Cognitive Architecture and Instructional Design

Article *in* Educational Psychology Review · September 1998

DOI: 10.1023/a:1022193728205

Jeroen J. G. Van Merrienboer
Maastricht University
501 PUBLICATIONS **29,057** CITATIONS

Fred Paas
Erasmus University Rotterdam and University of Wollongong
306 PUBLICATIONS **29,730** CITATIONS

Educational Psychology Review, Vol. 10, No. 3, 1998

Cognitive Architecture and Instructional Design

John Sweller,[1,4] Jeroen J. G. van Merrienboer,[2] and
Fred G. W. C. Paas[3]

Cognitive load theory has been designed to provide guidelines intended to assist in the presentation of information in a manner that encourages learner activities that optimize intellectual performance. The theory assumes a limited capacity working memory that includes partially independent subcomponents to deal with auditory/verbal material and visual/2- or 3-dimensional information as well as an effectively unlimited long-term memory, holding schemas that vary in their degree of automation. These structures and functions of human cognitive architecture have been used to design a variety of novel instructional procedures based on the assumption that working memory load should be reduced and schema construction encouraged. This paper reviews the theory and the instructional designs generated by it.

KEY WORDS: cognition; instructional design; learning; problem solving.

INTRODUCTION

The expansion in knowledge of cognitive structures and processes in recent years has provided a new and very promising source of research hypotheses associated with instructional design principles. This paper reviews aspects of the instructional design literature that have been generated by cognitive load theory, a theory that emphasizes working memory constraints as determinants of instructional design effectiveness. We begin by discussing human cognitive architecture, its relations to some basic categories of information, and its consequences for learning phenomena such as

---

[1] University of New South Wales, Australia.
[2] University of Twente, the Netherlands. Now at University of Maastricht, the Netherlands.
[3] University of Maastricht, the Netherlands.
[4] Correspondence should be directed to John Sweller, School of Education Studies, University of New South Wales, Sydney 2052, Australia; e-mail: j.sweller@unsw.edu.au

generalization and transfer. This discussion of human cognitive architecture includes an outline of cognitive load theory and its general principles. Issues associated with measuring cognitive load are discussed next. Lastly, a description of the instructional design principles generated by cognitive load theory and the experiments providing evidence for the effectiveness of the principles are considered.

SOME ASPECTS OF HUMAN COGNITIVE ARCHITECTURE

Working Memory

Working memory can be equated with consciousness. Humans are conscious of and can monitor only the contents of working memory. All other cognitive functioning is hidden from view unless and until it can be brought into working memory. The limitations of human working memory are both well-known and widely accepted. Working memory is capable of holding only about seven items or elements of information at a time (Miller, 1956). Furthermore, because working memory is most commonly used to process information in the sense of organizing, contrasting, comparing, or working on that information in some manner, humans are probably only able to deal with two or three items of information simultaneously when required to process rather than merely hold information. Any interactions between elements held in working memory themselves require working memory capacity, reducing the number of elements that can be dealt with simultaneously.

Whereas, initially, working memory was considered a unitary construct, modern theories of working memory place a greater emphasis on partially independent processors. These processors are frequently associated with individual sensory modes. For example, Baddeley's theory (see Baddeley, 1992 for a summary) divides working memory into a "visual-spatial scratch pad" for dealing with visually based information and a "phonological loop" to deal with auditory, primarily speech-based, information. These two systems, in turn, are governed by a central executive that probably more closely resembles the unitary working memory originally proposed by working memory theories. As indicated below, under certain restricted conditions, working memory capacity may be increased by the use of multiple processors rather than by a single, working memory processor.

The implications of working memory limitations on instructional design can hardly be overestimated. All conscious cognitive activity learners engage in occurs in a structure whose limitations seem to preclude all but the most basic processes. Anything beyond the simplest cognitive activities appear to overwhelm working memory. Prima facie, any instructional design that flouts or merely ignores working memory limitations inevitably is deficient. It is this factor that provides a central claim of cognitive load theory.

Despite these apparent restrictions, the intellectual heights to which humans are capable indicate that structures other than working memory must play a critical role in human cognition. The seat of human intellectual skill may more likely reside in long-term rather than working memory.

Long-Term Memory

Humans are not directly conscious of long-term memory. Awareness of its contents and functioning is filtered through working (conscious) memory. It is possibly partly for this reason that knowledge concerning the characteristics of long-term memory and its importance developed somewhat more slowly than was the case for working memory. Furthermore, from the point of view of those concerned with either long-term memory or with instructional design, the topic that provided the initial source of that knowledge may appear bizarre—the game of chess.

De Groot (1966) studied the factors that distinguish the differential ability of chess grand masters and less able players. Although grand masters virtually always defeat weekend players, it was not clear what they knew or what they did that permitted the huge disparity in skill. De Groot established that differential search was not a relevant factor. Grand masters did not appear to consider a greater number of alternatives when searching for a move than did less able players. The only difference that could be established clearly was in memory of board configurations taken from real games. If more and less able players were shown a board configuration taken from a real game for about five seconds and then were asked to reproduce that configuration from memory, grand masters could correctly place most of those pieces, whereas less able players were only able to correctly place far fewer pieces. Furthermore, as Chase and Simon (1973) demonstrated, this difference was not due to a difference in working memory. Chase and Simon found that they could reproduce de Groot's results using configurations taken from real games but found no difference when random configurations were used. This result suggested that working memory was not a relevant factor.

Why should expert chess players be superior at reproducing board configurations taken from real games but not random configurations? Grand masters have spent many years of practice attaining their high level of expertise. De Groot's (1966) results indicate what they learn during those years: to recognize thousands of board configurations that can occur during chess games. According to Simon and Gilmartin (1973), they could learn as many as 100,000 such configurations. As a consequence, grand masters can easily and accurately reproduce configurations taken from real games because each configuration is one with which they are familiar, but they are no better than anybody else at reproducing random configurations with which they are unfamiliar.

Why should memory of board configurations result in superior playing skill? Skilled chess players recognize most of the board configurations they encounter, and they have learned the basic move associated with each configuration. Unlike less-skilled players, they do not have to search for good moves using limited working memory. Rather, they use knowledge of board configurations and the appropriate moves associated with those configurations. This knowledge, acquired after years of practice (see Ericsson and Charness, 1994), is stored in long-term memory and, on current evidence, may be the only factor determining different levels of skill.

Similar findings were obtained in a variety of domains during the late 1970s and early 1980s (e.g., Barfield, 1986; Egan and Schwartz, 1979; Jeffries, Turner, Polson, and Atwood, 1981; Sweller and Cooper, 1985). All studies confirmed that the major factor distinguishing novice from expert problem solvers was not knowledge of sophisticated, general problem-solving strategies but, rather, knowledge of an enormous number of problem states and their associated moves.

There are many instructional design implications that flow from these findings. Later sections of this paper will be concerned with specific design principles. At this point, we wish to emphasize some of the more general consequences that flow from the limited working memory/very large long-term memory model assumed by most cognitive researchers and outlined above. The human cognitive system can be characterized as one that places its primary emphasis on the ability to store seemingly unlimited amounts of information in long-term memory. This information does not just consist of small, isolated facts but can include large, complex interactions and procedures. (The nature of these interactions and procedures are discussed below.) From this view, human intellectual prowess comes from this stored knowledge, not from an ability to engage in long, complex chains of reasoning in working memory. Indeed, knowledge about working memory limitations suggest humans are particularly poor at complex reasoning unless most of the elements with which we reason have previously been stored in long-term memory. Working memory simply is incapable of highly complex interactions using novel (i.e., not previously stored in long-term memory) elements. It follows, that instructional designs and instructional recommendations that require learners to engage in complex reasoning processes involving combinations of unfamiliar elements are likely to be deficient.

Human working memory does not support such activity. Chess grand masters are successful, not because they engage in more sophisticated reasoning procedures than weekend players, but because they have access to knowledge unavailable to others. If anything, it is the less expert players who must engage in complex chains of reasoning but, of course, these are likely to overburden working memory. Novice players must engage in such reasoning, not because it is particularly effective but rather, because they do not have access to knowledge that is effective. When translated to the field of instructional design, it follows that instruction should facilitate domain specific knowledge acquisition, not very general reasoning strategies that cannot possibly be supported by human cognitive architecture.

Schema Construction

If knowledge is the basis of human intellectual skill, what form does that knowledge take? According to schema theory, knowledge is stored in long-term memory in the form of schemas. A schema categorizes elements of information according to the manner in which they will be used (see Chi, Glaser, and Rees, 1982). Thus, chess grand masters have schemas that categorize board pieces into patterns that tell them which moves are appropriate. Schemas can tell us that certain objects are trees to which we can react in a common way even though no two trees have identical elements. When reading, we can derive meaning from an infinite variety of marks on a page because we have schemas that allow us to appropriately categorize letters, words, and combinations of words. Schemas provide the elements of knowledge. According to schema theory, it is through the building of increasing numbers of ever more complex schemas by combining elements consisting of lower level schemas into higher level schemas that skilled performance develops. Often, this acquisition of schemas is an active, constructive process. Reading provides a clear example. In early school years, children construct schemas for letters that allow them to classify an infinite variety of shapes (as occurs in hand writing) into a very limited number of categories. These schemas provide the elements for higher order schemas when they are combined into words that in turn can be combined into phrases, and so forth. Ultimately, this process allows readers to rapidly scan a page filled with a hugely complex array of squiggles and derive meaning from it.

Schemas are stored in long-term memory. One of their obvious functions is to provide a mechanism for knowledge organization and storage. It is not their only function. Schemas also reduce working memory load. Recall, that working memory can process only a limited number of elements at a time. Although the number of elements is limited, the size, complexity, and sophistication of elements is not. A schema can be anything that has been learned and is treated as a single entity. If the learning process has occurred over a long period of time, the schema may incorporate a huge amount of information. Our schema for a restaurant includes extensive knowledge about food and its functions in human affairs; money and its role in exchanging goods and services; the basic architecture of buildings; furniture and how it is used; plus many other facts, functions, processes, and entities. This huge array of elements has been acquired over many years but can be held in working memory, as a single entity. No reader of this paper will find the concept of a restaurant intellectually demanding. It can be held and processed in working memory effortlessly because our restaurant schema acts as a single element. The subelements, or lower-level schemas that are incorporated in the higher-level schemas no longer require working memory capacity. In fact, the huge number of elements that are incorporated into a restaurant schema could not possibly be processed in working memory as individual elements. Because of schema construction, although there are limits on the number of elements that can be processed by working memory, there are no apparent limits on the amount of information that can be processed. A schema, consisting of a single element in working memory has no limits on its informational complexity. In summary, schema construction has two functions: the storage and organization of information in long-term memory and a reduction of working memory load. It can be argued that these two functions should constitute the primary role of education and training systems.

Schema Automation

Automation is an important process in the construction of schemas. All information can be processed either consciously or automatically (Schneider and Shiffrin, 1977; Shiffrin and Schneider, 1977). Conscious processing occurs in working memory and has all the characteristics described above. Automatic processing largely by-passes working memory and has quite different characteristics to conscious processing. Automaticity occurs after practice, normally extensive practice. With sufficient practice, a procedure can be carried out with minimal conscious effort (i.e., with minimal working memory load). For example, most adults can read without consciously processing the individual letters that make up the prose being read. The procedures involved in reading the letters became automated in childhood. In contrast, a young child just learning to read must consciously process each letter.

Kotovsky, Hayes, and Simon (1985) demonstrated the importance of automation in problem solving. They used problems that were structurally identical in that they could be mapped onto each other in a 1:1 fashion (isomorphic problems). These problems differed only in their surface structures. As a consequence of the differences in surface structure, the problem-solving rules (which are indistinguishable from schemas as the term is used in this paper), although structurally identical, differed in their descriptions. Some descriptions could easily be processed in an automatic fashion because they were familiar, whereas others required considerably more conscious processing because they did not connect with familiar procedures.

