Educational Psychology Review, Vol. 17, No. 2, June 2005 (© 2005)
DOI: 10.1007/s10648-005-3951-0

Cognitive Load Theory and Complex Learning: Recent Developments and Future Directions

Jeroen J. G. van Merriënboer[1,3] and John Sweller[2]

Traditionally, Cognitive Load Theory (CLT) has focused on instructional methods to decrease extraneous cognitive load so that available cognitive resources can be fully devoted to learning. This article strengthens the cognitive base of CLT by linking cognitive processes to the processes used by biological evolution. The article discusses recent developments in CLT related to the current view in instructional design that real-life tasks should be the driving force for complex learning. First, the complexity, or intrinsic cognitive load, of such tasks is often high so that new methods are needed to manage cognitive load. Second, complex learning is a lengthy process requiring learners' motivational states and levels of expertise development to be taken into account. Third, this perspective requires more advanced methods to measure expertise and cognitive load so that instruction can be flexibly adapted to individual learners' needs. Experimental studies are reviewed to illustrate these recent developments. Guidelines for future research are provided.

KEY WORDS: cognitive architecture; biological evolution; complex learning; cognitive load; instructional design; expertise; adaptive instruction.

Cognitive Load Theory (CLT) uses interactions between information structures and knowledge of human cognition to determine instructional design. The theory's initial development in the early 1980s provided instruction that differed from the prevailing orthodoxies of the time. An emphasis on instruction designed to reduce unnecessary or extraneous cognitive

[1]Open University of the Netherlands, Heerlen, The Netherlands.
[2]University of New South Wales, New South Wales, Australia.
[3]Correspondence should be addressed to Jeroen J. G. van Merriënboer, Open University of The Netherlands, Educational Technology Expertise Center, P.O. Box 2960, 6401 DL Heerlen, The Netherlands; e-mail: jeroen.vanmerrienboer@ou.nl.

load resulted in, for example, a recommendation to provide learners with many worked examples rather than problems to solve. By considering relations between working memory and long-term memory, it was possible to structure both worked examples and related instruction to further reduce cognitive load. A strict insistence that all recommendations be directly tested for effectiveness using controlled experimental designs also provided a point of departure from the prevailing orthodoxies which tended to recommend instruction based on process models alone without testing comparative instructional effectiveness.

The theory has developed substantially since the 1980s. Having established a variety of basic instructional designs, an increasing number of cognitive load theorists from around the world considered how those instructional designs interacted, first, with the characteristics of the information and tasks that learners were dealing with and, second, with the characteristics of the learners themselves. Those interactions have generated an entirely new set of instructional guidelines. In turn, the theory and its design guidelines were, for the first time, able to throw light on basic cognitive processes and their origins rather than merely using known cognitive processes to generate instructional designs.

The purpose of this article is to discuss and place into context some of the newer methods of instruction developed over the past few years. Specifically, we consider how CLT has dealt with variations in informational complexity and in learner knowledge levels.

BASIC COGNITIVE LOAD THEORY

In a 1998 article in *Educational Psychology Review*, Sweller, van Merriënboer, and Paas described CLT at that time, discussed the main effects predicted by the theory, and reviewed empirical studies providing support for those effects. CLT assumes a limited working memory that stores about seven elements but operates on just two to four elements. It is able to deal with information for no more than a few seconds with almost all information lost after about 20 s unless it is refreshed by rehearsal. The theory emphasizes that these working memory capacity and duration limitations only apply to novel information obtained through sensory memory. Working memory has no known limitations when dealing with information retrieved from long-term memory (Ericsson and Kintsch, 1995; Sweller, 2003, 2004). In effect, long-term memory alters the characteristics of working memory. Long-term memory holds cognitive schemata that vary in their degree of complexity and automation. Human expertise comes from knowledge stored in these schemata, *not* from an ability to engage in reasoning with many elements that have not been organized in long-term memory. Human working memory simply is not able to process many elements. Expertise develops as learners mindfully combine simple ideas into more complex ones. A chess expert, for example, combines simple ideas about the best positioning of individual pieces to develop complex schemata of how several chess pieces should be positioned concomitantly. These schemata organize and store knowledge, but also heavily reduce working memory load because even a highly complex schema can be dealt with as *one* element in working memory.

In this sense, schemata can act as a central executive, organizing information or knowledge that needs to be processed in working memory. It is under these circumstances that there are no limits to working memory. For instance, an expert chess player recognizes a particular mid-game position at a single glance. In contrast, when dealing with novel information for which no schema-based central executive is available, working memory has limitations. Thus, for a novice chess player, a particular mid-game position may be little more then an unstructured set of single chess pieces. If knowledge is completely unavailable to organize information, it must be organized randomly and the organization then tested for effectiveness. Working memory must inevitably be limited in capacity when dealing with novel, unorganized information because as the number of elements that needs to be organized increases linearly, the number of possible combinations increases exponentially. Random testing of the effectiveness of possible combinations based on many elements becomes effectively impossible due to a combinatorial explosion. This problem of exponential growth can only be accommodated by severely limiting the number of information units that can be processed simultaneously. That problem does not arise when dealing with information from long-term memory that is already organized (Sweller, 2003, 2004).

Constructed schemata may become automated if they are repeatedly applied. As is the case for schema construction, automation can free working memory capacity for other activities because an automated schema, acting as a central executive, directly steers behavior without the need to be processed in working memory. Because automation requires a great deal of practice, automated schemata only develop for those aspects of performance that are consistent across problem situations, such as routines for dealing with standard game positions in chess, for operating machines, and for using software applications. From an instructional design perspective, well-designed instruction should not only encourage schema construction but also schema automation for those aspects of a task that are consistent across problems (van Merriënboer *et al.*, 2002a, 2003).

Novel information must be processed in working memory in order to construct schemata in long-term memory. The ease with which information may be processed in working memory is a focus of CLT. Working memory load may be affected either by the intrinsic nature of the learning tasks themselves (intrinsic cognitive load) or by the manner in which the tasks are presented (extraneous cognitive load). According to the 1998 version of CLT, *intrinsic cognitive load* cannot be altered by instructional interventions because it is determined by the interaction between the nature of the materials being learned and the expertise of the learner. It depends on the number of elements that must be processed simultaneously in working memory, and the number of elements that must be processed simultaneously, in turn, depends on the extent of *element interactivity* of the materials or tasks that must be learned. Materials with high element interactivity are difficult to understand—and the only way to foster understanding is to develop cognitive schemata that incorporate the interacting elements. It follows that a large number of interacting elements for one person might be a single element for another more experienced person who has a schema that incorporates the elements. Thus, element interactivity can be determined only by counting the number of interacting elements that people deal with at a *particular* level of expertise.

Extraneous cognitive load, in contrast, is load that is not necessary for learning (i.e., schema construction and automation) and that can be altered by instructional interventions. Extraneous cognitive load may be imposed, for example, by using weak problem solving methods such as working backward from a goal using means-ends-analysis, integrating information sources that are distributed in place or time, or searching for information that is needed to complete a learning task. Overloading one of the processors that constitute working memory also may increase it. Visual and auditory working memory are partially independent. If multiple sources of information that are required for understanding are all presented in visual form (e.g., a written text and a diagram), they are more likely to overload the visual processor than if the written material is presented in spoken form, thus enabling some of the cognitive load to be shifted to the auditory processor (Mousavi *et al.*, 1995).

Extraneous cognitive load and intrinsic cognitive load are additive. Whether extraneous cognitive load presents students with a problem depends, in part, on the intrinsic load: If intrinsic load is high, extraneous cognitive load must be lowered; if intrinsic load is low, a high extraneous cognitive load due to an inadequate instructional design may not be harmful because the total cognitive load is within working memory limits.

Thus, carefully considered instructional design is particularly important for teaching difficult subject matter. In 1998, CLT had been used almost exclusively to study instruction intended to decrease extraneous cognitive load. Some of the major effects that yield better schema construction and higher transfer performance and that may be attributed to a decrease in extraneous cognitive load are briefly summarized in Table I.

