired exit behavior should be used to distinguish which information is supportive and which is prerequisite while the expertise of the learner should be used to decide which of this information should be presented JIT. When a more firm basis for the JIT information presentation model is established, we can elaborate our model further by trying to form hypotheses about the manner in which the information presentation takes place. Until now our model is about the “What?”, supportive and prerequisite information, and the “When?”, before or during practicing task clusters. But another set of questions relates to the “How?”.

Four relevant approaches regarding the “How?” of information presentation can be distinguished: deductive vs. inductive approaches and expository vs. inquisitory

approaches. First, the deductive and inductive approaches are based on deductive and inductive reasoning, respectively. In deductive reasoning the learner uses given generalities to come to conclusions (instances) that can be proven correct. In contrast, in inductive reasoning, learning is viewed as inducing generalities from limited instances and testing them. As long as new information can be predicted by the generality, it is retained. Only when incompatible information is encountered by the learner are new generalities generated (Mayer, 1983). When a deductive information presentation approach is used, generalities are presented before instances (e.g. illustrative examples). Using the inductive approach to information presentation, the instances are presented prior to the generalities. According to Evans, Homme, and Glaser (1962) a generality-instance format is best applied when the concept can be understood in abstract form while an instance-generality format is best used when the concept is difficult or too abstract for the learner (Fleming & Levie, 1979).

Second, the expository approach is characterized by the explicit presentation of relationships between pieces of information, while in the inquisitory approach the learners are asked to produce these relationships themselves. The inquisitory approach allows the learner to connect the new information to what is already known and leaves ample room for elaboration, the opposite is true for the expository approach. The latter should only be used when the available instruction time is severely limited. These four approaches are combined to four information presentation strategies: a deductive-expository strategy, a deductive-inquisitory strategy, an inductive-expository strategy and an inductive-inquisitory strategy. The 4C/ID model proposes that prerequisite information presentation should always be presented according to a deductive-expository strategy. Furthermore, the strategy best used to present supportive information depends on availability of instructional time, relevant experience of the learner and required level of understanding. As a default an inductive-expository strategy is proposed. But when instructional time is severely limited, the learner has enough relevant experience and deep understanding is not necessary, the deductive-expository strategy is recommended. The inductive-inquisitory strategy is suggested when there is ample instructional time and the learners are inexperienced and a deep level of understanding is required.

The practical implications of our model and preliminary results are quite clear. In teaching and teacher training, little or no attention is paid to timing aspects of information presentation. When attention is paid to this area, it is too often a just-in-case approach compounded with information redundancy which dominates. On the one hand, this may lead to a huge amount of information presented to learners before they start to practice on their learning tasks, with negative effects on learners' motivation. On the other hand, it may lead to a sub-optimal availability of prerequisite information *during* task performance because it is included in the just-in-case information and forgotten by the learners by the time they need to use it. As a result, teachers have to answer many questions and (again) present the prerequisite information (redundancy) while learners work on their learning tasks. This may be especially detrimental in distance learning and other independent learning situations where there is no or little synchronous contact between learners and teachers. For this reason, the authors include the distinction between prerequisite and supportive

information in their teacher training programs at the Open University of the Netherlands, and advise teachers to directly connect prerequisite information to the (Web-based) learning tasks they provide to their learners.

To conclude this article, future research will be aimed at establishing a firm basis for the use of the demand pull principle in the instructional design of learning tasks that aim at mastering a complex cognitive skill. An attempt has to be made to find further support for the assumptions that supportive information is best presented before practicing task clusters and prerequisite information is best presented during task practice. When firm support is found for these assumptions, the manner in which the supportive or prerequisite information is presented can be further studied.

References

- Ausubel, D. P. (1963). *The psychology of meaningful verbal learning*. New York: Grune & Stratton.
- Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8, 293–332.
- Cheng, T. C. E. & Podolsky, S. (1993). In P. Jarvis (Ed.), *Just-in-time manufacturing. An introduction*. London: Chapman & Hall.
- Cooper, G. (1998). *Research into cognitive load theory and instructional design at UNSW* (on-line). Available at: <http://www.arts.unsw.edu.au/education/clt.html>.
- De Croock, M. B. M. (1999). *The transfer paradox. Training design in troubleshooting skills*. Doctoral dissertation, University of Twente, the Netherlands.
- Evans, J. L., Homme, L. E., & Glaser, R. (1962). The rule system for the construction of programmed verbal learning sequences. *The Journal of Educational Research*, 55(9), 513–518.
- Fleming, M., & Levie, W. H. (1979). *Instructional message design. Principles from the behavioral sciences* (2nd ed.). New Jersey: Englewood Cliffs.
- Hoyt, B. R. (1996). *The just-in-time approach to effectively use business software in college business courses* (Report No. IR 018 261). Association of small computer users in education (ASCUE; ERIC Document Reproduction Service No. ED 405 821).
- Jonassen, D. H. (1982). *The technology of text. Principles for structuring, designing, and displaying text*. New Jersey: Englewood Cliffs.
- Mayer, R. E. (1979). Can advance organizers influence meaningful learning? *Review of Educational Research*, 49, 371–383.
- Mayer, R. E. (1983). Inductive reasoning: thinking as hypothesis testing. In R. E. Mayer (Ed.), *Thinking, problem solving, cognition* (2nd ed.; pp. 81–113). New York: Freeman & Company.
- 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., & 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.
- Reigeluth, C. M. (1999). The elaboration theory: guidance for scope and sequence decisions. In C. M. Reigeluth (Ed.), *Instructional design theories and models. A new paradigm of instructional* (1st ed.; pp. 425–453). Mahwah, NJ: Lawrence Erlbaum Associates.
- Reigeluth, C. M., & Stein, F. S. (1983). The elaboration theory of instruction. In C. M. Reigeluth (Ed.), *Instructional-design theories and models: an overview of their current status* (1st ed.; pp. 335–381). New Jersey: Lawrence Erlbaum Associates.
- 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., Van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251–296.
- Van Merriënboer, J. J. G. (1997). *Training complex cognitive skills. A four-component instructional design model for technical training*. New Jersey: Englewood Cliffs.