Kotovsky *et al.* found that the problems could differ in their difficulty by a factor of 16. They attributed this huge difference to differences in automation. Problem solvers using automated rules had substantial working memory reserves to search for a problem solution. When using nonautomated rules, most or perhaps all working memory capacity may have been devoted to retrieving the rules. With little capacity left to engage in a problem-solving search, the achievement of solutions was slow and cumbersome. The differences between the problem variants could be substantially decreased by having problem solvers memorize less familiar rules prior to attempting to solve the problems. Under these conditions, the rules became partially automated with an attendant increase in working memory capacity available for problem search. Indeed, when first attempting to solve these more difficult problems, problem solvers may in effect have unintentionally been engaged solely in the process of automating the rules. Because solution was effectively impossible until the rules had been at least partially automated, the only effect of the initial problem-solving attempts may have been to automate the rules to the point where solution eventually became feasible.

Schemas are examples of sophisticated rules. They probably become automated in exactly the same way as problem-solving rules. When faced with a problem such as $(a + b)/c = d$, solve for a, people may immediately and automatically know that this problem is solved by multiplying out the denominator as the first move. They have an automated schema for this problem that tells them immediately, without conscious processing, how the problem should be solved. In contrast, students who have just learned to solve this category of problem may need to actively attempt to recall the solution procedure: "Do I multiply out the denominator first on this type of problem or do I subtract the addend? I remember now, I multiply out the denominator." Learners who have a more automated schema have more working memory capacity available to use the schema to solve more sophisticated problems. Similarly, a reader who has automated the schemas associated with letters, words and phrases has working memory capacity available to devote to the meaning of the text, whereas less sophisticated readers may be able to read the text perfectly well but not have sufficient working memory capacity available to extract meaning from it.

Automation is therefore an important factor in schema construction. As is the case for schema construction, automation can free working memory capacity for other activities. With automation, familiar tasks are performed accurately and fluidly, whereas unfamiliar tasks—that partially require the automated processes—can be learned with maximum efficiency because maximum working memory capacity is available. Without automation, a previously encountered task may be completed, but performance is likely to be slow and clumsy. Novel tasks may prove to be impossible to complete until prerequisites have not only been learned but also automated, because without automation there may be insufficient working memory capacity to even begin learning and performing the new task. From an instructional design perspective, it follows that designs should not only encourage the construction of schemas, but also the automation of schemas that steer those aspects of a task that are consistent from problem to problem (van Merrienboer, 1997; van Merrienboer, Jelsma, and Paas, 1992). Techniques that encourage automation are discussed in subsequent sections.

Human cognitive architecture can be summarized as follows. We have a limited working memory that deals with all conscious activities and an effectively unlimited long-term memory that can be used to store schemas of varying degrees of automaticity. Intellectual skill comes from the construction of large numbers of increasingly sophisticated schemas with high degrees of automaticity. Schemas both bring together multiple elements that can be treated as a single element and allow us to ignore myriads of irrelevant elements. Working memory capacity is freed, allowing processes to occur that otherwise would overburden working memory. Automated schemas both allow fluid performance on familiar aspects of tasks and—by freeing working memory capacity—permit levels of performance on unfamiliar aspects that otherwise might be quite impossible.

The next question that needs to be asked is how do information structures interact with the cognitive structures discussed in this section. That question is addressed in the next section and provides the core of cognitive load theory.

COGNITIVE LOAD THEORY: SOME INFORMATION STRUCTURES AND THEIR COGNITIVE LOAD CONSEQUENCES

The cognitive architecture discussed in the previous section suggests that prime goals of instruction are the construction and the automation of schemas that are useful for solving the problems of interest. Although schemas are stored in long-term memory, in order to construct them, information must be processed in working memory. Relevant sections of the information must be extracted and manipulated in working memory before being stored in schematic form in long-term memory. The ease with which information may be processed in working memory is a prime concern of cognitive load theory. Working memory load may be affected either by the intrinsic nature of the material (intrinsic cognitive load), or alternatively, by the manner in which the material is presented, or the activities required of students (extraneous cognitive load). Intrinsic cognitive load cannot be altered by instructional interventions because it is intrinsic to the material being dealt with, whereas extraneous cognitive load is unnecessary cognitive load and can be altered by instructional interventions. Extraneous cognitive load is determined by the instructional design (see Sweller, 1994). A further distinction can be made between extraneous cognitive load and germane cognitive load. Although both can be altered by instructional interventions, extraneous cognitive load reflects the effort required to process poorly designed instruction, whereas germane cognitive load reflects the effort that contributes to the construction of schemas. Appropriate instructional designs decrease extraneous cognitive load but increase germane cognitive load.

Intrinsic Cognitive Load and Element Interactivity

Curriculum materials can differ substantially in the extent to which they impose a working memory load. The working memory load imposed 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. An element is anything that has been or needs to be learned, most frequently a schema.

Consider someone who has to learn a new vocabulary, such as a foreign language vocabulary, computer terminology, or chemical symbols. The task may be difficult because there may be a large number of vocabulary items that require learning. Nevertheless, it does not impose a heavy cognitive load. Each element of the task can be learned without reference to any of the other elements. Learning that Fe is the symbol for iron can be accomplished without reference to, for example, the fact that Cu is the symbol for copper or indeed, any other chemical symbols. The task is low in element interactivity in that the elements that must be learned do not interact and so can be learned in isolation. When noninteracting elements can be learned in isolation, intrinsic cognitive load is low because working memory load due to the intrinsic nature of the task is low.

Low-element interactivity tasks allow elements to be learned serially rather than simultaneously. The tasks can be fully understood and learned without holding more than a few elements in working memory at a time. High-element interactivity tasks are at the other end of the continuum. To understand and learn these tasks, several elements must be manipulated in working memory simultaneously. For example, whereas learning the vocabulary of a foreign language is a low element interactivity task, learning the grammatical properties is likely to be a high element interactivity task because the elements interact and because learning them as individual elements may make no sense. As an example, we only can understand the required order of words in the English language by considering all of the words in a phrase. It would make little sense to consider one word at a time. Knowing that *all of the words in a phrase* is the correct order while *in a phrase the words all of* is incorrect, cannot be learned by considering each word individually. Failing to relate each word to the others results in failure to learn the task. Learning word order is a high element interactivity task because all of the elements must be processed in working memory simultaneously. A heavy cognitive load is a consequence.

Mathematical tasks tend to be high in element interactivity. Consider a student who is learning to multiply out a denominator in an equation such as $a/b = c$. To understand the procedure as opposed to rote memorizing it, the student must simultaneously consider that both sides of the equation can be multiplied by b with the equality still retained and that the term, b/b on the left side is a result of multiplying by b and cancelling to *1*, leaving the equation, $a = cb$. Because the elements interact, it makes no sense for students to attempt to learn this procedure sequentially, one element at a time. Attempting to learn what happens on the left side of the equation without simultaneously considering what happens on the right side will result in a lack of mathematical understanding. For students who, unlike most readers of this paper, do not have a schema for this process and who must hold all of the elements individually and simultaneously in working memory, the intrinsic cognitive load may be overwhelmingly high. Not surprisingly, some students learn this procedure merely by learning that the b is shifted to the right to give cb. Learning the procedure in this fashion substantially reduces the cognitive load but at the expense of understanding.

Understanding

The term understanding is applied only when dealing with high element interactivity material (see Marcus, Cooper, and Sweller, 1996). Material is hard to understand when it consists of many interacting elements that cannot be readily held in working memory. Material that can be easily held in working memory is easy to understand. It follows that low element interactivity material is easy to understand. Indeed, if element interactivity is sufficiently low, understanding is so straightforward that we tend not to even use the term understanding. We normally do not refer to someone understanding or failing to understand the translation of the word *cat* from one language to another. Failure to correctly translate normally is said to be due to failure of memory or failure to learn, never failure to understand. In contrast, a student learning to multiply out a denominator in the expression, $a/b = c$, who ends up with the expression, $a/b = cb$, equally has failed to learn or remember part of the procedure. Nevertheless, in this case, the failure tends to be encompassed in the term *failed to understand*. Understanding occurs when high element interactivity material can be held simultaneously in working memory. The mechanism by which this process occurs is discussed next.

Expertise, Schema Construction, and Element Interactivity

Most readers of this paper have no difficulty holding all of the elements associated with the expression, $a/b = c$, along with the procedure for its transformation into $a = cb$, in working memory. We are experts in this particular domain and understand the procedure. As a consequence, we can hold the material in working memory despite the large number of elements involved. The mechanism that permits this accomplishment is schema construction, discussed above. Once a schema has been constructed, the interacting elements are incorporated within the schema and do not need to be considered individually within working memory. The schema can act as a single element in working memory and will impose minimal working memory demands, especially if it is automated. For most readers of this paper, multiplying out a denominator is considered a single entity or element rather than multiple, interacting elements. Furthermore, once constructed, this schema can act as an interacting element in higher order schemas. More complex algebra involving the multiplying out of a denominator interacting with other elements can be held in working memory.

It follows from this discussion, that levels of element interactivity cannot be determined merely by analyzing the instructional material. A large number of interacting elements for one person may be a single element for someone with more expertise. Element interactivity can be determined only by counting the number of interacting elements with which people at a particular level of expertise are likely to deal. Knowing the probable characteristics of a potential target population of learners is essential when determining element interactivity (see Sweller and Chandler, 1994 for procedures used to estimate element interactivity). For the same reason, target group analysis should be integrated with knowledge analysis (hierarchical analysis of the material to be learned) when designing instruction, so that the knowledge can be communicated to the learners at the right grain size (van Merrienboer, 1997).

Intrinsic cognitive load through element interactivity is determined by an interaction between the nature of the material being learned and the expertise of the learners. It cannot be directly influenced by instructional designers although as discussed below, it certainly needs to be considered by designers. From a learner's perspective, intrinsic cognitive load is added directly to extraneous cognitive load. Extraneous cognitive load, discussed below, provides the core of cognitive load theory and is under the direct control of instructional designers. The distinction between extraneous cognitive load and germane cognitive load also is discussed below.

Extraneous Cognitive Load and Instructional Design

The design of practice and the organization and presentation of information is the domain of instructional designers. Although there are many factors that a designer may consider, the major thesis of this paper is that the cognitive load imposed by instructional designs should be the pre-eminent consideration when determining design structures. Limited working memory is one of the defining aspects of human cognitive architecture and, accordingly, all instructional designs should be analyzed from a cognitive load perspective. We argue that many commonly used instructional designs and procedures, because they were designed without reference to working memory limitations, are inadequate.

In this subsection, we simply summarize some instructional design considerations, with particular reference to designs that are inadequate because they impose a high extraneous load which is not relevant for learning. In the next subsection, we consider instructional designs that also increase cognitive load, but which nonetheless are optimal because they direct learners' attention to cognitive processes that are relevant for learning or schema construction (i.e., instructional designs that increase germane cognitive load). Subsequently, we will provide more details, including alternative designs that take working memory limitations into account. We also summarize the experimental evidence demonstrating the effectiveness of alternative instructional designs.

Consider a student who is learning a new topic in mathematics, science, or technology. Typically, the student is presented some new material, shown one or two worked examples, and then practices the procedures by solving many problems. For conscientious students, most time is likely spent and most learning likely occurs during the problem-solving phase. As a consequence, questions concerning the effectiveness of problem solving as a schema construction and automation device become pertinent.

Problem solving search places heavy demands on working memory (Sweller, 1988). The strategy most commonly used by people faced with novel problems for which they do not have previously constructed schemas—means–ends analysis—requires problem solvers to consider a current problem state (e.g., $a/b = c$), consider the goal state (e.g., $a = ?$), extract differences between the two states, and find a problem-solving operator (e.g., the rules of algebra such as subtracting an addend or multiplying out a denominator) that can be used to reduce or eliminate differences between a current problem state and the goal state. In addition, any subgoals that have been established need to be kept in mind. This particular problem-solving search strategy, although an efficient means of solving the problem, bears little relation to learning. In an educational context, learning is the primary goal of the exercise. Except in a test, attaining a problem goal is not directly relevant. Furthermore, not only is a means–ends, problem-solving search strategy distant from any schema construction goals, the strategy imposes a very heavy cognitive load that interferes with learning (Sweller, 1988). Under these circumstances, alternatives to conventional problem solving practice are needed. Several of these are discussed later in the section on instructional procedures.