Table I. Some Effects Studied by Cognitive Load Theory and Why They Reduce Extraneous Cognitive Load[a]

Effect	Description	Extraneous load
Goal-free effect	Replace conventional problems with goal-free problems that provide learners with an a-specific goal	Reduces extraneous cognitive load caused by relating a current problem state to a goal state and attempting to reduce differences between them; focus learner's attention on problem states and available operators
Worked example effect	Replace conventional problems with worked examples that must be carefully studied	Reduces extraneous cognitive load caused by weak-method problem solving; focus learner's attention on problem states and useful solution steps
Completion problem effect	Replace conventional problems with completion problems, providing a partial solution that must be completed by the learners	Reduces extraneous cognitive load because giving part of the solution reduces the size of the problem space; focus attention on problem states and useful solution steps
Split attention effect	Replace multiple sources of information (frequently pictures and accompanying text) with a single, integrated source of information	Reduces extraneous cognitive load because there is no need to mentally integrate the information sources
Modality effect	Replace a written explanatory text and another source of visual information such as a diagram (unimodal) with a spoken explanatory text and a visual source of information (multimodal)	Reduces extraneous cognitive load because the multimodal presentation uses both the visual and auditory processor of working memory
Redundancy effect	Replace multiple sources of information that are self-contained (i.e., they can be understood on their own) with one source of information	Reduces extraneous cognitive load caused by unnecessarily processing redundant information

[a]Reported by Sweller *et al.*, 1998.

Over the last 5 years, CLT followed two distinct developmental routes. First, the theory and instructional findings were used to develop our knowledge of human cognitive architecture. As a result, we better understand why our cognitive architecture followed its particular evolutionary route. This work was intended to strengthen the cognitive foundations of CLT rather than to directly address instructional issues.

The second route continued the traditional task of CLT: namely, the generation of new instructional effects (see Kirschner, 2002; Paas *et al.*, 2003a, 2004). In contrast to most previous work within the CLT framework, this work examined instructional methods that affect intrinsic and so-called germane cognitive load rather than extraneous load.

Three major instructional developments are discussed in this article. First, an increasing number of instructional theories view rich, real-life learning tasks as a basis for complex learning (Merrill, 2002). The cognitive load imposed by such tasks is often excessive for novice learners and may seriously hamper learning. Conventional methods to decrease extraneous cognitive load (see Table I) might not lower the total load imposed by rich learning tasks to an acceptable level and thus not leave enough cognitive resources for schema construction and automation. Therefore, new instructional methods are beginning to be studied that affect intrinsic cognitive load and/or decrease extraneous cognitive load by providing problem-solving support. Second, recent studies place less emphasis on short-term laboratory experiments and focus more on lengthy training programs. Emphasizing lengthy training programs increases the importance of students' motivation. Effective instructional methods encourage learners to invest free processing resources to schema construction and automation, evoking germane cognitive load. In addition, lengthier training programs reveal the importance of the expertise reversal effect: instructional methods that work well for novice learners may have neutral or even negative effects when expertise increases. This effect necessitates the formulation of instructional strategies that make the application of particular instructional methods dependent on learners' expertise. The third development is closely related to the expertise reversal effect. It is difficult or even impossible for preplanned instruction to take the expertise of individual learners fully into account. The aim of this line of research is to develop new methods for the assessment of expertise based both on performance and cognitive load and to use assessment to develop adaptive eLearning applications. By continuously adapting instruction to levels of expertise, the difficult task of predicting subsequent levels of expertise prior to the commencement of an instructional sequence is obviated. Following a discussion of the origins of human cognitive architecture, the remainder of this article addresses these three instructional issues.

ORIGINS OF HUMAN COGNITIVE ARCHITECTURE

Evolution by natural selection is an information processing system that solves problems. The problem faced by a genetic information processing system is how to survive in an inevitably variable environment. If the currently "known" information is sufficient for survival, the system does not need to alter. If it is not sufficient for survival, it must alter satisfactorily or reach a dead-end. The system has three critical characteristics. First, it includes a massive store of information in the form of a genome that permits some individuals and species to survive in their variable environments. In other words, all problems faced by the genome to that point have been solved because the genome includes information relevant to the direct solution of those problems. Second, all changes to a genome can be traced back to random mutations characterized as a generate-and-test problem-solving exercise. Any mutation is tested for effectiveness with effective mutations retained and ineffective ones jettisoned. A successful mutation can solve new problems without losing the ability to solve critical old problems. Third, species' genomes alter slowly because a rapid, large, random alteration is unlikely to be adaptive. A large, rapid alteration in a genome is likely to leave it unable to solve problems that previously could be solved readily. A species' genome is unlikely to alter if the species can handle readily the variations in its environment.

In effect, a genetic system can be seen as a creative, information-processing system designed to decide which genetic modifications are usable and which are not. This system has driven the evolution of human cognitive architecture. In this section, it is argued that the information-processing properties of human cognition have evolved to mimic the information-processing properties of biological evolution (for a complete discussion, see Sweller, 2003, 2004). There may be only one natural information-processing system on earth and that system is found in both the procedures of biological evolution and in some of the results of biological evolution—sentient beings.

The analogy between biological evolution and human cognition is readily established. Just as an evolutionary system requires an enormous store of information to function, so does the human cognitive system. The huge store of schematically organized information in human long-term memory is central to all human cognitive activity. Virtually everything humans see, hear, or even think about is critically dependent on information stored in long-term memory. We suggest that information plays the same role in human cognition as genetic information plays in biological evolution. Both guide the actions of their respective systems and both act as a central executive determining what and when actions should be taken.

Furthermore, there is no other logically possible central executive. Any "hard-wired" central executive inevitably leads to an infinite regression of central executives (Sweller, 2003) known as the "Droste effect."[4]

As an example, consider the schema for the complex set of squiggles needed to read the word "cat." We have a schema for these squiggles because they can take an infinite variety of forms (e.g., when hand-written) that we process in exactly the same way. Furthermore, despite the complexity of the squiggles, they can be treated as one single element in working memory. The schema is automated because we no longer have to use working memory to process the geometric shapes to extract meaning from them. When we were learning to read, each of the shapes were processed individually and combined consciously to extract the appropriate meaning. This resulted in a heavy working memory load. Because the schema was eventually automated, that is no longer necessary, thereby reducing cognitive load. Furthermore, the schema for the squiggles "cat" includes its own executive functions that tell us what to do and how to do it. It tells us how to interpret the squiggles (how to "read"' them), how they connect to other similar squiggles, and what they refer to. The schema for "cat" tells us how to react to that particular animal. Until the schema for "cat" is automated, all this activity occurs in working memory. In other words, the executive functions associated with the word "cat" must be learned. No other executive functions are available or, indeed, needed.

Just as the information stored in a genome may not be sufficient to guide all needed activity, the information stored in long-term memory also proves insufficient under many circumstances. Mechanisms to alter the information store are needed. Again, the mechanisms used in evolution by natural selection and in human cognition are analogous. Evolution uses a variety of mechanisms to alter a genome but all genetic variation between individuals of a species and between species can be traced back to random mutation. Precisely the same mechanisms are used by the human cognitive system. Humans alter the contents of long-term memory by learning. There are two ways in which new information is acquired. We can obtain information directly from another human by instruction or we can generate new information by a process of problem solving. At some point, if knowledge is unavailable, all problem solving requires random generation of problem-solving moves followed by tests of effectiveness (Sweller, 2003). No other mechanism is available and this mechanism is identical to the one used in evolution by natural selection. Furthermore, even when knowledge is acquired by instruction from other humans, at some point that knowledge

[4]The Dutch chocolate maker Droste is famous for the visual effect on its boxes of cocoa. The image on the box contains itself on a smaller scale, which contains itself on a smaller scale, and so on.

was generated by a random generation followed by an effectiveness-testing mechanism. In other words, the initial generation of usable information by humans is closely analogous to the initial generation of information by biological evolution.

The last point of analogy concerns the speed with which large, natural stores of information are altered. A large store of information cannot be rapidly altered by random procedures without losing its functionality. Accordingly, successful genomes are stable and so, for example, humans share almost all of their genes with apes despite last having a common ancestor millions of years ago. Equivalently, human long-term memory does not alter rapidly. Human cognition has a specific structure to ensure that rapid alterations to long-term memory do not occur: A limited working memory. Working memory can be used to test the effectiveness of only a small number of combinations of elements. For example, using the logic of permutations, three elements can be combined in six different ways ($3! = 6$) whereas 10 elements can be combined in over 3.5 million different ways ($10! = 3,628,800$). A working memory that could deal with more than a few elements of information would not be functional. A working memory that was used to effect large, random alterations to long-term memory would be even less functional.

The treatment of human cognitive architecture as an example of a natural information-processing system is not intended to have direct instructional implications. It is intended to strengthen the cognitive architecture assumed by CLT and that architecture, in turn, is intended to have direct instructional implications. If human cognitive architecture includes a massive long-term memory holding uncountable schemas and if working memory must be limited to ensure the important information in long-term memory is not corrupted by random processes, then the aim of instruction should be to accumulate rapidly systematized, coherent knowledge in long-term memory. Aiding the accumulation of usable rather than random knowledge in long-term memory means that information need not be freely discovered by learners but rather be conveyed in a manner that reduces unnecessary working memory load. CLT is entirely concerned with these aims. The remainder of this article considers newly developed instructional methods that deal with changes from (1) simple to complex learning, (2) short experiments to lengthy training programs, and (3) preplanned instruction to adaptive eLearning.

FROM SIMPLE TO COMPLEX LEARNING

Complex learning aims at the integration of knowledge, skills, and attitudes; the coordination of qualitatively different constituent skills; and the transfer of what is learned to daily life or work settings (van Merriënboer *et al.*, 2003). Current instructional theories are expanding to incorporate authentic learning tasks that are based on real-life tasks as the driving force for such complex learning (for a discussion of theories, see Merrill, 2002). Whereas CLT has been able to generate effective instructional methods for tasks in relatively well-structured procedural and conceptual domains, its applications to the field of complex learning are just beginning. Because authentic learning tasks are often characterized by high element interactivity, there is a need to develop new instructional methods that allow the designer to optimize cognitive load for such highly complex tasks.

Knowledge Progression and Intrinsic Cognitive Load

The most important characteristic of complex learning is that students must learn to deal with materials incorporating an enormous number of interacting elements. In conceptual domains, there are many interacting knowledge structures that must be processed simultaneously in working memory in order to be understood. In skill domains, there are many interacting constituent skills that must be coordinated in working memory for a coherent performance. Earlier studies showed that instruction designed to decrease extraneous cognitive load has negligible effects on learning if element interactivity is low; however, such instruction positively affects learning and transfer performance for complex materials with a high level of element interactivity (Carlson *et al.*, 2003; Marcus *et al.*, 1996; Sweller and Chandler, 1994; Tindall-Ford *et al.*, 1997). The explanation is that for materials with low element interactivity, there is no need to decrease extraneous cognitive load because there are sufficient cognitive resources available for learning. For materials with high element interactivity, the decrease of extraneous cognitive load is necessary to free up processing resources that can be devoted subsequently to learning. The major question addressed in this section is what can be done if even after the removal of all sources of extraneous cognitive load, the element interactivity of the material is still too high to allow learning?

Earlier research indicated that the organization of instructional texts affects the allocation of cognitive capacity (e.g., Britton and Glynn, 1982). Pollock *et al.* (2002) were the first to study the effects of sequencing in the context of CLT and to test techniques for reducing intrinsic cognitive load. They presented learners with a complex sequence of instruction in two parts. In the first part, cognitive load was reduced by not presenting all information at once. Instead, isolated elements that could be processed serially were presented. In the second part, however, all information was presented at once, including the interactions among the elements. Thus, the elements had to be processed simultaneously in working memory. As hypothesized, although understanding was lower in the first phase of instruction when elements were presented in isolation, this deficiency was compensated in the second phase when the full set of interacting elements was presented. Thus, presenting the full set of interacting elements in both phases resulted in less understanding than presenting isolated elements in the first phase followed by the full set in the second phase. In general, the progressive method is an appropriate technique to use for novice learners who are confronted with highly complex materials but who lack the rudimentary schemata for dealing with those materials.

This study (Pollock *et al.*, 2002) shows that it is possible and sometimes beneficial to alter *intrinsic* cognitive load by artificially reducing element interactivity (see also Bannert, 2002). At first sight, reducing intrinsic cognitive load seems to contradict the original suggestion that intrinsic cognitive load cannot be altered by instruction (Sweller *et al.*, 1998). This suggestion is still valid in the sense that by artificially reducing intrinsic cognitive load, understanding is also reduced. For learners to fully understand the material they must ultimately be presented with the materials in their full complexity, with all element interactivity that is typical of the domain. Understanding complex information may not be necessary or even possible in the early stages of learning. The isolated-followed-by-interacting-elements approach suggests that when dealing with complex information, the intrinsic cognitive load of the material should be reduced by eliminating the interactions among the information elements. Those interactions impose a high intrinsic load because they require that elements be processed simultaneously. If the interactivity among the elements is reduced, then discrete information units are created that can be processed serially—substantially reducing working-memory load. Learners may need to be presented with the materials in their full complexity only in a later learning phase.

Although there is clear evidence that sequencing instruction from isolated elements to interacting elements is beneficial for learning, the level at which interacting elements should be broken down into isolated elements is still unclear. Information elements themselves are little more than sets of interactions. Pollock *et al.* (2002) did not give specific guidelines indicating the points at which interacting elements should be converted into isolated elements, but a study of their instructional materials suggests that their isolated information elements are really sets of interactions that are already familiar to the learners. Reigeluth (1999) and van Merriënboer (1997) suggested a slightly different sequencing technique that focuses, in the early phases of learning, on those elements that are most fundamental and representative for the whole complex task. Focusing on fundamental elements should allow the learners to obtain a quick impression of the whole task that can be further elaborated in the course of the training program.

Consequently, CLT does not contradict the common finding that whole-task sequencing is more suitable than part-task sequencing for tasks with a high level of integration and coordination (for reviews, see Goettl and Shute, 1996; Peck and Detweiler, 2000; van Merriënboer, 1997). Both sequencing techniques work from less to more interacting elements, but part-task sequencing practice does so by progressing from part-tasks to the whole task, whereas whole-task sequencing does so by progressing from simplified versions to more complex versions of the whole task. CLT merely indicates that if even the simplest version of the whole task is too demanding for beginning training so that part-task sequencing is inevitable to further decrease cognitive load to an acceptable level, then the part-tasks should be chosen in such a way that they stimulate the construction of schemata with a cognitive structure facilitating the learning of subsequent schemata. In this respect, a fast automation of the part-tasks may be detrimental to learning the whole task because this blocks successful integration later (Schwartz *et al.*, in press). Instead, already acquired schemata should act as a central executive and organize the new knowledge that needs to be processed in working memory—a view that is consistent with the idea of a long-term working memory (Ericsson and Kintsch, 1995). Future research should compare various sequencing techniques and empirically determine what effects they have on intrinsic cognitive load, the structure of acquired schemata, and the transfer of learning.

Problem-Solving Support

Rich learning tasks often require problem-solving and reasoning skills. The most common method used to help students with problem solving is to provide them with some kind of *process worksheet* that provides a description of the phases one should go through when solving a problem and that provides hints or rules-of-thumb that help students successfully complete each phase. Students can consult the process worksheet while they are working on the learning task(s), and they may use it to note down intermediate results of the problem solving process. It should be clear that both the phases and the hints have a heuristic nature: They may help solve the problem but do not guarantee a correct solution (this distinguishes them from a procedure, which is algorithmic in nature). Nadolski *et al.* (2001) used a process worksheet for law students who were trained to plea in court. It contained phases such as ordering the documents in the file, getting acquainted with the file, and analyzing the situation for preparing and conducting the plea. For each phase, it also contained rules-of-thumb. For the first phase, ordering the documents in the file, there were hints to order the files in legal categories (documents, letters, notes), in chronological order, and in terms of relevance for the case.

Although process worksheets are frequently used in educational practice, little is known about their effects on learning. Nadolski *et al.* (in press) studied the effects of process worksheets with law students and found that the availability of a process worksheet had positive effects on the coherence and content of prepared pleas. However, the availability of a process worksheet had no effects on cognitive load and transfer task performance (i.e., a new plea that had to be prepared without any available support). Thus, the availability of process worksheets does not decrease extraneous cognitive load, and, although it helps in performing the learning task, there is no indication it helps in constructing cognitive schemata that promote transfer of learning.

Kester *et al.* (2001; see also Kester *et al.*, 2004, 2005) explained why process worksheets were not effective. Because the information provided in a process worksheet typically has high element interactivity, simultaneously performing the learning tasks and consulting the worksheet may be too demanding. Working memory demands may be increased further because learners must split their attention between the task and the process worksheet. It may be better if learners thoroughly study the recommended phases and hints before they start to work on the learning tasks as suggested by the results of studies on the positive effects of advance organizers (for an overview, see Williams and Butterfield, 1992). By initially studying the recommended phases and hints of a worksheet, a cognitive schema may be constructed in long-term memory that can subsequently be activated in working memory during task performance. Retrieving the already constructed schema should be less cognitively demanding than activating the externally presented complex information in working memory during task performance.