Consider another common instructional design. Frequently, instruction includes multiple sources of information such as a combination of mutually referring diagrams and text. In order to understand the diagram or the text, it may be necessary to mentally integrate them. Such mental integration likely imposes a heavy, extraneous cognitive load (Sweller, Chandler, Tierney, and Cooper, 1990). The load is extraneous because it is caused entirely by the format of the instruction rather than by the intrinsic characteristics of the material. We discuss alternative designs that reduce extraneous cognitive load when discussing instructional procedures.

Intrinsic cognitive load due to element interactivity and extraneous cognitive load due to instructional design are additive. Whether extraneous cognitive load presents students with a problem depends, at least in part, on the intrinsic cognitive load (Sweller and Chandler, 1994). A combination of high intrinsic and high extraneous cognitive load may be fatal to learning because working memory may be substantially exceeded. Because intrinsic cognitive load cannot be altered, it may be essential to design instruction in a manner that reduces extraneous cognitive load. If, in contrast, the intrinsic cognitive load is low due to low element interactivity, a high extraneous cognitive load due to inadequate instructional procedures may be less harmful. The total cognitive load may be well within working memory limits. If, for example, students are required to mentally integrate diagrams and text in an environment where the intrinsic cognitive load is low, there may be no adverse implications because the total working memory load may not be high. However, in such cases one might also encourage learners to invest extra effort in processes that are directly relevant to learning, such as schema construction. This process also increases cognitive load, but it is germane cognitive load that contributes to, rather than interferes with, learning.

Germane Cognitive Load and Instructional Design

Until now, cognitive load theory research almost exclusively has studied instructional designs intended to decrease extraneous cognitive load. Recently, some studies have been conducted in which germane cognitive load was increased for processes considered to be directly relevant to schema construction. The basic assumption is that an instructional design that results in unused working memory capacity because of a low intrinsic cognitive load imposed by the instructional materials, low extraneous cognitive load due to appropriate instructional procedures, or a combination of both, may be further improved by encouraging learners to engage in conscious cognitive processing that is directly relevant to the construction of schemas. Obviously, this approach will only work if the total cognitive load associated with the instructional design, or the sum of intrinsic cognitive load plus extraneous cognitive load plus germane cognitive load, stays within working memory limits. The combination of decreasing extraneous cognitive load and at the same time increasing germane cognitive load involves redirecting attention: Learners' attention must be withdrawn from processes not relevant to learning and directed toward processes that are relevant to learning and, in particular, toward the construction and mindful abstraction of schemas (see van Merrienboer, 1997).

Consider, for instance, the following instructional design. Frequently, instruction includes a range of examples for problem solutions which students have to study before they start to work on problems. Studying those examples may not be an activity that generates an excessive cognitive load (see section on worked examples). Nonetheless, research has shown that students often skip those examples and only consult them when they are unable to solve the related problems (van Merrienboer and Paas, 1990; Pirolli and Anderson, 1985). The examples may not induce students to engage in cognitive processes that help them to construct schemas. Instructional procedures, such as asking questions about the examples or making examples incomplete so that students have to complete them, may then help students learn by increasing germane cognitive load. Some details of designs that not only decrease extraneous cognitive load but also increase germane cognitive load are discussed subsequently.

In summary, students face several sources of difficulty when presented with some new material. The material may consist of a large number of elements that need to be assimilated over a long period of time, perhaps years. If that material is low in element interactivity, it is easy to understand because individual elements are easy to learn. They can be learned serially without imposing a heavy working memory load. When dealing with this type of material, instruction designed to reduce cognitive load may not be an issue. Material that is high in element interactivity is hard to understand because understanding requires working memory to process many interacting elements simultaneously, rather than serially. Understanding may only occur fully once the interacting elements have been incorporated into a higher-order schema that can be held more easily in working memory. When dealing with high element interactivity material, because intrinsic cognitive load is high, it may be vital to reduce extraneous cognitive load in order to reduce total cognitive load to manageable proportions. If total cognitive load is not excessive, teachers might direct learners' attention to processes that are relevant to learning, or to the construction of schemas (increasing germane cognitive load) as well as redirect learners' attention from processes not directly relevant to learning (causing extraneous cognitive load). Appropriate instructional designs can reduce extraneous cognitive load and redirect learners' attention to cognitive processes that are directly relevant to the construction of schemas.

Cognitive load theory consists of the cognitive architecture outlined earlier together with the information structures and their instructional implications outlined in this section. The primary purpose of the theory has been to provide a framework for instructional design. Later, we will discuss some of the direct instructional consequences of the theory together with the validating empirical evidence. In order to assess that evidence, techniques for estimating comparative cognitive load are required. Those techniques are discussed in the next section.

MEASURING COMPARATIVE COGNITIVE LOAD

The Concept of Cognitive Load

Cognitive load is generally considered a construct representing the load that performing a particular task imposes on the cognitive system. It can be conceptualized as a task-based dimension (i.e., mental load) and a learner-based dimension (i.e., mental effort), both of which affect performance. *Mental load* refers to the load that is imposed by task (environmental) demands. These demands may pertain to task-intrinsic aspects, such as element interactivity, which are relatively immune to instructional manipulations and to task-extraneous aspects associated with instructional design. *Mental effort* refers to the amount of cognitive capacity or resources that is actually allocated to accommodate the task demands. *Performance* refers to the associated learner's performance. For a detailed theoretical account of the concept of cognitive load the reader may refer to Paas and van Merrienboer (1994b).

Cognitive Load and Mental Effort

The question of how to determine cognitive load is difficult for researchers, because of its multidimensional character and the complex interrelationships between performance, mental load, and mental effort. The complexity may be illustrated by the observation that, within the limits of their cognitive capacity, students can compensate for an increase in mental load (e.g., increasing task complexity) by investing more mental effort, thereby maintaining performance at a constant level. Consequently, the cognitive costs associated with a certain performance level cannot be consistently inferred from task- and performance-based measures. Instead, measures of mental effort can reveal important information about cognitive load which is not necessarily reflected in performance and mental-load measures. Based on these arguments, a combination of the intensity of mental effort being expended by learners and the level of performance attained by the learners, constitutes the best estimator of instructional efficiency.

Measuring Cognitive Load

Three major categories of mental-effort measurement techniques can be classified (Wierwille and Eggemeier, 1993). These include subjective, physiological, and task- and performance-based indices. Each category incorporates a number of individual assessment techniques.

Subjective techniques are based on the assumption that people are able to introspect on their cognitive processes and to report the amount of mental effort expended. Among others, Gopher and Braune (1984) found that people can introspect on their cognitive processes and have no difficulty in assigning numerical values to the imposed mental load or invested mental effort. Typically, these techniques use rating scales to report the experienced effort or the capacity expenditure. The results of empirical and theoretical studies on scaling suggest that the kind of scale used is not critical; the choice of uni- or multidimensional category scales, magnitude estimation, and the presence or absence of verbal labels makes no difference (e.g., Borg, 1978; Hendy, Hamilton, and Landry, 1993).

Physiological techniques are based on the assumption that changes in cognitive functioning are reflected in physiological measures. These techniques include measures of heart rate and heart rate variability, brain activity (e.g., brain evoked potentials), and eye activity (e.g., pupillary dilation, blink rate).

Task- and performance-based techniques include two subclasses of techniques: primary task measurement, which is based on learner performance of the task of interest, and secondary task methodology, which is based on performance when a second task is performed concurrently with the primary task. These techniques use objective task characteristics (e.g., number of elements that need to be considered such as the number of if–then conditions in a propositional reasoning task) and performance levels (e.g., differential learning times, errors) to obtain information on mental effort.

Whereas the central claim of cognitive load theory is that any instructional design should incorporate efficient use of working memory capacity, it is not common to measure cognitive load while conducting research on instruction. At least until 1992, instructional research in the context of cognitive load theory was exclusively concerned with performance- and task-based estimates of cognitive load. For example, using computational models and secondary tasks, Sweller (1988) provided evidence of a substantial reduction in cognitive load when using a goal-free as opposed to a conventional, means–ends strategy (see below). However, recently, in addition to such performance- and task-based estimates of cognitive load, subjective and physiological measurement techniques have been applied.

With respect to subjective techniques used in the context of cognitive load theory, Paas (1992) and Paas and van Merrienboer (1994a) have used a modified version of Bralfisch, Borg, and Dornic's (1972) rating scale for measuring perceived task difficulty. Learners had to report their invested mental effort on a unidimensional ninth-grade symmetrical category scale by translating the perceived amount of mental effort into a numerical value. The numerical values and labels assigned to the categories ranged from

"1" to "9," corresponding to "very, very low mental effort" to "very, very high mental effort." A comparable seventh-grade rating scale was used by Marcus, Cooper, and Sweller (1996), Kalyuga, Chandler, and Sweller (1998), and Tindall-Ford, Chandler, and Sweller (1997).

A physiological method for measuring the intensity of mental effort, applied in a study by Paas and van Merrienboer (1994a), subjected the variability in heart rate to spectral analysis. Spectral analysis is a mathematical method to investigate whether or not a signal contains periodical components. The time between successive heart beats seems to be determined by three different feedback mechanisms connected with respiration, blood pressure, and body-temperature regulation. Controlled processing is related to a specific cardiovascular state that manifests itself in the heart-rate variability power-spectrum band, which is related to blood pressure regulation, the so-called midfrequency band from 0.07 to 0.14 Hz. Intensity of effort is directly related to controlled processing, which in turn causes a change in this power spectrum. This blood pressure-related component has been found to decrease with increasing mental effort. Among others, Aasman, Mulder, and Mulder (1987) and Mulder (1988) have validated this technique with several cognitive tasks (e.g., multidimensional classification, sentence comprehension, and continuous working memory).

Comparisons between subjective and physiological measurement techniques have been carried out. On the basis of the empirical studies of Paas (1992) and Paas and van Merrienboer (1994a), the sensitivity, reliability, construct validity, and intrusiveness of the subjective and physiological measurement techniques were evaluated by Paas, van Merrienboer, and Adam (1994). This evaluation showed that the subjective rating scale was sensitive to relatively small differences in cognitive load, and that it was valid, reliable, and nonintrusive. The psychophysiological measure, based on spectral-analysis of heart-rate variability, turned out to be nonintrusive but unreliable, invalid, and only sensitive to relatively large differences in cognitive load. It was concluded that subjective rating-scale measurement is the most promising technique for research in the context of cognitive load theory.

Instructional Efficiency

Although, at present, absolute values that indicate acceptable vs. unacceptable levels of cognitive load are not available, relative measures are available that compare instructional techniques. A recently developed technique measures instructional efficiency by combining cognitive load with performance measures. To obtain information on the relative mental efficiency of instructional conditions, Paas and van Merrienboer (1993) developed a computational method which combines the intensity of mental effort being expended by learners with the level of performance attained. This efficiency approach is based on the conversion of raw mental effort data and raw performance data to *z*-scores, which can be displayed in a M(ental effort)-P(erformance) cross of axes. The combined effects on mental effort and performance of experimental instructional conditions can be deduced from the relative position of points on the display. As can be seen in Fig. 1, this procedure makes the method very useful for visualizing the differential mental efficiency of instructional conditions.