A recent study on learning to troubleshoot electrical circuits provided some empirical support for this explanation (Kester *et al.*, in press). In this study, a distinction was made between *supportive information* with high element interactivity, such as a description of phases and hints that help solve a troubleshooting problem (cf., a process worksheet), and *procedural information* with low element interactivity, such as step-by-step instructions for manipulating an electrical circuit (e.g., toggling a switch or replacing a lamp). An interaction effect was found on cognitive load. Load was relatively high when both supportive and procedural information were presented *before* practice on the troubleshooting problems; load was relatively low when supportive information was presented before practice and when procedural information was presented during practice. Similarly, load was relatively high when both supportive and procedural information were presented *during* practice; load was relatively low when procedural information was presented before practice and when supportive information was presented during practice. In other words, presenting supportive and procedural information together increased cognitive load. The same interaction pattern was found for transfer test performance. As expected, the highest transfer test performance occurred when supportive information was presented before practice and when procedural information was presented during practice.

In conclusion, it seems that high element interactivity information such as a description of recommended problem-solving phases and hints is best studied by students before they start to practice, rather than during practice. However, future research may yield alternative instructional methods that are equally or more effective. For instance, students may first study one or more process-oriented worked examples, that is, examples that show how an expert solves a particular category of problems and articulate why the problem is solved this way. Such examples may provide students with the preknowledge that enables them to make more effective use of process worksheets during practice (van Gog *et al.*, 2004). As another example, performance constraints or "training wheels" may prove more effective than process worksheets. The basic idea is to make unavailable to learners the cognitive processes and actions that are not relevant in a particular phase of the problem-solving process. Those processes and actions only are made available after the learners successfully complete the previous phase(s) and begin to work on the phase for which the processes and actions are relevant. For instance, law students who are learning to prepare a plea for court would not be allowed to start reading documents (phase 2) before they have ordered all documents in the file in an acceptable fashion (phase 1), and they would not be allowed to perform a situational analysis (phase 4) before they have thoroughly studied the entire file (phase 3). In contrast to process worksheets, performance constraints reduce the student's problem space in a particular phase of the problem-solving process and perhaps decrease the extraneous cognitive load caused by considering irrelevant actions (Leutner, 2000; see also van Merriënboer, 2000).

FROM SHORT EXPERIMENTS TO LENGTHY TRAINING PROGRAMS

Recent applications of CLT in the field of complex learning focus less on short laboratory experiments and more on lengthy training programs.

Students may be more inclined to employ their processing resources when they participate in relatively short—and occasionally paid—laboratory experiments, where they may try to please the experimenter (acquiescence bias) or exhibit desirable behaviors (best light phenomenon), than when they participate in regular training programs. Students' reluctance to use processing resources while learning may make it necessary to encourage them to invest mental effort in schema construction and automation (processes that increase germane cognitive load—see below). Furthermore, the expertise reversal effect is of particular importance during longer training programs because it indicates that instructional methods have to change as learners acquire more expertise. Some examples of instructional strategies that take the expertise reversal effect into account are discussed.

Germane Cognitive Load

In their 1998 article, Sweller, van Merriënboer, and Paas introduced the concept of germane cognitive load. This new construct was necessary to explain the effects of variability in materials presented to learners. Variability of problem situations encourages learners to construct cognitive schemata because variability increases the probability that similar features can be identified and that relevant features can be distinguished from irrelevant ones. Instructional designs that incorporate high variability ensure that a task is practiced under conditions that require the performance of different variants of the task across problem situations or across conditions that increase variability along other task dimensions, such as the manner in which the task is presented, the saliency of defining characteristics, the context in which the task is performed, or the familiarity of the task. It is well documented that variability of practice may result in beneficial effects on schema construction and transfer of training (e.g., McKeough *et al.*, 1995). These findings were first related to CLT by Paas and van Merriënboer (1994a) and by Quilici and Mayer (1996).

The results of studies on variability initially seemed to contradict CLT. High variability *increased* cognitive load during practice but yielded better schema construction and transfer of learning as indicated by a better ability to solve problems that were not solved before. Paas and van Merriënboer (1994a) hypothesized that the increase in cognitive load was due to processes directly relevant to learning (schema construction and automation) instead of processes that resulted in extraneous cognitive load. Exposure to a highly varied sequence of problems and solutions to those problems helps learners determine the range of applicability of constructed schemata. Although determining the applicability of a schema is clearly useful, it requires the mindful engagement of the learners and increases cognitive load. This new element in CLT indicates that it may be necessary to increase learners' motivation and encourage them to invest germane cognitive load in learning. Consequently, instructional manipulations to improve learning by diminishing extraneous cognitive load and by freeing up cognitive resources is only effective if students are motivated and actually invest mental effort in learning processes that use the freed resources.

The germane load construct in CLT explains the results of studies on contextual interference and self-explanation. Contextual interference is a special kind of variability (van Merriënboer *et al.*, 1997). If adjacent problems rely on exactly the same skills, then contextual interference is low. If, on the other hand, adjacent problems rely on different skills, then interference is high and will stimulate learners to construct schemata. For instance, suppose that learners are trained in a troubleshooting task where four types of malfunctions (m1, m2, m3, m4) occur in four different components of a technical system (c1, c2, c3, c4). Low contextual interference would imply a sequence of problems where the skills for troubleshooting one particular type of malfunction are fully practiced before the skills for troubleshooting other types of malfunctions are introduced, yielding a "blocked" practice schedule like this:

m1c1, m1c2, m1c3, m1c4
m2c1, m2c2, m2c3, m2c4
m3c1, m3c2, m3c3, m3c4
m4c1, m4c2, m4c3, m4c4

In contrast, a high contextual interference condition would sequence the 16 problems in such a way that each requires a different solution using a "random" practice schedule. De Croock *et al.* (1998) and van Merriënboer *et al.* (2002b) reported a number of studies on contextual interference. As expected, high contextual interference (i.e., a random practice schedule) increased cognitive load during training and improved transfer performance. In one study, it was hypothesized that *redirecting attention* from extraneous to germane processes improves training efficiency, that is, positively affects the balance between cognitive load during training and transfer test performance. Redirecting attention was realized by replacing conventional problems (i.e., high extraneous load) sequenced in a blocked order (i.e., low germane load) with completion problems (i.e., low extraneous load) sequenced in a randomized order (i.e., high germane load). As expected, training efficiency was highest for the completion group practicing under high contextual interference.

A second line of research on germane cognitive load has examined the elaboration of worked examples and the self-explanation effect. Stark *et al.* (2002) conducted a study on learning from worked examples. Students exhibited three different elaboration strategies: passive, cognitive, and metacognitive. Students who used predominantly cognitive and metacognitive elaboration strategies invested more mental effort than students who used a passive strategy and also did best on the subsequent transfer tests, indicating that effective example elaboration is associated with a higher germane cognitive load.

A common instructional problem is that many students use a passive elaboration strategy. Renkl (1997) reported that many learners do not effectively use their available processing resources and do not spontaneously provide fruitful self-explanations when they study worked examples. To solve this problem, Stark *et al.* (2002) included a short training session focusing on aspects of cognitive and metacognitive elaboration. The researchers found that such a training session enhanced the quality of example elaborations and improved learning outcomes. Renkl and Atkinson (2001; see also Hummel and Nadolski, 2002) prompted students to self-explain. In their study, students worked on incomplete worked examples in the field of statistics (probability). In the prompted group, students were explicitly asked at each worked-out step which probability rule was applied. The researchers found higher transfer test performance in the prompted group.

The Expertise Reversal Effect

A newly described effect that is particularly relevant when CLT is applied to the design of courses with a longer duration became known as the expertise reversal effect (for an overview, see Kalyuga *et al.*, 2003). This effect is an interaction between several basic cognitive load effects (split-attention, modality, and worked example effects) and level of expertise. The effect is demonstrated when instructional methods that work well for novice learners have no effects or even adverse effects when learners acquire more expertise. For example, Kalyuga *et al.* (1998) demonstrated the usual split-attention effect. Novice students presented diagrams and text in a format that separated the two sources of information learned less than novice students given materials that integrated the texts into the diagrams. Physical integration reduced the need for mental integration and reduced extraneous cognitive load. As levels of expertise increased, the difference between the separate and integrated conditions first disappeared and eventually reversed with the separate condition superior to the integrated condition. Indeed, rather than integrating the diagrams and text, totally eliminating the text was superior. The text had become redundant for these more expert learners.

Initially, for complete novices, both the diagrams and text were essential for learning. Under such conditions, extraneous cognitive load could be reduced by physically integrating the two sources of information in order to reduce the need for learners to mentally integrate them. Reducing the need for mental integration reduces extraneous cognitive load. As expertise increased, the textual explanations became less and less important. Eventually they were unnecessary, but if presented to learners with more experience in the area in an integrated format, they were hard to ignore. Processing such redundant information imposed an extraneous cognitive load reducing further learning. The redundancy effect had replaced the split-attention effect as expertise increased, providing an example of the expertise reversal effect.

Similar results were obtained by Yeung *et al.* (1998) using purely textual materials. McNamara *et al.* (1996), who were not working within a CLT framework, found that low-knowledge learners benefited from additional explanatory text intended to increase coherence whereas high-knowledge learners benefited most from the more sparse text. Kalyuga *et al.* (2000) found that among novices, dual mode, auditory/visual presentations were superior to visual only presentations, demonstrating the modality effect. With more experience, the auditory component became redundant and was best eliminated. In their experiment with novices, Kalyuga *et al.* (2001) and Tuovinen and Sweller (1999) demonstrated the worked example effect in which worked examples were superior to solving the equivalent problems. With increasing knowledge, the effect first disappeared and then reversed. Worked examples become redundant for more knowledgeable learners and imposed an extraneous cognitive load. (See Kalyuga *et al.*, 2003, for other examples of the expertise reversal effect.)

All of these examples of the expertise reversal effect have strong implications for the design of instruction, but the elimination and reversal of the worked examples effect with increased knowledge has particularly important implications. It suggests that a training program best starts with worked examples and smoothly works up to conventional problems. In the early 1990s, van Merriënboer (1990; van Merriënboer and de Croock, 1992) introduced the so-called *completion strategy* to reach this goal. The key to this strategy is the use of completion problems, in which a problem solution is presented in partially completed form with the learner required to find the remaining steps. The training starts with full worked examples that include all of the solution steps, proceeds with a series of completion problems that provide fewer and fewer of the required steps, and ends with conventional problems that require all of the required steps. In two studies in the domain of computer programming, the completion strategy was effective in obtaining transfer of learning.

Renkl and Atkinson (2003) reviewed recent studies that applied the completion strategy, which they call a fading guidance procedure, in the domains of physics (electricity) and statistics (probability). They explicitly related the strategy to the expertise reversal effect and to different phases in skill acquisition. In the early phases of skill acquisition, germane cognitive load is increased by self-explanations of illustrated principles and by generalization over examples; in later phases, germane cognitive load is increased by the anticipation of solution steps and imagining; and in the final phases, it is increased by genuine problem solving. Four reviewed studies (described in Renkl and Atkinson, 2001; Renkl *et al.*, 2000; Renkl *et al.*, 2002) indicated that the completion strategy is effective in managing cognitive load throughout the various phases of skill acquisition. Compared to the use of traditional example-problem pairs, the completion strategy yielded superior transfer performance in all four studies.

Although germane cognitive load is a valuable construct, the studies reviewed in this section also indicate that this concept alone cannot fully describe the observed phenomena because the level of expertise of the learner directly influences what constitutes germane load. Future CLT research should provide a more detailed description of how cognitive structures develop with expertise, how they enable learners to adapt increasingly to task demands (Ericsson *et al.*, 2000), and how they relate to improvements in self-regulated learning (Zimmerman, 2002).

FROM PREPLANNED INSTRUCTION TO ADAPTIVE eLEARNING

CLT researchers have mainly studied the effects of preplanned instruction on cognitive load and transfer performance. However, the expertise reversal effect in combination with the application of CLT to courses with a longer duration require a more dynamic approach, wherein instruction can be adapted in real-time to the increasing levels of expertise of individual learners. This approach requires two new developments. First, assessment methods are needed to measure the levels of expertise of learners in such a way that cognitive load is taken into account. Second, research is needed to indicate how these measures can promote effective, efficient, and appealing forms of adaptive instruction, which today take the form of adaptive eLearning.

Assessment of Expertise

Traditionally, assessment in education primarily deals with *performance*, defined in terms of a learner's achievement as measured by the number of correct test answers, the number of errors, or the time on task. Performance is one assessment dimension of cognitive load, because a higher cognitive load often increases the number of errors and slows down performance (see Paas and van Merriënboer, 1994b). However, CLT stresses that other dimensions are at least equally important for the assessment of expertise. They include *mental load*, which originates from the interaction between task characteristics (e.g., task format, multimedia, task complexity) and learner characteristics (e.g., age, prior knowledge, spatial ability), and yields an a priori estimate of cognitive load and *mental effort*, which refers to the cognitive capacity that is actually allocated to accommodate the demands imposed by the task (Paas and van Merriënboer, 1993, 1994b). The intensity of effort being expended by learners is considered essential to obtaining a reliable estimate of cognitive load. As indicated in the previous section, mental effort may yield important information that is not necessarily reflected in mental load and performance measures. For instance, it is feasible for two people to attain the same performance levels with one person needing to work laboriously through an effortful process to arrive at the correct answers, whereas the other person reaches the same answers with minimum effort. Although both people demonstrate identical performance, "expertise" might be higher for the person who performs the task with minimum effort than for the person who exerts substantial effort.

An appropriate assessment of expertise should at least then include measures of mental effort and performance. Paas *et al.* (2003b) discuss different measurement techniques for mental effort, including rating scales, secondary task methods, and psychophysiological measures. Whereas more recent techniques such as secondary task methods (e.g., Brünken *et al.*, 2003) and psychophysiological measures (e.g., van Gerven *et al.*, 2004) are used to measure cognitive load, to this point, most researchers have used rating scales. On the basis of a comprehensive review of about 30 studies, Paas *et al.* (2003b) conclude that "...the use of rating scales to measure mental effort and cognitive load remains popular, because they are easy to use; do not interfere with primary task performance; are inexpensive; can detect small variations in workload (i.e., sensitivity); are reliable; and provide decent convergent, construct, and discriminate validity" (p. 68).

Although mental effort measurement using rating scales is a straightforward and widely used experimental method, the measurement of complex performance is still a difficult and time-consuming process. In particular, CLT is interested in measures that reflect the quality of available cognitive schemata, and most performance measures are not particularly designed with this end in mind. Kalyuga and Sweller (2004) proposed a "rapid assessment test" to measure the quality of learners' schemata that guide their problem-solving process. The rapid assessment test asked students to indicate their *first* step towards solution of a task. As a simple example, when presented the algebraic problem $(3x - 5)/2 = 5$, students may respond in one of the following ways when asked to report their first step:

- The first step reported is incorrect or the student indicates that (s)he "doesn't know the answer"—this student is categorized as a prenovice with no relevant schemata.
- $3x - 5 = 10$ is indicated as the first step by a student who first multiplies both sides of the equation by 2—this student is categorized as a novice.
- $3x = 15$ is indicated as the first step by a student who mentally multiplies both sides of the equation by ten and adds 5 to both sides of the equation—this student is categorized as having intermediate ability.
- $x = 15/3$ is indicated as the first step by a student who mentally divides both sides by 3 and immediately writes the final answer—this student is categorized as an advanced student.
- $x = 5$ is indicated as the first step by a student who has automated the whole procedure—this student is categorized as an expert.

The assumption that higher quality schemata allow for the skipping of steps is central to this type of rapid assessment. The skipping of steps is an important characteristic of higher levels of expertise because well learned or automated solution procedures chunk cognitive rules together that consistently follow each other in performing particular tasks (Blessing and Anderson, 1996; Sweller *et al.*, 1983). Kalyuga and Sweller (in press) found high correlations (up to .92) between performance on "rapid assessment tests" and traditional performance tests that required complete solutions of corresponding tasks. This indicates that the first step is indeed a good indicator of the quality of available problem-solving schemas. Furthermore, Kalyuga and Sweller (2004) presented data indicating that task selection based on the rapid assessment of expertise facilitates learning.

A final step in the assessment of expertise is the difficult task of combining a student's mental effort and performance measures, because a meaningful interpretation of a certain level of cognitive load can only be made in the context of its associated performance and vice versa. Paas and van Merriënboer (1993; see also Paas *et al.*, 2003b) developed a computational approach to combine measures of mental effort with measures of associated performance in order to compare the efficiency of instructional conditions—under the assumption that learners' behavior in a particular condition is more efficient if their performance is higher than might be expected on the basis of their invested mental effort or, equivalently, if their invested mental effort is lower than might be expected on the basis of their performance. Using this approach, high task performance associated with low effort is called high-instructional efficiency, whereas low task performance with high effort is called low-instructional efficiency. Unfortunately, this approach is based on standardized measures and can only be used after gathering all data from a group of students working under different instructional conditions. Alternative methods are needed for the continuous assessment of expertise of individual learners. Such alternatives are currently under development in the context of adaptive eLearning.