Successful applications of the method can be found in the studies by Paas and van Merrienboer (1994a), Marcus, Cooper, and Sweller (1996), Kalyuga, Chandler, and Sweller (1998), and Tindall-Ford, Chandler, and

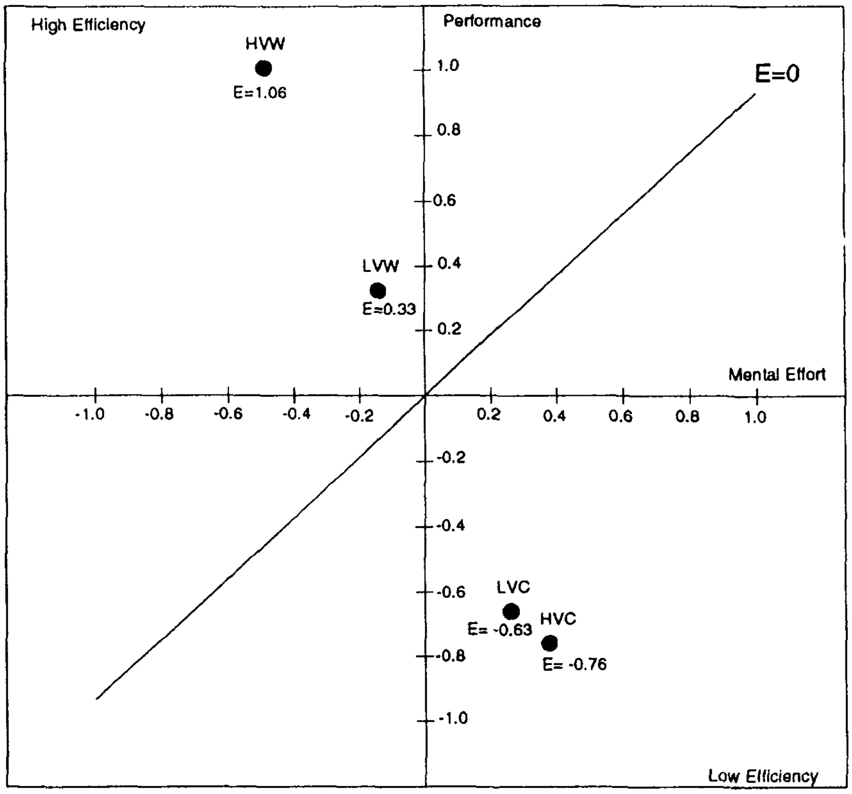


Fig. 1. Cross of axes for instructional efficiency.

Sweller (1997). The efficiency data depicted in Fig. 1 were obtained by Paas and van Merrienboer (1994a) in the domain of geometrical problem solving. They compared a high- and low-variability conventional problem condition to a high- and low-variability worked example condition. It was hypothesized and found that instruction with worked examples leads to superior schema construction and higher efficiency compared to instruction with conventional examples.

SOME INSTRUCTIONAL PROCEDURES

Cognitive load theory has been used to generate instructional techniques for the last 15 years with the work carried out in several centers around the globe. The theory has continued to develop during those years with the earlier work relying on little more than knowledge of a limited working memory and a large long-term memory. Two points concerning the use of the theory for instructional design need to be noted. First, as indicated previously, most of the earlier work did not attempt to directly and independently measure comparative cognitive load. Rather, the theory was used to generate instructional techniques that were tested experimentally. Successful tests were assumed to strengthen the theory that generated the techniques. As suggested, the only relatively direct evidence for cognitive load came from differential learning times. Direct measures of cognitive load were used in subsequent work. Second, all techniques discussed below have been supported by multiple, overlapping experiments using a variety of materials and a variety of populations. As a consequence, we believe there are some grounds for optimism concerning the strength and stability of the effects studied. We begin with the goal-free effect, the first cognitive load effect described.

The Goal-Free Effect

This effect comes under a variety of names including the no-goal effect and the reduced goal-specificity effect. Consider a physics student practicing the following problem while studying kinematics:

A car is uniformly accelerated from rest for 1 min. Its final velocity is 2 km/min. How far has it traveled?

A student faced with this problem is unlikely to have a schema to generate a solution, and so is likely to use problem solving search through means–ends analysis. That process requires the student to consider the goal, distance traveled; consider the givens, uniform acceleration, time traveled, and final velocity; consider differences between the givens and the goal; and find problem-solving operators to reduce those differences. To accomplish the process, relevant kinematics equations that can be used to provide a chain of equations bridging the goal and the givens must be found. Assume that two of the relevant equations learned by the student at this point are *average velocity* = .5 *final velocity* and *distance* = *average velocity* × *time*. The equation, *distance* = *average velocity* × *time*, includes the goal variable, distance. If values for the variables, *average velocity* and *time,* can be found, the problem is solved. Time is known, but average velocity is not. Average velocity must be set as a subgoal. The equation, *average velocity* = .5 *final velocity* contains the subgoal variable, average velocity. If final velocity can be found, average velocity can be calculated. Final velocity is a given. The student has worked backward from the goal constructing a series of equations connecting the goal to the givens. He or she now can work forward calculating values for the unknown variables. This process provides a value for the goal variable and a solution to the problem.

Problem-solving search through means–ends analysis is an efficient way of attaining a problem goal in the absence of a schema. Nevertheless, as indicated above, it is a process that is exceptionally expensive of working memory capacity (Sweller, 1988), and it bears little relation to schema construction processes that are concerned with learning to recognize problem states and their associated moves. In this context, learning and problem solving are different and incompatible processes. Solving conventionally structured problems may be an inefficient technique for constructing the schemas associated with expertise. A heavy, extraneous cognitive load is imposed.

Goal-free problems have been devised (Sweller and Levine, 1982 using puzzle problems) to alter learner activities in a manner that reduces the extraneous cognitive load caused by means–ends analysis and to encourage schema construction. A goal-free version of the conventional kinematics problem above is identical except that the last statement is replaced by the statement, *Calculate the value of as many variables as you can.* The function of this statement is to eliminate means–ends search and its attendant cognitive load. Goal-free problems do not permit problem solvers to extract differences between a current problem state and the goal state because no goal state is specified, short-circuiting the entire means–ends process. In order to solve goal-free problems, problem solvers must find an alternative strategy to means–ends analysis. That strategy cannot rely on a goal specified as a problem state. The most obvious strategy is to consider each problem state encountered and find any problem-solving operator that can be applied. Once an operator has been applied, a new problem state has been generated and the process can be repeated.

Consider this goal-free strategy as applied to the goal-free version of the above kinematics problem:

A car is uniformly accelerated from rest for 1 min. Its final velocity is 2 km/min. Calculate the value of as many variables as you can.

In order to solve this problem, a problem solver must find any operator, in this case an equation, that can be applied to the givens. The givens indicate that time =1 min and final velocity = 2 km/min. Assume that the equation, *average velocity = .5 final velocity,* is the only equation known by the problem solver that can be immediately applied to the givens. Once applied, a new problem state exists and the problem solver again must find an equation that can be applied. In this case, the equation, *distance = average velocity* × *time,* can be applied, providing a value for distance.

Several points concerning this process need to be noted. In following a goal-free strategy, the problem solver has ended up calculating exactly the same values using the same equations as a problem solver using means–ends analysis. Nevertheless, although the end result is identical, the problem-solving strategy used and the cognitive processes that occur are quite different. With the presence of a goal and the use of means–ends analysis, a problem solver must continually hold and process in working memory, the current problem state, the goal state, relations between them, problem-solving operators that could reduce differences, and any subgoals. In contrast, a goal-free strategy requires nothing more than each problem state and any operator that can be applied to that state. Furthermore, it is precisely this combination that is required for schema construction. Sweller (1988), using computational models and secondary tasks, provided evidence of a substantial reduction in cognitive load when using a goal-free as opposed to a conventional, means–ends, strategy.

Many experiments have demonstrated the effectiveness of goal-free problems as an instructional design tool. Sweller, Mawer, and Ward (1983) ran several experiments using kinematics and geometry problems with secondary students. The kinematics problems were similar to the one used as an example above. The geometry problems used theorems such as *vertically opposite angles are equal* and *the external angles of a triangle equal the sum of the opposite internal angles.* Conventional geometry problems required students to find a value for a particular angle in a diagram, whereas goal-free problems asked students to find the values of as many angles as they could. The general procedure was to provide a conventional group with relevant instruction in kinematics or geometry, followed by an acquisition phase involving practice at solving conventional problems. An identical procedure was followed by goal-free groups except that the practice session used goal-free rather than conventional problems. Common tests using conventional problems were then used to assess learning. Results consistently indicated that goal-free groups were superior in terms of schema construction. Similar results were obtained by Owen and Sweller (1985) in the domain of trigonometry. Although direct measures of cognitive load were not taken, large differences in speed of performance on the acquisition problems, favoring the goal-free groups, were used to infer differences in cognitive load. Bobis, Sweller, and Cooper (1994) found that primary school students taught geometrical paper-folding tasks were superior after practice on tasks in which the final model that was the goal of the task was absent. Model presence or absence in this task was considered analogous to goal presence or absence. Ayres (1993) found that on two-step geometry tasks with conventional problems, most errors occurred during the subgoal rather than the goal phase. Working memory load is highest at the subgoal phase because more elements must be considered at this phase than at the goal phase. In contrast, fewer errors were made by students practicing on goal-free problems with the reduction due to a reduction during the "subgoal phase." (There are, of course, no subgoals on goal-free problems.) Vollmeyer, Burns, and Holyoak (1996) used biology-based problems to demonstrate that learning was retarded when tertiary students solved problems using means–ends analysis compared to goal-free strategies.

Evidence for the effectiveness of goal-free problems is strong, with the effect obtained under a very wide variety of conditions. We believe there are cogent grounds for encouraging instructional designers to consider including goal-free problems in their repertoire of techniques when dealing with areas such as mathematics and science in which practice at solving problems is an important instructional procedure.

Worked Example Effect

The goal-free effect occurs because goal-free problems reduce extraneous cognitive load and facilitate schema construction in comparison to solving problems by mean–ends analysis. Studying worked examples also eliminates means–ends search, and so a heavy use of worked examples as a substitute for solving problems may be also beneficial. In contrast to conventional problems, worked examples focus attention on problem states and associated operators (i.e., solution steps), enabling learners to induce generalized solutions or schemas. As nothing else has to be attended to, extraneous cognitive load should be low. This reasoning leads to the counterintuitive prediction that studying worked examples may facilitate schema construction and transfer performance more than actually solving the equivalent problems (Jelsma, van Merrienboer, and Bijlstra, and 1990). In addition, it may be desirable to identify the critical features in the worked examples by annotating them with what they are supposed to illustrate (e.g., Anderson, Boyle, Corbett, and Lewis, 1990).

Support for the prediction that worked examples facilitate learning and problem solving more than solving the equivalent problems has been found in several subject-matter domains. Sweller and Cooper (1985) and Cooper and Sweller (1987) studied the use of worked examples as a substitute for conventional problem solving in learning algebra. In their studies, the use of worked examples improved schema construction and the ability to solve new algebra problems more than conventional problem solving. As another, impressive illustration, Zhu and Simon (1987) found in a series of long-term studies that worked examples could replace conventional classroom teaching. In one study, they found that a 3-year mathematics course was completed in 2 years by emphasizing worked examples.

Paas (1992) performed a study in the domain of statistics, in which students who solved conventional problems had the opportunity to study a worked example for a particular problem if they failed to find a solution to this problem. Despite the opportunity to study the worked examples, the conventional condition yielded inferior schema construction and lower transfer performance by students than did the worked example condition. A possible explanation for this finding is that inferior transfer performance in the conventional condition was caused by the incorporation of the failed solutions in the constructed schemas. Paas and van Merrienboer (1994a) performed an experiment in geometry problem solving in which, in contrast to previous studies in which the worked condition also contained a number of conventional problems, a pure worked condition in which students only had to study worked examples was used. Also in this study, the worked examples yielded lower extraneous cognitive load scores, better schema construction, and higher transfer performance than the conventional condition. In addition, the instructional efficiency of the worked examples condition proved to be superior to the conventional condition.

In other recent work directly testing the effects of studying worked examples as opposed to using means–ends search to solve problems, Trafton and Reiser (1993) found that college students learning aspects of the LISP programming language benefitted more after studying worked examples than after solving the equivalent problems. Similarly, Carroll (1994), using remedial mathematics students, found a superiority of worked examples over solving problems that required English expressions to be translated into algebraic equations.

The general conclusion from these studies is that a far heavier than usual use of worked examples is beneficial to learning outcomes and transfer. This conclusion particularly applies because learners often view worked examples, rather than explanatory texts, as the primary and most natural source of learning material (e.g., Lieberman, 1986; Pirolli, 1991; Segal and Ahmad, 1993). This is not to say that there are no disadvantages to the use of worked examples. A lack of training with genuine problem-solving tasks may have negative effects on learners' motivation. A heavy use of worked examples can provide learners with stereotyped solution patterns that may inhibit the generation of new, creative solutions to problems (Smith, Ward, and Schumacher, 1993). For this reason, goal-free problems and completion problems, which are discussed in the next section, may offer a good alternative to an excessive use of worked examples.