Adaptive eLearning

Salden *et al.* (in press) described adaptive eLearning as a straightforward two-step cycle: (1) assessment of a learner's expertise, and (2) the dynamic selection of the next learning task. With regard to the ongoing assessment of expertise, they differentiated between a learner who must work laboriously to attain a certain performance level and a learner who attains the same performance level with little mental effort. Only the second learner who solved the problem with minimal mental effort should be presented with a more complex learning task. With regard to task selection, given the learner's expertise, one might select tasks that are less, equally, or more difficult than the previous task; one might vary the format of the task (e.g., worked examples, completion problems, or conventional problems), or one might vary the amount of problem-solving support that is given to the learner.

In the domain of Air Traffic Control (ATC), Camp *et al.* (2001) and Salden *et al.* (2004) compared the effectiveness of four approaches: (1) a fixed sequence of learning tasks ordered from simple to complex, (2) dynamic task selection based on efficiency, (3) dynamic selection based on performance, and (4) dynamic selection based on mental effort. It was hypothesized that dynamic task selection results in higher test performance than a fixed sequence of learning tasks and that dynamic task selection, on the basis of efficiency, results in higher test performance than the other two dynamic conditions (mental effort and performance). Mental effort and performance were both measured on a 5-point scale. The data for the fixed group were gathered first and used as a baseline for the other three dynamic groups.

In the dynamic conditions, learners received ATC tasks at 10 levels of difficulty, starting at the lowest level. Depending on the assessment results, the next task was selected. For instance, a student who attains a performance score of 4 whereas his or her cognitive load is 3 is presented with a learning task that is one complexity level higher than the previous task; another student who attains a performance score of 4 whereas his or her cognitive load is 1 is presented with a learning task that is two complexity levels higher than the previous task. In both studies (Camp *et al.*, 2002; Salden *et al.*, 2004), adaptive eLearning proved superior to the use of a fixed sequence of tasks. But dynamic task selection on the basis of efficiency was not more effective than the other dynamic conditions. However, in the study of Salden *et al.* (2004), the mental efficiency condition appeared more effective during training than the mental effort and performance conditions. The participants in the mental efficiency condition needed fewer learning tasks to reach the highest complexity level, reached a higher complexity level, and made larger jumps to higher complexity levels than students in the other dynamic conditions.

In order to obtain an optimal indicator of a learner's expertise, Kalyuga and Sweller (in press) took a different approach to combining performance and mental effort measures than Camp *et al.* (2001) and Salden *et al.* (2004). In the domain of algebra, they used the "rapid assessment test" as described in the previous section to measure performance and used a 9-point rating scale to measure mental effort. *Cognitive efficiency* (E) was defined as $E = P/R$, where R is the mental effort rating and P is the performance measure on the same task. On the one hand, this indicator is similar to the efficiency indicator as defined by Paas and van Merriënboer (1993), because both indicators increase as similar levels of performance are reached with less effort, or as higher levels of performance are reached with the same effort. On the other hand, this new indicator makes it unnecessary to use a baseline group such as for the efficiency indicator of Paas and van Merriënboer (1993), because there is no need to standardize measures across groups.

In the study of Kalyuga and Sweller (in press), the cognitive efficiency indicator was used to monitor learners' progress during instruction. Learners were presented tasks at different levels of difficulty. For each level, a critical level of cognitive efficiency (Ecr) was arbitrarily defined as the maximum performance score (which was different per task level) divided by the maximum mental effort score (which was always 9). Cognitive efficiency is positive if $E >$ Ecr and negative if $E <$ Ecr. Thus, if someone invests maximum mental effort in a task but does not display the maximum level of performance, then expertise should be regarded as suboptimal; if someone performs at the maximum level with less than a maximum mental effort, then expertise should be regarded as optimal.

In the adaptive eLearning condition in the study of Kalyuga and Sweller (in press), learners were presented algebra tasks at three different levels. If their cognitive efficiency was negative for tasks at the lowest level, then they continued with the study of worked examples; if their cognitive efficiency was positive for tasks at the lowest level but negative for tasks at the second level, then they continued with simple completion problems; if their cognitive efficiency was positive for tasks at the lowest and second level but negative for tasks at the third level, then they continued with difficult completion problems; and, finally, if their cognitive efficiency was positive for tasks at all three levels, then they continued with conventional problems. Similar adaptive methods were applied when students were working on the worked examples, completion problems, or conventional problems. Each student in the adaptive eLearning condition was paired to a student in the control condition, who served as a yoked control. Kalyuga and Sweller (in press) report higher gains in algebraic skills from pretest to posttest and higher gains in cognitive efficiency for the adaptive eLearning group than for the control group. Thus, in agreement with Camp *et al.* (2001) and Salden *et al.* (2004), adaptive eLearning was found to be superior to non-adaptive learning.

DISCUSSION

In this article, we have described basic CLT and reviewed important developments of this theory driven by both new theoretical questions and changes in the field of instructional design. Four major developments in current CLT research were discussed, namely: (1) a strengthening of the cognitive underpinnings of the theory by more closely tying it to biological evolution, (2) research on methods to decrease the high cognitive load associated with the use of learning tasks that are based on real-life tasks, (3) research to take learners' motivation and their development of expertise during lengthy courses or training programs into account, and (4) research to assess learners' expertise on the basis of performance and cognitive load in order to flexibly adapt instruction to the needs of individuals. These new developments provide direction for future research and should eventually increase the practicability of CLT.

Without its particular view of human cognitive architecture, CLT would not have generated its research programs or its instructional effects. The theory about cognitive architecture is strengthened by closely tying it to the information processes of biological evolution. By emphasizing that both biological evolution and human cognition constitute variants of a single natural information processing system, it is possible to place both cognitive architecture and instructional design in a wider, more sophisticated context.

Because modern instructional theories often incorporate real-life learning tasks as a driving force, cognitive load considerations are becoming increasingly important. A first research line may develop instructional methods to manage the cognitive load associated with those tasks, while leaving the essence of the "whole" task intact. Methods to gradually increase the element interactivity and the associated intrinsic cognitive load in the course of a training program are growing in importance. A key question is: with which elements, or sets of interactions, should the instruction commence? Other methods support the problem-solving process. Preliminary results show that "external" support structures, such as process worksheets, are not ideal for reducing extraneous cognitive load when students are working on rich learning tasks. Future research may investigate if alternative methods, such as making ample use of process-oriented worked examples or reducing students' problem space by setting performance constraints, yield the desired effects on cognitive load and transfer.

In long duration training programs, students are not always inclined to devote their processing resources, freed up by decreasing intrinsic and/or extraneous cognitive load, to learning. A second research line may study instructional methods that motivate students to invest effort in processes that generate germane cognitive load, such as schema construction and automation. In addition to variability of practice, promising methods include the use of (pre) training in elaboration strategies and the use of prompts for self-explanations. Other instructional strategies are related to the expertise reversal effect. At the very least, such strategies should take learners' expertise development into account and prescribe different instructional methods for the various phases of a training program. Several studies have demonstrated that the completion strategy, which starts with worked examples and smoothly moves to conventional problems, is an effective approach. Alternative guidance-fading strategies can be devised and empirically tested for their effectiveness.

Even more powerful approaches designed to take learners' expertise into account require an ongoing measurement of the level of expertise for each student, so that instruction can be flexibly adapted to this level. This third research line should first identify optimal methods to measure learner's expertise so that cognitive load is taken into account. There are several approaches to gather and combine real-time cognitive load and performance measures while students are working on their learning tasks (Kalyuga and Sweller, 2004; Salden *et al.*, 2004). Future research should reveal the strengths and weaknesses of using cognitive load and performance measures to determine expertise and should make critical, data-based comparisons. Instructional methods should also permit a dynamic adjustment of instruction to a particular student's level of expertise. Preliminary results indicate that adaptive eLearning is superior to non-adaptive eLearning. However, adaptive instruction can take many different forms. For instance, adaptation may refer to the difficulty of the next learning task or to the nature and amount of support given for this task. The wide range of adaptive instructional forms opens the door for a systematic research program on CLT and adaptive eLearning.

The new research lines discussed in this article are derived from CLT but also indicate where the theory requires expansion and modification to remedy limitations. If CLT is to guide the design of instruction for increasingly advanced learners with decreasing instructional support, then the extensions require a theory of how complex schemata mediating the performance of advanced learners are gradually acquired or constructed. A future version of CLT should explain how acquired schemata allow learners to better adapt to task demands (Ericsson *et al.*, 2000) and perhaps how expertise is related to self-regulated learning (Zimmerman, 2002). The assessment of learners' expertise should focus on accessible cognitive structures and diagnose learners' ability to monitor their own learning. In effect, these suggested developments require a deeper knowledge of human cognition. As indicated above, CLT has expanded in this direction by considering the evolutionary implications of human cognition. This work is intended to provide a strengthened cognitive underpinning for instructional design. But additional research is needed to revise and expand CLT to fully account for complex learning.