The design of good worked examples is also difficult. For instance, worked examples that require learners to integrate different sources of information (e.g., text and diagrams) are often not effective because they yield a high extraneous cognitive load (Sweller, Chandler, Tierney, and Cooper, 1990; Tarmizi and Sweller, 1988; Ward and Sweller, 1990). The same may be true for worked examples that convey redundant information or a low-variability sequence of worked examples that does not allow for an effective construction of schemas (Paas and van Merrienboer, 1994a). These constraints to the design and use of worked examples are also further discussed below.

Completion Problem Effect

One major disadvantage of worked examples is that they do not force learners to carefully study them. For this reason, worked examples are typically intermixed with conventional problems. Nevertheless, learners may only briefly look at the worked examples before they start to work on the conventional problems. Whereas higher-ability learners tend to fully process and even elaborate worked examples, lower-ability learners may only carefully study the worked examples when they encounter problems in solving the conventional problems (Chi, Bassok, Lewis, Reimann, and Glaser, 1989; Lefevre and Dixon, 1986). Consulting worked examples at the same time as attempting to solve a problem requires both the worked example and the problem to be simultaneously processed in working memory, resulting in a possible working-memory overload. As an alternative, van Merrienboer and Krammer (1987, 1990) suggested the use of completion problems. Completion problems are problems for which a given state, a goal state, and a partial solution are provided to learners who must complete the partial solution. Completion problems seem to be especially useful in design-oriented subject-matter domains such as software design, design of electronic circuits, planning production processes, Computer Numerically Controlled (CNC) programming, and architecture. Completion problems also provide a bridge between worked examples and conventional problems. Worked examples are completion problems with a full solution, and conventional problems are completion problems with no solution. Particular instructional strategies may start with completion problems that provide almost complete solutions, and gradually work to completion problems for which all or most of the solution must be generated by the learners. Such a strategy is known as the "completion strategy."

Completion problems combine the strong points of worked examples and conventional problems. Like worked examples, they typically decrease extraneous cognitive load. Although learners are not explicitly induced to study worked examples, they must carefully study the partial worked examples provided in completion problems because they otherwise will not be able to complete the solution correctly. Paas (1992) compared the effects of conventional problems, worked examples, and completion problems on training performance, transfer performance, and cognitive load in statistical problem solving. He found that a cognitive structure resulting from instruction emphasizing practice with worked examples or completion problems yielded a more efficient knowledge base for solving transfer problems than one resulting from instruction emphasizing conventional problems. Training with worked examples or completion problems led to less effort-demanding and higher transfer performance than training with conventional problems. In this short study, no significant differences between effects of worked examples and completion problems were found.

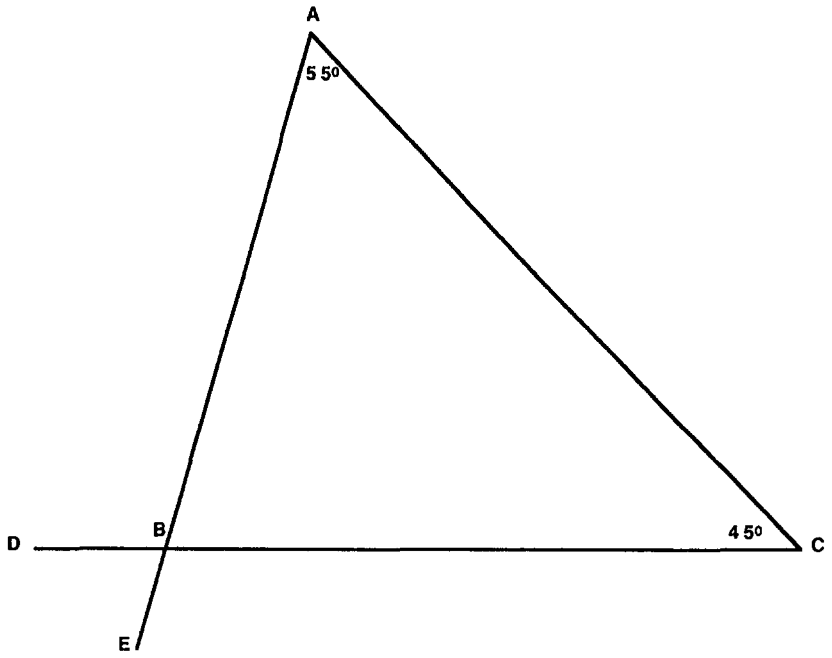
Van Merrienboer (1990) studied the use of completion problems in a course of longer duration than that of Paas (1992). In a ten-lesson introductory computer-programming course, the differential effects on learning outcomes and transfer were studied when comparing an instructional strategy based on completion problems (i.e., the modification and extension of existing computer-programs) and a strategy based on conventional problems (i.e., the design and coding of new computer programs). After the course, the completion group was clearly superior on measures concerning the construction of new programs. In addition, a detailed analysis of the data revealed that the completion assignments facilitated the use of programming templates (i.e., stereotyped patterns of programming code). This finding provides strong support for a better construction of schemas as a result of the use of completion problems. The results were replicated in a subsequent study using computer-assisted instruction for introductory computer programming (van Merrienboer and de Croock, 1992). In addition, log files indicated that students who were working on conventional problems frequently had to search for useful examples while they were performing their program-generation tasks. In contrast, the information needed to perform the program completion tasks appeared to be largely available in the to-be-completed programs, decreasing extraneous cognitive load and enhancing learning.

Summarizing the results of studies on completion problems (see van Merrienboer, 1992), there is considerable evidence that, compared to conventional problems, they decrease extraneous cognitive load, facilitate the construction of schemas, and lead to better transfer performance. In short duration studies, results indicated that completion problems are equally effective as worked examples intermixed with conventional problems. In studies of a longer duration, completion problems may better help learners to maintain motivation and focus their attention on useful solution steps that are available in the partial examples.

One drawback of completion problems is that they can be time-consuming to construct. An instructional designer must consider which part of the solution is presented to the learners, or from the opposite perspective, which part is left for learners to complete. Although a good completion problem typically requires that the learners (a) must understand the partial solution before they are able to complete it, and (b) have to perform a nontrivial completion, this still leaves the instructional designer with both considerable freedom and, consequently, a considerable number of decisions.

Split-Attention Effect

While the completion problem effect is closely related to the worked example effect, the split-attention effect was derived directly from the worked-example effect. Although the worked-example effect can be obtained on a considerable range of problems, on some categories of problems, it is more difficult to obtain than on others. The worked-example effect occurs because worked examples reduce extraneous cognitive load in comparison to solving problems by means–ends analysis. There can be no guarantee that all worked examples appreciably reduce cognitive load compared to means–ends search. Consider a conventionally structured geometry worked example consisting of a diagram and its associated solution statements (see Fig. 2). The diagram alone reveals nothing about the solution to the problem. The statements, in turn, are unintelligible until they have been integrated with the diagram. In order to understand the two sources of information, the diagram and statements, learners must mentally integrate them. For example, in order to derive any meaning from a statement, learners must read the statement, hold it in working memory, and then search the diagram for the appropriate referents. This process can be cognitively demanding, and there is no *a priori* guarantee that it is less



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

Solution:

Angle ABC = 180^0 - Angle BAC - Angle BCA (Internal angles of a triangle sum to 180^0)
= 180^0 - 55^0 - 45^0
= 80^0

Angle DBE = Angle ABC (Vertically opposite angles are equal)
= 80^0

Fig. 2. Example demonstrating split attention.

demanding than means–ends search. Furthermore, the demands imposed by such split-attention formats constitute an extraneous cognitive load that occurs purely because of the particular format conventionally used to present this area of the curriculum.

An alternative to the split-attention format of Fig. 2 is an integrated format as exemplified in Fig. 3. This format obviates the need to search for relations between the diagram and statements. Instead of having to use working memory resources to mentally integrate the two sources of infor-

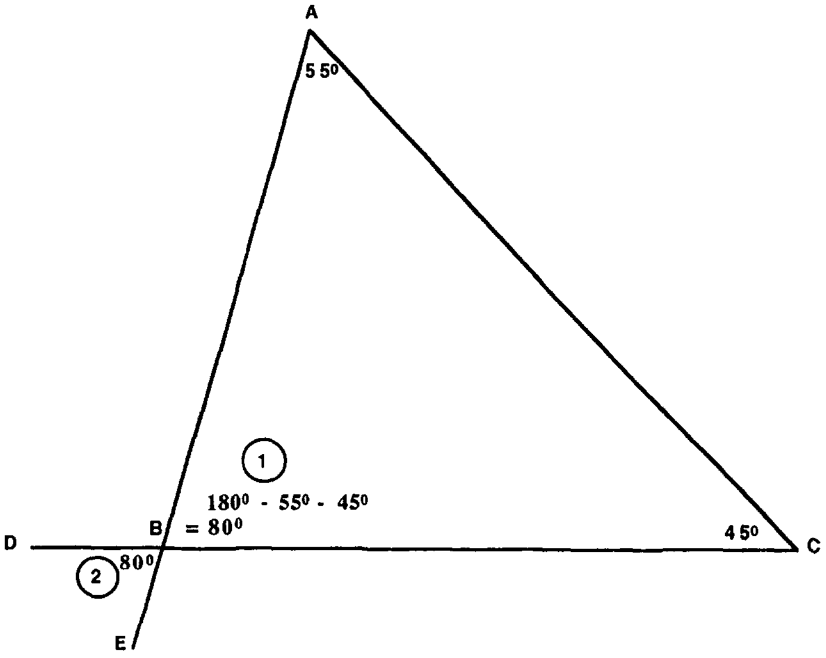


Fig. 3. Integrated example with no split attention.

mation, physical integration is used. The information in Figs. 2 and 3 is identical. Only the instructional designs differ, with Fig. 3 intended to reduce extraneous cognitive load by physically integrating disparate sources of information and so reducing the need for mental integration. Training conditions using conventional split-attention formats similar to Fig. 2, when compared to the integrated format of Fig. 3, can be predicted to yield results demonstrating the superiority of the integrated formal. This result is known as the split-attention effect.

Many experiments using both worked examples and other forms of instruction provide evidence for the split-attention effect. Tarmizi and Sweller (1988) failed to find the worked-example effect using conventional geometry examples. The effect could be obtained only by the use of integrated examples that proved superior to both split-attention examples—demonstrating the split-attention effect—and conventional problems, demonstrating the worked-example effect. Ward and Sweller (1990) obtained similar results to Tarmizi and Sweller using kinematics problems. Figure 4 provides examples of conventional and integrated kinematics worked examples.

a. Conventional Kinematics Worked Example

A vehicle moving from rest reaches a speed of 40 m/s after 20 seconds. What is the acceleration of the vehicle?

$u = 0$ m/s
$v = 40$ m/s
$t = 20$ s

$v = u + at$
$a = (v - u)/t$
$a = (40 - 0)/20$
$a = 2$ m/s^2

b.Integrated Kinematics Worked Example

A vehicle moving from rest (u) reaches a speed of 40 m/s (v) after 20 seconds (t): [$v = u + at$, $a = (v - u)/t = (40 - 0)/20 = 2$ m/s^2]. What is the acceleration of the vehicle?

Fig. 4. Examples of conventional and integrated kinematics worked examples.

Sweller, Chandler, Tierney, and Cooper (1990) demonstrated the split-attention effect using both worked examples and more general forms of instruction in coordinate geometry and numerical control programming. Bobis, Sweller, and Cooper (1993) used a primary school paper-folding task, whereas Chandler and Sweller (1992) looked at split-attention effects that occur in the writing of scientific reports. Mwangi and Sweller (in press) used primary school arithmetic word problem-solving tasks and in addition, using student self-explanations, found that a split-attention format interfered with inferential processes required to understand the problems. Mayer and Anderson (1991, 1992) found that animation and associated narration needed to be temporally coordinated, a result that probably mimics the spatial coordination investigated in the other work reported here.

Lastly, Sweller and Chandler (1994) and Chandler and Sweller (1996) demonstrated that students learning a computer application learned better if all of the material was presented in an appropriately structured manual without a computer present on which to practice, as opposed to having a manual and computer on which to work. It was argued that the conventional technique of having both a computer screen and a manual resulted in split-attention. Of considerable importance, the split-attention effect was obtained only when high element interactivity material was used, providing the first evidence of the importance of intrinsic, as well as extraneous, cognitive load. When intrinsic cognitive load was low, working memory was not overloaded and the extraneous cognitive load imposed by the split-attention instructional design did not matter. It only mattered when element interactivity and hence, intrinsic cognitive load, were high. These results do not indicate that computer applications must be learned without access to a computer. Cerpa, Chandler, and Sweller (1996) found that instruction was just as effective if all of the information was placed on a computer screen and presented in computer-assisted instruction mode. It is split-attention between the screen and manual that creates a problem, not the use of a computer screen *per se*. The problem can be overcome by integrated instruction. It is irrelevant whether that instruction is placed entirely in a manual or entirely on the screen. The only important point is that learners should not have to use scarce working memory resources to search for information on a screen that is referred to in a manual, or *vice versa*.

Split-attention occurs very commonly in instructional contexts. On the basis of dozens of experiments under a wide variety of conditions, the evidence suggests overwhelmingly that it has negative consequences and should be eliminated wherever possible. Sometimes, of course, it is not possible to provide integrated conditions. Nevertheless, frequently it is possible, and instructional designers should seriously consider the adverse consequences of unnecessarily requiring learners to search for information.

Modality Effects

The effects discussed to this point all rely on instructional designs and procedures intended to reduce an unnecessary load on working memory. It was assumed that although the amount of information that could be processed in working memory could be altered by acquiring more sophisticated schemas, working memory capacity was fixed in the sense that the number of schemas, elements or chunks that could be dealt with was unalterable. Although generally valid, there is one set of conditions under which this assumption may not be entirely true. Working memory can be subdivided into partially independent components, streams, or processors. Major theories such as that of Baddeley (1992) include working memory subcomponents consisting or a "phonological loop" to deal with verbal material based on an auditory working memory and a "visual-spatial scratch pad" to deal with diagrammatic/pictorial information and based on a visual working memory. Furthermore, as Penney (1989) indicated in a detailed review of experimental psychological literature, effective working memory capacity can be increased by using both visual and auditory working memory rather than either memory stream alone. Although less than purely additive, there seems to be an appreciable increase in capacity available by the use of both, rather than a single, processor.

These theoretical and empirical considerations have implications for the split-attention effect and for instructional designs intended to ameliorate its negative consequences. Split-attention occurs when two or more sources of information must be processed simultaneously in order to derive meaning from material. The working memory load imposed by the need to mentally integrate the disparate sources of information interferes with learning. In the previous section, we discussed reducing this problem by physical integration of the disparate sources of information. If effective working memory can be increased by using dual-modality presentation techniques (i.e., presentation techniques in which some of the material is presented in auditory and some of the material in visual form rather than having all of the material presented in visual form), then this procedure may be just as effective in facilitating learning as physically integrating two sources of visually-based information. The modality effect derives from the split-attention effect. It occurs under split-attention conditions when a written source of information, that must be integrated with another source of visually presented information such as a diagram, is presented in auditory rather than visual (written) mode.

Mousavi, Low, and Sweller (1995) tested for the modality effect using geometry instruction. Consider again the information in Fig. 2. The diagram obviously must be presented in visual form, but the textual information could be presented in either written (and hence visual) or auditory form. A visually presented diagram and auditorily presented text may increase effective working memory over conditions where visual working memory alone must be used to process all of the information. As a consequence, learning may be enhanced. In a series of experiments, Mousavi *et al.* obtained this result. Audio/visual instructions were consistently superior to visual/visual instructions, demonstrating the modality effect. Tindall-Ford, Chandler, and Sweller (1997) replicated this finding in another series of experiments using electrical engineering instructional materials. In addition, these experiments differentiated between material that was low or high in element interactivity. It was predicted that low element interactivity material with its low intrinsic cognitive load would not demonstrate the modality effect because increasing effective working memory would be irrelevant under conditions where the information that had to be processed did not strain working-memory capacity. The modality effect was obtained only when using high element interactivity materials. Lastly, Tindall-Ford *et al.* used subjective ratings to assess comparative cognitive load. They obtained strong evidence that the cognitive load was higher under visual/visual than under audio/visual conditions but only when the material was high in element interactivity. Instructional efficiency calculations indicated that audio/visual conditions were substantially more efficient than visual/visual conditions under high element interactivity conditions. There were no differences under low element interactivity conditions.

The findings on the modality effect have both theoretical and practical implications. From a theoretical perspective, the results provide further evidence that to some extent, effective working memory may be increased and this increase can be used to reduce cognitive load and facilitate learning. From a practical perspective, the results provide a new instructional technique. Under split-attention conditions, rather than physically integrating disparate sources of information, learning may be facilitated by presenting a written source of information in auditory mode. This effect may be especially important in areas such as the use of multimedia. Multimedia instruction is becoming increasingly popular but much of the work is a-theoretical. Findings associated with the modality effect can provide a coherent theoretical base for further multimedia investigations and applications.

Redundancy Effect

As was the case for the modality effect, the redundancy effect grew from and was associated with work on the split-attention effect, although it subsequently turned out that the redundancy effect had a much longer history than cognitive load theory. Split-attention occurs when learners are faced with multiple sources of information that must be integrated before they can be understood. The individual sources of information cannot be used by learners if considered in isolation, hence the need for integration. The redundancy effect occurs when multiple sources of information are self-contained and can be used without reference to each other. Frequently, the same material is presented on two or more occasions, in different forms. Again, a diagram and text can be used to provide an example. In the case of split-attention, the student can learn little or nothing by referring to the diagram or text in isolation without integrating them. In the case of redundancy, either the diagram or text, or frequently both, are fully intelligible in isolation and, indeed, may provide all of the information required by the learner. Chandler and Sweller (1991) provided an example of a diagram demonstrating the flow of blood in the heart, lungs, and rest of the body together with statements indicating, for example, "Blood from the lungs flows into the left atrium." The diagram included arrows demonstrating that blood flowed from the lungs into the left atrium and was readily intelligible in isolation. Because the diagram could be understood in isolation, the statements were redundant.

The distinction between multiple sources of information that can or cannot be understood in isolation is critical from an instructional design perspective. When dealing with multiple sources of information that cannot be understood in isolation, cognitive load can be reduced by physical integration to reduce split-attention. When dealing with multiple sources of information that can be understood in isolation, integration can increase, rather than decrease, cognitive load. Integrated information is very hard to ignore. If students find it difficult to ignore integrated material that they do not need to process because it is redundant, cognitive load is increased. It is easier to ignore nonintegrated than integrated material. It is easier again not to even have to consider redundant information. Chandler and Sweller (1991) using both electrical engineering and biology instructional material found that when dealing with redundancy, the best instructional design is one that eliminates redundancy or at the very least, allows learners to ignore the redundant material because it is separated from more relevant information. The redundancy effect occurs when students who are not presented with redundant information perform better on tests than students who are presented with redundant information.

There have been many experiments dealing with the redundancy effect since Chandler and Sweller (1991). Bobis, Sweller, and Cooper (1993) found that redundant verbal information associated with diagrams designed to teach primary school students geometrical paper-folding tasks, interfered with learning. Sweller and Chandler (1994) found that the presence of a computer on which students could work was redundant and interfered with learning from an instructional manual compared to the manual alone. This result was obtained using high- but not low-element interactivity material, demonstrating again that extraneous cognitive load may not be a problem if intrinsic cognitive load is low. Chandler and Sweller (1996) replicated this result and, using secondary task analyses, provided evidence that redundancy increased cognitive load. Cerpa, Chandler, and Sweller (1996) placed the computer manual information on screen using computer-assisted learning and found that, in this case, the manual was redundant. Mayer, Bove, Bryman, Mars, and Tapangco (1996) found that students presented a pictorial summary of a scientific process performed better than students presented the summary along with the full text or the full text alone.

Recent studies provided evidence of the close relations between redundancy and split-attention. Kalyuga, Chandler, and Sweller (1998) provided novice electrical apprentices with a wiring diagram to which a textual description of the same diagram had been added. The textual description merely re-described the circuit diagram and so might be expected to be redundant. In fact, for the novices, it was found that the textual material was essential. They could not understand the diagram alone and required the text. As a result, the two sources of information were best presented in integrated form to reduce split-attention. In contrast, once the learners gained additional experience resulting in a more generalized knowledge of circuit diagrams, textual material re-describing a novel circuit was redundant and it was better to fully eliminate the textual material for expert learners rather than integrate it with the diagrams. The additional experience allowed students to learn best from a diagram alone, rather than a diagram plus integrated text. Similar results were obtained by Yeung, Jin, and Sweller (1998) using instructional material that assisted students in text comprehension. These results demonstrated that material that is redundant for some learners and so best eliminated, may be essential for less experienced learners and best integrated. Ideal instructional designs may be heavily dependent on instructors accurately assessing their students' levels of expertise.

All of the above work was generated within a cognitive load framework. In fact, inspection of the literature reveals that over many decades, a considerable number of results have been obtained that can be interpreted as examples of the redundancy effect. Miller (1937) found that young children learning to read nouns made more progress if the words were presented alone rather than in conjunction with the relevant pictures. The use of pictures still occurs commonly today. This picture/word result has been replicated many times over the last few decades (e.g., Solman, Singh, and Kehoe, 1992). In other findings, Reder and Anderson (1980, 1982) ran many experiments indicating that the contents of physics texts were learned more effectively if students were merely presented with a summary rather than with the entire contents. Another example of the redundancy effect comes from Schooler and Engstler-Schooler (1990). They found that if learners had to verbalize a visual stimulus, subsequent recognition performance was impaired. The requirement to verbalize was redundant and so imposed an extraneous cognitive load that interfered with learning. Work on minimal computer manuals also can be interpreted within a redundancy framework. This work has consistently found that less information can be more effective than more information (e.g., Carrol, 1990; Lazonder and van der Meij, 1993).

Lastly, McNamara, Kintsch, Singer, and Kintsch (1996), in work that can be closely related to that of Kalyuga, Chandler, and Sweller (1998) and Yeung, Jin, and Sweller (1998), described above, found that additions to biology textual material intended to increase coherence benefitted low-knowledge readers but impeded high-knowledge readers. The authors argued that minimally coherent text without additions forced high-knowledge readers to engage in active processing that facilitated learning. Providing high-knowledge readers with additional material reduced active processing and so reduced learning. In fact, the additional material may have been redundant to high-knowledge readers, increasing extraneous cognitive load that interfered with learning. The mental effort ratings taken by Kalyuga *et al.* (1998) and Yeung *et al.* (1998) indicated a higher cognitive load when more experienced learners were presented with material containing redundancy. If additional material reduced active processing, ratings of mental effort should have decreased, not increased. These results suggest that in this context, a redundancy/cognitive load explanation may be more plausible than an active processing argument.

The redundancy effect has been discovered, forgotten, and rediscovered over many decades. We believe there are at least two reasons for this history. First, the effect is counterintuitive. Most people intuitively believe that at worst, redundant materials might have neutral effects and that there is every possibility that they could be beneficial. It can be difficult to accept that redundancy can have substantial, negative consequences. Second, the effect has never been placed within a detailed theoretical context. As a consequence, it has been easy to dismiss individual results as aberrations that can be ignored. We believe the redundancy effect is readily placed within a cognitive load framework and is related to other cognitive load effects. In this context, we hope the effect continues to influence the field.

Redundancy is a major effect that should be considered seriously by instructional designers. A large range of experimental results indicate the negative consequences of including redundant material when designing instruction. We know of no experimental work demonstrating advantages of redundancy, and we suspect that such a result only could be obtained under conditions where one set of instructional materials was so poor that any redundant alternative would inevitably confer benefits.

Variability Effect

It has been well documented that variability of practice may result in beneficial effects on transfer of training (e.g., see Cormier and Hagman, 1987; Jelsma, van Merrienboer, and Bijlstra, 1990; Singley and Anderson, 1989). Variability over problem situations is expected to encourage learners to develop schemas, because it increases the probability that similar features can be identified and that relevant features can be distinguished from irrelevant ones. Instructional designs that apply variability of practice ensure that a task is practiced under conditions that require the performance of different variants of the task over problem situations, or under 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, the familiarity of the task, and so forth.