The studies discussed in this article also signify the growing practicability of CLT for an increasing number of design levels, media, and target groups. With regard to design levels, CLT is slowly evolving from a theory for instructional *message* design, with a strong focus on presentation formats, to a full-fledged instructional design model. CLT is becoming useful for designing larger courses and training programs that are characterized by a high level of interactivity, as indicated by methods for sequencing and providing problem-solving support, for encouraging students to invest germane load in learning and taking their levels of expertise into account, and for student assessment and the development of adaptive forms of instruction. van Merriënboer's four-component instructional design model (4C/ID-model, 1997; van Merriënboer *et al.*, 2002a) includes many instructional methods that were developed in the context of CLT. Recent efforts to further integrate CLT with the 4C/ID-model are described by van Merriënboer *et al.* (2003; see also van Merriënboer and Paas, 2003).

With regard to media, CLT is no longer limited to straightforward instructional presentations or linear media but has found several applications in the design of multimedia systems (for an overview, see Mayer and Moreno, 2002, 2003). These applications include highly interactive, learner-controlled (hyper) media (e.g., Gerjets and Scheiter, 2003; Leahy *et al.*, 2003; Tabbers *et al.*, 2004) as well as learning environments for Computer Supported Collaborative Learning (CSCL; see van Bruggen *et al.*, 2002). CLT is one of the few theories beginning to generate a coherent and research-based set of instructional guidelines for multimedia systems and (adaptive) eLearning applications.

With regard to target groups, applications of CLT are extending from students in primary and secondary education to adult learners, including the elderly. van Gerven *et al.* (2000) reported that CLT is of particular relevance for designing instruction for target groups characterized by impaired working-memory functions. Research has demonstrated that cognitive aging is accompanied by a reduction of working-memory capacity, a general slowing of mental processes, and a decline in the ability to repress irrelevant information. Because instruction based on CLT deals with the need for an efficient use of available resources due to cognitive limitations, CLT based instruction is especially effective when elderly people are involved. Indeed, the results of recent studies indicate that instruction based on CLT is relatively more effective for older learners than for younger learners (Paas *et al.*, 2001; van Gerven *et al.*, 2002, 2003).

As can be seen from this review, CLT has undergone major developments over the last 5 years. New instructional methods were derived from the theory, and because there is a close association between these methods and the cognitive architecture assumed by CLT, we believe that applications of CLT will yield instruction that is compatible with human cognitive processing. Critical to the further development of CLT is a rigorous testing program based on replicated and controlled experimental designs. Such rigor in experimental methods is too often missing in educational research but is needed to develop sound instructional theories capable of making a real difference to educational practice. There is no substitute for evidence-based instructional design.

REFERENCES

Bannert, M. (2002). Managing cognitive load: Recent trends in cognitive load theory. *Learn. Instruct.* 12: 139–146.

Blessing, S. B., and Anderson, J. R. (1996). How people learn to skip steps. *J. Exp. Psychol.: Learn., Memory, Cognit.* 22: 576–598.

Britton, B. K., and Glynn, S. M. (1982). Effects of text structure on use of cognitive capacity during reading. *J. Educ. Psychol.* 74: 51–61.

Brünken, R., Plass, J. L., and Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educ. Psychol.* 38: 53–62.
Camp, G., Paas, F., Rikers, R., and van Merriënboer, J. J. G. (2001). Dynamic problem selection in air traffic control training: A comparison between performance, mental effort and mental efficiency. *Comput. Hum. Behav.* 17: 575–595.
Carlson, R., Chandler, P., and Sweller, J. (2003). Learning and understanding science instructional material. *J. Educ. Psychol.* 95: 629–640.
De Croock, M. B. M., van Merriënboer, J. J. G., and Paas, F. (1998). High vs. low contextual interference in simulation-based training of troubleshooting skills: Effects on transfer performance and invested mental effort. *Comput. Hum. Behav.* 14: 249–267.
Ericsson, K. A., and Kintsch, W. (1995). Long-term working memory. *Psychol. Rev.* 102: 211–245.
Gerjets, P., and Scheiter, K. (2003). Goal configurations and processing strategies as moderators between instructional design and cognitive load: Evidence from hypertext-based instruction. *Educ. Psychol.* 38: 33–42.
Goettl, B. P., and Shute, V. J. (1996). Analysis of part-task training using the backward-transfer technique. *J. Exp. Psychol.: Appl.* 2: 227–249.
Hummel, H. G. K., and Nadolski, R. J. (2002). Cueing for schema construction: Designing problem-solving multimedia practicals. *Contemp. Educ. Psychol.* 27: 229–249.
Kalyuga, S., Ayres, P., Chandler, P., and Sweller, J. (2003). The expertise reversal effect. *Educ. Psychol.* 38: 23–31.
Kalyuga, S., Chandler, P., and Sweller, J. (1998). Levels of expertise and instructional design. *Hum. Factors* 40: 1–17.
Kalyuga, S., Chandler, P., and Sweller, J. (2000). Incorporating learner experience into the design of multimedia instruction. *J. Educ. Psychol.* 92: 126–136.
Kalyuga, S., Chandler, P., Tuovinen, J., and Sweller, J. (2001). When problem solving is superior to studying worked examples. *J. Educat. Psychol.* 93: 579–588.
Kalyuga, S., and Sweller, J. (in press). Rapid dynamic assessment of expertise to optimize the efficiency of e-learning. *Educ. Technol., Res. Dev*.
Kalyuga, S., and Sweller, J. (2004). Measuring knowledge to optimize cognitive load factors during instruction. *J. Educ. Psychol.* 96: 558–568.
Kester, L., Kirschner, P. A., and van Merriënboer, J. J. G. (2004). Timing of information presentation in learning statistics. *Instruct. Sci.* 32: 233–252.
Kester, L., Kirschner, P. A., and van Merriënboer, J. J. G. (2005). The management of cognitive load during complex cognitive skill acquisition by means of computer simulated problem solving. *Br. J. Educ. Psychol.* 75: 71–86.
Kester, L., Kirschner, P. A., and van Merriënboer, J. J. G. (in press). Just-in-time information presentation: Improving learning a complex troubleshooting skill. *Contemp. Educ. Psychol.*
Kester, L., Kirschner, P. A., van Merriënboer, J. J. G., and Baumer, A. (2001). Just-in-time information presentation and the acquisition of complex cognitive skills. *Comput. Hum. Behav.* 17: 373–391.
Kirschner, P. A. (ed.). (2002). Cognitive load theory. *Learn. Instruct.* 12(special issue): 1–154.
Leahy, W., Chandler, P., and Sweller, J. (2003). When auditory presentations should and should not be a component of multimedia instruction. *Appl. Cognit. Psychol.* 17: 401–418.
Leutner, D. (2000). Double-fading support: A training approach to complex software systems. *J. Comput. Assist. Learn.* 16: 347–357.
Marcus, N., Cooper, M., and Sweller, J. (1996). Understanding instructions. *J. Educ. Psychol.* 88: 49–63.
Mayer, R. E., and Moreno, R. (2002). Aids to computer-based multimedia learning. *Learn. Instruct.* 12: 107–119.
Mayer, R. E., and Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educ. Psychol.* 38: 43–52.

McKeough, A., Marini, A., and Lupart, J. L. (eds.). (1995). *Teaching for Transfer: Fostering Generalization in Learning*. Hillsdale, NJ: Erlbaum.

McNamara, D., Kintsch, E., Songer, N. B., and Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognit. Instruct.* 14: 1–43.

Merrill, M. D. (2002). First principles of instructional design. *Educ. Technol., Res. Dev.* 50: 43–59.

Mousavi, S., Low, R., and Sweller, J. (1995) Reducing cognitive load by mixing auditory and visual presentation modes. *J. Educ. Psychol.* 87: 319–334.

Nadolski, R. J., Kirschner, P. A., and van Merriënboer, J. J. G. (in press). Optimizing the number of steps in learning tasks for complex skills. *Br. J. Educ. Psychol.*

Nadolski, R. J., Kirschner, P. A., van Merriënboer, J. J. G., and Hummel, H. G. K. (2001). A model for optimizing step size of learning tasks in competency-based multimedia practicals. *Educ. Technol., Res. Dev.* 49: 87–103.

Paas, F., Camp, G., and Rikers, R. (2001). Instructional compensation for age-related cognitive declines: Effects of goal specificity in maze learning. *J. Educ. Psychol.* 93: 181–186.

Paas, F., Renkl, A., and Sweller, J. (eds.). (2003a). Cognitive load theory. *Educ. Psychol.* 38(special issue): 1–71.

Paas, F., Renkl, A., and Sweller, J. (eds.). (2004). Cognitive load theory. *Instruct. Sci.* 32(1–2; special issue): 1–182.

Paas, F., Tuovinen, J. E., Tabbers, H., and van Gerven, P. (2003b). Cognitive load measurement as a means to advance cognitive load theory. *Educ. Psychol.* 38: 63–71.

Paas, F., and van Merriënboer, J. J. G. (1993). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Hum. Factors* 35: 737–743.

Paas, F., and van Merriënboer, J. J. G. (1994a). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive load approach. *J. Educ. Psychol.* 86: 122–133.

Paas, F., and van Merriënboer, J. J. G. (1994b). Instructional control of cognitive load in the training of complex cognitive tasks. *Educ. Psychol. Rev.* 6: 51–71.

Peck, A. C., and Detweiler, M. C. (2000). Training concurrent multistep procedural tasks. *Hum. Factors* 42: 379–389.

Pollock, E., Chandler, P., and Sweller, J. (2002). Assimilating complex information. *Learn. Instruct.* 12: 61–86.

Quilici, J. L., and Mayer, R. E. (1996). Role of examples in how students learn to categorize statistics word problems. *J. Educ. Psychol.* 88: 144–161.

Reigeluth, C. M. (1999). The elaboration theory: Guidance for scope and sequence decisions. In Reigeluth, C. M. (ed.), *Instructional design theories and models: A new paradigm of instructional theory*, Vol. 2, Erlbaum, Mahwah, NJ, pp. 424–453.

Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognit. Sci.* 21: 1–29.

Renkl, A., and Atkinson, R. K. (2001, August). *The effects of gradually increasing problem-solving demands in cognitive skill acquisition*. Paper presented at the 9th Conference of the European Association for Research on Learning and Instruction (EARLI), Fribourg, Switzerland.

Renkl, A., and Atkinson, R. K. (2003). Structuring the transition from example study to problem solving in cognitive skill acquisition: A cognitive load perspective. *Educ. Psychol.* 38: 15–22.

Renkl, A., Atkinson, R. K., and Maier, U. H. (2000). From studying examples to solving problems: Fading worked-out solution steps helps learning. In Gleitman, L., and Joshi, A. K. (eds.), *Proceedings of the 22nd Annual Conference of the Cognitive Science Society*, Erlbaum, Mahwah, NJ, pp. 393–398.

Renkl, A., Atkinson, R. K., Maier, U. H., and Staley, R. (2002). From example study to problem solving: Smooth transitions help learning. *J. Exp. Educ.* 70: 293–315.

Salden, R. J. C. M., Paas, F., Broers, N. J., and van Merriënboer, J. J. G. (2004). Mental effort and performance as determinants for the dynamic selection of learning tasks in air traffic control training. *Instruct. Sci.* 32(1–2): 153–172.

Salden, R. J. C. M., Paas, F., and van Merriënboer, J. J. G. (in press). A comparison of approaches to learning task selection in the training of complex cognitive skills. *Comput. Hum. Behav.*

Schwartz, D. L., Bransford, J. D., and Sears, D. (in press). Efficiency and innovation in transfer. In Mestre, J. (ed.), *Transfer of Learning: Research and Perspectives,* Information Age Publishing, Greenwich, CT.

Stark, R., Mandl, H., Gruber, H., and Renkl, A. (2002). Conditions and effects of example elaboration. *Learn. Instruct.* 12: 39–60.

Sweller, J. (2003). Evolution of human cognitive architecture. In Ross, B. (ed.), *The Psychology of Learning and Motivation*, Vol. 43, Academic Press, San Diego, pp. 215–266.

Sweller, J. (2004). Instructional design consequences of an analogy between evolution by natural selection and human cognitive architecture. *Instruct. Sci.* 32(1/2): 9–31.

Sweller, J., and Chandler, P. (1994). Why some material is difficult to learn. *Cognit. Instruct.* 12: 185–233.

Sweller, J., Mawer, R., and Ward, M. (1983). Development of expertise in mathematical problem solving. *J. Exp. Psychol.: Gen.* 112: 639–661.

Sweller, J., van Merriënboer, J. J. G., and Paas, F. (1998). Cognitive architecture and instructional design. *Educ. Psychol. Rev.* 10: 251–296.

Tabbers, H., Martens, R., and van Merriënboer, J. J. G. (2004). Multimedia instructions and cognitive load theory: Effects of modality and cueing. *Br. J. Educ. Psychol.* 74: 71–81.

Tindall-Ford, S., Chandler, P., and Sweller, J. (1997). When two sensory modes are better than one. *J. Exp. Psychol.: Appl.* 3: 257–287.

Tuovinen, J., and Sweller, J. (1999). A comparison of cognitive load associated with discovery learning and worked examples. *J. Educ. Psychol.* 91: 334–341.

van Bruggen, J. M., Kirschner, P. A., and Jochems, W. (2002). External representations of argumentation in CSCL and the management of cognitive load. *Learn. Instruct.* 12: 121–138.

van Gerven, P. W. M., Paas, F., van Merriënboer, J. J. G., Hendriks, M., and Schmidt, H. G. (2003). The efficiency of multimedia learning into old age. *British J. Educ. Psychol.* 73: 489–505.

van Gerven, P. W. M., Paas, F., van Merriënboer, J. J. G., and Schmidt, H. G. (2000). Cognitive load theory and the acquisition of complex cognitive skills in the elderly: Towards an integrative framework. *Educ. Gerontol.* 26: 503–521.

van Gerven, P. W. M., Paas, F., van Merriënboer, J. J. G., and Schmidt, H. G. (2002). Cognitive load theory and aging: Effects of worked examples on training efficiency. *Learn. Instruct.* 12: 87–105.

van Gerven, P. W. M., Paas, F., van Merriënboer, J. J. G., and Schmidt, H. G. (2004). Memory load and the cognitive pupillary response in aging. *Psychophysiology* 41: 167–174.

van Gog, T., Paas, F., and van Merriënboer, J. J. G. (2004). Process-oriented worked examples: Improving transfer performance through enhanced understanding. *Instruct. Sci.* 32(1/2): 83–98.

van Merriënboer, J. J. G. (1990). Strategies for programming instruction in high school: Program completion vs. program generation. *J. Educ. Comput. Res.* 6: 265–287.

van Merriënboer, J. J. G. (1997). *Training Complex Cognitive Skills*, Educational Technology Publications, Englewood Cliffs, NJ.

van Merriënboer, J. J. G. (2000). The end of software training? *J. Comput. Assist. Learn.* 16: 347–357.

van Merriënboer, J. J. G., Clark, R. E., and de Croock, M. B. M. (2002a). Blueprints for complex learning: The 4C/ID-model. *Educ. Technol., Res. Dev.* 50(2): 39–64.

van Merriënboer, J. J. G., and de Croock, M. B. M. (1992). Strategies for computer-based programming instruction: Program completion vs. program generation. *J. Educ. Comput. Res.* 8: 365–394.

van Merriënboer, J. J. G., de Croock, M. B. M., and Jelsma, O. (1997). The transfer paradox: Effects of contextual interference on retention and transfer performance of a complex cognitive skill. *Percept. Motor Skills* 84: 784–786.
van Merriënboer, J. J. G., Kirschner, P. A., and Kester, L. (2003). Taking the load off a learner's mind: Instructional design for complex learning. *Educ. Psychol.* 38: 5–13.
van Merriënboer, J. J. G., and Paas, F. (2003). Powerful learning and the many faces of instructional design: Toward a framework for the design of powerful learning environments. In de Corte, E., Verschaffel, L., Entwistle, N., and van Merriënboer, J. (eds.), *Unravelling Basic Components and Dimensions of Powerful Learning Environments,* Elsevier Science, Oxford, UK, pp. 1–20.
van Merriënboer, J. J. G., Schuurman, J. G., de Croock, M. B. M., and Paas, F. (2002b). Redirecting learners' attention during training: Effects on cognitive load, transfer test performance and training efficiency. *Learn. Instruct.* 12: 11–37.
Williams, T. R., and Butterfield, E. C. (1992). Advance organizers: A review of the research. *J. Tech. Writing Commun.* 22: 259–272.
Yeung, A., Jin, P., and Sweller, J. (1998). Cognitive load and learner expertise: Split-attention and redundancy effects in reading with explanatory notes. *Contemp. Educ. Psychol.* 23: 1–21.