Quilici and Mayer (1996) provided a recent example of the advantages of variability of practice. They compared the effects of low and high variability on learning to solve statistics word problems. After instruction, students had to place problems into categories requiring either a *t*-test, correlation, or chi-square for solution. Variability of practice was realized by distinguishing between structure-emphasizing and surface-emphasizing example sets. In the high variability, structure-emphasizing set, different surface stories were used for each example of a particular test (e.g., three different surface stories with each problem requiring a *t*-test). This set of surface stories was also used for the correlation and chi-square test types, with of course, an appropriate structure for the relevant test. In the low variability surface-emphasizing set, the same surface story was used for each example of a particular test, but it was different from those used for the other tests. As expected, students confronted with the high variability, structure-emphasizing problem sets were better able to categorize statistics word problems following the instruction.

Consistent findings were reported by Jelsma and Bijlstra (1988) and Jelsma and van Merrienboer (1989). They compared the effects of low and high variability on learning and transfer performance in the training of a troubleshooting task. Students had to diagnose particular malfunctions that appeared in a computer-based simulation of a water-alcohol distillery plant (Jelsma and Bijlstra, 1990). In these studies, variability of practice was realized by placing the problems that students had to solve in a particular order. For high variability, problems that required different diagnoses were presented in a random order. Such a random practice schedule is also referred to as practice with high contextual interference, because the skills performed for one particular problem interfere with the skills performed for adjacent problems. For low variability, sets of problems that require similar solutions are presented in a blocked order. Using this schedule, students see all of the problems of a particular type before seeing all of the problems of the next type.

The results of studies on variability initially seemed to contradict cognitive load theory. High variability increased cognitive load during practice but yielded better transfer of learning, indicated by a better ability to diagnose faults that were not practiced before. It was hypothesized that the increase in cognitive load was directly relevant to learning and thus represented an increase in germane cognitive load instead of an increase in extraneous cognitive load. Confrontation with a highly varied sequence of problems and solutions to those problems helps learners extend or restrict the range of applicability of acquired schemas, but seems to require the mindful engagement of the learners, which increases cognitive load. An increase in cognitive load followed by enhanced transfer led to the description of the transfer paradox. This paradox suggests that instructional procedures that improve transfer performance typically have a negative effect on the cognitive load imposed on the learners during practice, or on the number of problems or training time needed to reach a pre-specified level of performance (van Merrienboer, de Croock, and Jelsma, 1997).

Initially, training problem variability and the resulting transfer paradox seem to contradict most of the results reported in this review, because these findings, discussed above, indicated that an increase of cognitive load during practice typically impairs learning. Paas and van Merrienboer (1994a, b) reinterpreted the results on variability of practice in the light of cognitive load theory by distinguishing between extraneous and germane cognitive load. They hypothesized that variability of practice would have a positive effect on learning and transfer in situations in which extraneous cognitive load was low, because in such situations the total cognitive load would be within limits, irrespective of the fact that variability increases germane cognitive load. In contrast, they predicted that variability of practice would have a negative effect on learning and transfer in situations in which extraneous cognitive load was high, because the total cognitive load would then overburden the learners' working memory. In Paas and van Merrienboer's (1994a) study, students had to learn a number of procedures in the domain of geometry. They had to solve conventional problems that imposed a high extraneous cognitive load under either high variability and thus high germane cognitive load conditions or under low variability and thus low germane cognitive load conditions. Alternatively, they had to study worked examples that imposed a low extraneous cognitive load, again combined with either high variability (high germane cognitive load) or low variability (low germane cognitive load) conditions. As predicted, a significant interaction was found between problem format and variability. A problem format associated with high extraneous cognitive load (conventional problems) showed no positive effects on transfer performance under conditions of high variability, but a problem format associated with low extraneous cognitive load (worked examples) yielded an acceptable level of cognitive load during practice and showed superior transfer performance under conditions of high variability (see Fig. 1 for a representation of those data).

These results fully accord with cognitive load theory. They first indicate that instructional designs should decrease extraneous cognitive load. But as an additional implication, they indicate that instructional designs that are successful in decreasing extraneous cognitive load may become even more effective if they increase germane cognitive load, provided that total cognitive load stays within limits. The results also call into question some common instructional designs for sequencing tasks or problems. A typical approach to lower cognitive load when students have to solve conventional problems is to use a simple-to-complex ordering of problems. However, with the possible exception of very high intrinsic cognitive load conditions, this approach is at variance with cognitive load theory: Although it prevents cognitive overload—an important consideration under very high intrinsic cognitive load conditions—it neither decreases extraneous cognitive load nor increases germane cognitive load. Cognitive load theory predicts more effective learning if problem formats are used that decrease extraneous cognitive load (e.g., goal-free problems, worked examples, completion assignments) in combination with more random sequences that increase germane cognitive load. This prediction is currently under investigation using a computer-based instructional program for teaching computer programming (CASCO; van Merrienboer, Krammer, and Maaswinkel, 1994; van Merrienboer and Luursema, 1995; van Merrienboer, Luursema, Kingma, Houweling, and de Vries, 1995).

CONCLUSIONS

In this paper, we have described a cognitive architecture and its instructional consequences. The architecture assumes several structures and processes. A working memory is used to process all instructional material. Working memory is very limited with respect to the number of elements it can handle but its capacity may be enhanced if information is processed using both the visual and auditory channel. All material handled by working memory can be transferred to long-term memory. Everything that is learned as a consequence of information that is processed in working memory is stored in an effectively limitless long-term memory in the form of schemas that can vary in their degree of automaticity. Both schema construction and automation have the dual function of storing information in long-term memory and reducing the load on working memory.

The architecture described in this paper is restricted in that we omitted any cognitive structures (e.g., sensory memory) that we felt did not have clear implications for instructional design. In addition, we have omitted structures and details that may not yet be well understood. As a consequence, we believe that the architecture discussed is likely to be widely accepted and quite noncontroversial. While it is not the only cognitive architecture discussed in the literature, it is widely accepted. Few are likely to disagree with the characteristics of the particular structures described.

Cognitive load theory assumes the above architecture and, in addition, makes assumptions concerning the structure of information. All information that must be processed in working memory can be placed on an element interactivity continuum. The elements of low element interactivity material interact minimally and so can be learned serially without imposing a heavy working memory load. These materials can be characterized as having a low intrinsic cognitive load. In contrast, interactions between the elements of high element interactivity material require them to be processed simultaneously in working memory resulting in a high intrinsic cognitive load. Learning reduces that load by embedding interacting elements in schemas that can be related as a single element in working memory.

These cognitive and information structures can be used to generate appropriate instructional designs that accord, rather than conflict, with our cognitive architecture. The practical implications were derived directly from the theory. Because the instructional designs (described previously) were generated by the cognitive architecture, we believe there is a closer association between cognitive processes and instructional design than would otherwise occur. Most of the instructional designs and procedures described were intended to reduce an extraneous cognitive load in the presence of a high intrinsic load. We will now summarize the instructional designs discussed in this paper.

In those areas of the curriculum where problem-solving performance is critical, an emphasis on goal-free problems, worked examples, and completion problems can all be effective. These designs were devised as alternatives to solving conventional problems that normally require means–ends search to attain solution. Means–ends search places heavy demands on limited working memory and these demands are largely irrelevant to schema construction and storage in long-term memory, or to automation of stored schemas, both essential for the development of expertise.

The split-attention effect derived from the worked example effect. Instructional designs requiring students to split their attention between and mentally integrate multiple sources of information can place as heavy a load on working memory as means–ends search, thus negating the positive effects of worked examples. Physically integrated instructional designs reduce cognitive load and so facilitate schema construction and attention. The elimination of split-attention is important for all forms of instruction, not just when using worked examples.

The modality effect partly derived from the split-attention effect. Rather than reducing the load on working memory by physically integrating split-sources of information, effective working memory can be increased by presenting some verbal sources of information in auditory rather than visual form. Dual modality presentation under split-attention conditions facilitates schema construction and automation as effectively as physical integration of visually presented materials. At present, there is no empirical evidence nor theoretical reasons for supposing that dual mode presentation is effective except under split-attention conditions.

The redundancy effect also derived from the split-attention effect. Multiple sources of information that do not require split-attention and integration, because one or more sources are redundant, should not be physically integrated. Working memory load is reduced and schema construction and automation enhanced by using designs that eliminate redundant sources of information. How redundant material is dealt with should depend on learners' level of expertise. Redundant material should be eliminated for expert learners but the same material might need to be integrated for novice learners. Material that is redundant for more experienced learners may be essential for less experienced learners.

Lastly, the variability effect was predicated on the assumption that while an increased extraneous cognitive load had negative consequences, increases in germane cognitive load could be beneficial. Variability increases germane cognitive load but that increase in load on working memory benefits schema construction.

In addition, experimental results have indicated that split-attention, redundancy, and the use of single- rather than dual-modality materials do not have appreciable negative consequences under conditions of low element interactivity. If intrinsic cognitive load is low due to low element interactivity, extraneous cognitive load may not be an important consideration for instructional designers. It becomes critical under conditions of high element interactivity. In addition to these findings, cognitive load theory yields many guidelines for the broader field of instructional systems design (ISD) which were not discussed in this review (van Merrienboer, 1997; van Merrienboer and Dijkstra, 1997; van Merrienboer, Jelsma, and Paas, 1992).

No techniques have been described in this paper that have not been extensively tested experimentally for effectiveness. We consider that testing program—based on replicated, controlled experimental designs—to be a major strength of cognitive load theory and the instructional procedures that it has generated. Furthermore, whereas early work within a cognitive load framework provided only indirect indicators of cognitive load, such as study times, later work used more direct measures such as secondary tasks and subjective rating scales of load. Work on subjective rating scales is beginning to develop a comprehensive set of statistical procedures specifically associated with cognitive load theory.

As can be seen from this review, most work within a cognitive load framework has been concerned with techniques designed to reduce extraneous cognitive load. Later work has been concerned with relations between extraneous and intrinsic cognitive load. There has been some work on procedures designed to increase cognitive load by increasing relevant mental effort (germane cognitive load) in order to facilitate schema construction. The use of problem variability is a major representative of this class of techniques. Nevertheless, to this point, that area does not have the same extensive repertoire of techniques and experiments as are associated with extraneous cognitive load. For future research, germane cognitive load will become a major focus of attention, and we expect considerable progress in this area over the next few years. Another important topic for further research concerns the implications of cognitive load theory for specific target groups with impaired cognitive capacity. For instance, there is a growing body of evidence to support the hypothesis that age-related declines in cognitive performance are most likely to occur in complex cognitive tasks requiring effortful processing. Because these tasks are highly dependent upon the availability of sufficient attentional resources for their successful completion, they are disproportionally compromised by age-related declines in cognitive capacity and age-related declines in the ability to inhibit irrelevant information. Instructional methods based on cognitive load theory, which can be argued to be more cognitive-capacity efficient, could be expected to compensate for these age-related declines. Research currently is commencing in this area.

ACKNOWLEDGMENTS

The work reported in this paper was supported by grants to the first author by the Royal Dutch Academy of Arts and Sciences, and the Australian Research Council.

REFERENCES

Aasman, J., Mulder, G., and Mulder, L. J. M. (1987). Operator effort and the measurement of heart-rate variability. *Human Fact.* 29: 161-170.

Anderson, J. R., Boyle, C. F., Corbell, A., and Lewis, M. (1990). Cognitive modelling and intelligent tutoring. *Artif. Intell.* 42: 7-49.

Ayres, P. (1993). Why goal-free problems can facilitate learning. *Contemp. Educ. Psychol.* 18: 376-381.

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

Barfield, W. (1986). Expert-novice differences for software: Implications for problem solving and knowledge acquisition. *Behav. Inform. Technol.* 5: 15-29.

Bobis, J., Sweller, J., and Cooper, M. (1993). Cognitive load effects in a primary school geometry task. *Learn. Instruct.* 3: 1-21.

Bobis, J., Sweller, J., and Cooper, M. (1994). Demands imposed on primary-school students by geometric models. *Contemp. Educ. Psychol.* 19: 108-117.

Borg, G. (1978). Subjective aspects of physical work. *Ergonomics* 21: 215-220.

Bratfisch, O., Borg, G., and Dornic, S. (1972). Perceived Item-Difficulty in Three Tests of Intellectual Performance Capacity. Report No. 29, Institute of Applied Psychology, Stockholm, Sweden.

Carroll, J. M. (1990). *The Nurnberg Funnel: Designing Minimalist Instruction for Practical Computer Skill*, MIT Press, Cambridge, MA.

Carroll, W. (1994). Using worked examples as an instructional support in the algebra classroom. *J. Educ. Psychol.* 86: 360-367.

Cerpa, N., Chandler, P., and Sweller, J. (1996). Some conditions under which integrated computer-based training software can facilitate learning. *J. Educ. Comput. Res.* 15: 345-367.

Chandler, P., and Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognit. Instr.* 8: 293-332.

Chandler, P., and Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *Brit. J. Educ. Psychol.* 62: 233-246.

Chandler, P., and Sweller, J. (1996). Cognitive load while learning to use a computer program. *Appl. Cognit. Psychol.* 10: 151-170.

Chase, W. G., and Simon, H. A. (1973). Perception in chess. *Cognit. Psychol.* 4: 55-81.

Chi, M., Bassok, M., Lewis, M., Reimann, P., and Glaser, R. (1989). Self-explanations: How students study and use examples in problem solving. *Cognit. Sci.* 13: 145-182.

Chi, M., Glaser, R., and Rees, E. (1982). Expertise in problem solving. In Sternberg, R. (ed.), *Advances in the Psychology of Human Intelligence*, Erlbaum, Hillsdale, NJ, pp. 7-75.

Cooper, G., and Sweller, J. (1987). The effects of schema acquisition and rule automation on mathematical problem-solving transfer. *J. Educ. Psychol.* 79: 347-362.

Cormier, S. M. and Hagman, J. D. (eds.) (1987). *Transfer of learning: Contemporary research and applications*, Academic Press, San Diego, CA.

De Groot, A. D. (1966). Perception and memory versus thought: Some old ideas and recent findings. In Kleinmuntz, B. (ed.), *Problem Solving*, Wiley, New York.

Egan, D. E., and Schwartz, B. J. (1979). Chunking in recall of symbolic drawings. *Mem. Cognit.* 7: 149-158.

Ericsson, K. A., and Charness, N. (1994). Expert performance: Its structure and acquisition. *Am. Psychol.* 49: 725-747.

Gopher, D., and Braune, R. (1984). On the psychophysics of workload: Why bother with subjective measures? *Human Fact.* 26: 519-532.

Hendy, C. H., Hamilton, K. M., and Landry, L. N. (1993). Measuring subjective workload: When is one scale better than many? *Human Fact.* 35: 579-601.

Jeffries, R., Turner, A., Polson, P., and Atwood, M. (1981). Processes involved in designing software. In Anderson, J. R. (ed.), *Cognitive Skills and Their Acquisition*, Erlbaum, Hillsdale, NJ, pp. 255-283.

Jelsma, O., and Bijlstra, J. P. (1988). Training for Transfer in Learning to Detect, Diagnose, and Compensate System Failures. Proceedings of the Seventh European Annual Conference on Human Decision Making and Manual Control, Paris, France, pp. 256-262.

Jelsma, O., and Bijlstra, J. P. (1990). PROCESS: Program for research on operator control in an experimental simulated setting. *IEEE Trans. Syst. Man, Cybernet.* 20: 1221-1228.

Jelsma, O., and van Merrienboer, J. J. G. (1989). Contextual interference: Interactions with reflection-impulsivity. *Percept. Motor Skills* 68: 1055-1064.

Jelsma, O., van Merrienboer, J. J. G. and Bijlstra, J. P. (1990). The ADAPT design model: Towards instructional control of transfer. *Instr. Sci.* 19: 89-120.

Kalyuga, S., Chandler, P., and Sweller, J. (1998). Levels of expertise and instructional design. *Human Fact.* 40: 1-17.

Kotovsky, K., Hayes, J. R., and Simon, H. A. (1985). Why are some problems hard? Evidence from Tower of Hanoi. *Cognit. Psychol.* 17: 248-294.

Lazonder, A. W., and Van der Meij, H. (1993). The minimal manual: Is less really more? *Int. J. Man Machine Stud.* 39: 729-752.

LeFevre, J., and Dixon, P. (1986). Do written instructions need examples? *Cognit. Instr.* 3(1): 1-30.

Lieberman, H. (1986). An example-based environment for beginning programmers. *Instr. Sci.* 14: 277-292.

Marcus, N., Cooper, M., and Sweller, J. (1996). Understand instructions. *J. Educ. Psychol.* 88: 49-63.

Mayer, R. E., and Anderson, R. B. (1991). Animations need narrations: An experimental test of dual-coding hypothesis. *J. Educ. Psychol.* 83: 484-490.

Mayer, R. E., and Anderson, R. B. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *J. Educ. Psychol.* 84: 444-452.

Mayer, R., Bove, W., Bryman, A., Mars, R., and Tapangco, L. (1996). When less is more: Meaningful learning from visual and verbal summaries of science textbook lessons. *J. Educ. Psychol.* 88: 64-73.

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. Instr.* 14: 1-43.

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

Miller, W. (1937). The picture crutch in reading. *Elem. Eng. Rev.* 14: 263-264.

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

Mulder, L. J. M. (1988). Assessment of Cardiovascular Reactivity by Means of Spectral Analysis. Unpublished doctoral dissertation, University of Groningen, Groningen, The Netherlands.

Mwangi, W., and Sweller, J. (in press). Learning to solve compare word problems: The effect of example format and generating self-explanations. *Cognit. Instr.*

Owen, E., and Sweller, J. (1985). What do students learn while solving mathematics problems? *J. Educ. Psychol.* 77: 272-284.

Paas, F. G. W. C. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive load approach. *J. Educ. Psychol.* 84: 429-434.

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

Paas, F. G. W. C., and van Merrienboer, 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. G. W. C., and van Merrienboer, J. J. G. (1994b). Instructional control of cognitive load in the training of complex cognitive tasks. *Educ. Psychol. Rev.* 6: 351-371.

Paas, F. G. W. C., van Merrienboer, J. J. G., and Adam, J. J. (1994). Measurement of cognitive load in instructional research. *Percept. Motor Skills* 79: 419-430.

Penney, C. (1989). Modality effects and the structure of short-term verbal memory. *Mem. Cognit.* 17: 398-422.

Pirolli, P. L. (1991). Effects of examples and their explanations in a lesson on recursion: A production system analysis. *Cognit. Instr.* 8: 207-259.

Pirolli, P. L. and Anderson, J. R. (1985). The role of learning from examples in the acquisition of recursive programming skills. *Can. J. Psychol.* 3: 240-272.

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.

Reder, L., and Anderson, J. R. (1980). A comparison of texts and their summaries: Memorial consequences. *J. Verb. Learn. Verb. Behav.* 19: 121-134.

Reder, L., and Anderson, J. R. (1982). Effects of spacing and embellishment on memory for main points of a text. *Mem. Cognit.* 10: 97-102.

Schneider, W., and Shiffrin, R. (1977). Controlled and automatic human information processing: I. Detection, search and attention. *Psychol. Rev.* 84: 1-66.

Schooler, J., and Engstler-Schooler, L. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognit. Psychol.* 22: 36-71.

Segal, J., and Ahmad, K. (1993). The role of examples in the leaching of programming languages. *J. Educ. Comput. Res.* 9: 115-129.

Shiffrin, R., and Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychol. Rev.* 84: 127-190.

Simon, H., and Gilmartin, K. (1973). A simulation of memory for chess positions. *Cognit. Psychol.* 5: 29-46.

Singley, M. K., and Anderson, J. R. (eds.) (1989). *The Transfer of Cognitive Skill*, Harvard University Press, Cambridge, MA.

Smith, S. M., Ward, T. B., and Schumacher, I. S. (1993). Constraining effects of examples in a creative generation task. *Mem. Cognit.* 21: 837-845.

Solman, R., Singh, N., and Kehoe, E. J. (1992). Pictures block the learning of sightwords. *Educ. Psychol.* 12: 143-153.

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

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

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

Sweller, J., and Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognit. Instr.* 2: 59-89.

Sweller, J., and Levine, M. (1982). Effects of goal specificity on means–ends analysis and learning. *J. Exp. Psychol: Learn., Mem., Cognit.* 8: 463-474.

Sweller, J., Mawer, R., and Ward, M. (1983). Development of expertise in mathematical problem solving. *J. Exp. Psychol. Gen.* 112: 634-656.

Sweller, J., Chandler, P., Tierney, P., and Cooper, M. (1990). Cognitive load and selective attention as factors in the structuring of technical material. *J. Exp. Psychol.: Gen.* 119: 176-192.

Tarmizi, R., and Sweller, J. (1988). Guidance during mathematical problem solving. *J. Educ. Psychol.* 80: 424-436.

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

Trafton, J. G., and Reiser, B. J. (1993). The Contribution of Studying Examples and Solving Problems to Skill Acquisition. Proceedings of the 15th Annual Conference of the Cognitive Science Society, Erlbaum, Hillsdale, NJ, pp. 1017-1022.

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

Van Merrienboer, J. J. G. (1992). Training strategies for teaching introductory computer programming. In Engel, F. L., Bouwhuis, D. G., Bsser, T., and d'Ydewalle, G. (eds.), *Cognitive Modelling and Interactive Environments in Language Learning*, Springer Verlag, Berlin, pp. 81-88.

Van Merrienboer, J. J. G. (1997). *Training Complex Cognitive Skills: A Four-Component Instructional Design Model for Technical Training*, Educational Technology Publications, Englewood Cliffs, NJ.

Van Merrienboer, 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 Merrienboer, J. J. G, de Croock, M. B. M., and Jelsma, O. (1997). The transfer paradox: Effects of contextual interference on retention and transfer performance in learning a complex cognitive skill. *Percept. Motor Skills* 84: 784-786.

Van Merrienboer, J. J. G., and Dijkstra, S. (1997). The four-component instructional design model for training complex cognitive skills. In Tennyson, R. D., and Schott, F. (eds.), *Instructional Design: Theory and Research* (Vol. 1), Lawrence Erlbaum, Hillsdale, NJ.

Van Merrienboer, J. J. G., Jelsma, O., and Paas, F. G. W. C. (1992). Training for reflective expertise: A four-component instructional design model for training complex cognitive skills. *Educ. Technol. Res. Devel.* 40(2): 23-43.

Van Merrienboer, J. J. G., and Krammer, H. P. M. (1987). Instructional strategies and tactics for the design of introductory computer programming courses in high school. *Instr. Sci.* 16: 251-285.

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

Van Merrienboer, J. J. G., Krammer, H. P. M., and Maaswinkel, R. M. (1994). Automating the planning and construction of programming assignments for teaching introductory computer programming. In Tennyson, R. D. (ed.), *Automating Instructional Design, Development, and Delivery* (NATO ASI Series F, Vol. 119), Springer Verlag, Berlin, pp. 61-77.

Van Merrienboer, J. J. G., and Luursema, J. J. (1995). Implementing instructional models in computer-based learning environments: A case study in problem selection. In Liao, T. T. (ed.), *Advanced Educational Technology: Research Issues and Future Potential* (NATO ASI Series F, Vol. 145). Springer Verlag, Berlin.

Van Merrienboer, J. J. G., Luursema, J. J., Kingma, H., Houweling, F., and de Vries, A. P. (1995). Fuzzy logic instructional models: The dynamic construction of programming assignments in CASCO. In Tennyson, R. D., and Barron, A. E. B. (eds.), *Automating Instructional Design: Computer-Based Development and Delivery Tools* (NATO ASI Series F), Springer Verlag, Berlin.

Van Merrienboer, J. J. G., and Paas, F. G. W. C. (1990). Automation and schema acquisition in learning elementary compute programming: Implications for the design of practice. *Comput. Human Behav.* 6: 273-289.

Vollmeyer, R., Burns, B., and Holyoak, K. (1996). The impact of goal specificity on strategy use and the acquisition of problem structure. *Cognit. Sci.* 20: 75-100.

Ward, M., and Sweller, J. (1990). Structuring effective worked examples. *Cognit. Instr.* 7: 1-39.

Wierwille, W. W., and Eggemeier, F. L. (1993). Recommendations for mental workload measurement in a test and evaluation environment. *Human Fact.* 35: 263-281.

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.

Zhu, X., and Simon, H. (1987). Learning mathematics from examples and by doing. *Cognit. Instr.* 4: 137-166.