2

Cognitive Load Theory: Recent Theoretical Advances

JOHN SWELLER

Cognitive Load Theory (CLT) began as an instructional theory based on our knowledge of human cognitive architecture. It proved successful in generating a series of cognitive load effects derived from the results of randomised, controlled experiments (Clark, Nguyen, & Sweller, 2006). This chapter summarises the theory, including its general instructional implications. Many of the theory's specific instructional implications, which provide its prime function and purpose, are discussed in other chapters in this volume and therefore will not be discussed in detail in this chapter (see Table 2.1 for a summary).

HUMAN COGNITION

The processes of human cognition constitute a natural information-processing system that mimics the system that gave rise to human cognitive architecture: evolution by natural selection. Both human cognition and biological evolution create novel information, store it for subsequent use, and are capable of disseminating that information indefinitely over space and time. By considering human cognition within an evolutionary framework, our understanding of the structures and functions of our cognitive architecture are being transformed. In turn, that cognitive architecture has profound instructional consequences. CLT is an amalgam of human cognitive architecture and the instructional consequences that flow from that architecture.

From an evolutionary perspective, there are two categories of human knowledge: biologically primary and biologically secondary knowledge (Geary, 2007, 2008). Biologically primary knowledge is knowledge we have evolved to acquire over many generations. Examples are general problem-solving techniques, recognising faces, engaging in social relations, and

TABLE 2.1. *Cognitive load effects*

Cognitive load effect	Description	Primary cognitive load source
Worked-Example	Studying worked examples results in better performance on subsequent tests of problem solving than solving the equivalent problems (Renkl, 2005).	Extraneous
Completion	Requiring learners to complete partially solved problems can be just as effective as worked examples (Paas & van Merriënboer, 1994).	Extraneous
Split-Attention	Multiple sources of information that are unintelligible in isolation result in less learning when they are presented in split-attention as opposed to integrated format (Ayres & Sweller, 2005).	Extraneous
Modality	Multiple sources of information that are unintelligible in isolation result in less learning when they are presented in single-modality as opposed to dual-modality format (Low & Sweller, 2005).	Extraneous
Redundancy	The presence of sources of information that do not contribute to schema acquisition or automation interfere with learning (Sweller, 2005).	Extraneous
Expertise reversal	With increasing expertise, instructional procedures that are effective with novices can lose their effectiveness, whereas ineffective techniques can become effective (Kalyuga, 2005).	Extraneous
Guidance fading	With increasing expertise, learners should be presented worked examples followed by completion problems and then full problems rather than worked examples alone (Renkl, 2005).	Extraneous
Goal-Free	Problems presented in goal-free form enhance learning compared with conventional problems (Paas, Camp, & Rikers, 2001).	Extraneous
Element interactivity	Cognitive load effects are only obtainable using high rather than low element interactivity material (Sweller, 1994).	Intrinsic
Isolated/interacting elements	Learning is enhanced if very high element interactivity material is first presented as isolated elements followed by interacting elements versions rather than as interacting elements form initially (Pollock, Chandler & Sweller, 2002).	Intrinsic
Variable examples	Examples with variable surface features enhance learning compared with examples with similar features (Paas & van Merriënboer, 1994).	Germane
Imagination	Imagining procedures or concepts enhance learning compared with studying materials (Leahy & Sweller, 2004).	Germane

listening to and speaking our native language. Primary knowledge is modular in that we have independent, cognitive modules that allow us to acquire the relevant knowledge unconsciously, effortlessly, and without external motivation simply by membership in a human society. Learning to speak our native language provides a clear example. We are not normally explicitly taught how to organise our lips, tongue, voice, and breath when learning to speak. We have evolved to learn these immensely complex procedures just by listening to others speaking.

In contrast, biologically secondary knowledge is culturally dependent. We have evolved to acquire such knowledge in a general sense rather than having evolved to acquire particular knowledge modules such as speaking. Biologically secondary knowledge is acquired consciously and usually requires mental effort. In modern times, we invented educational institutions to impart biologically secondary knowledge precisely because, unlike biologically primary knowledge, it tends not be learned simply by immersion in a functioning society. Virtually everything taught in educational institutions consists of biologically secondary knowledge. For example, unlike listening and speaking, few people are likely to learn to read and write without being explicitly taught to read and write. Simple immersion in a reading and writing society is unlikely to be sufficient.

CLT and, indeed, instructional design in general, applies to biologically secondary knowledge (Sweller, 2007, 2008). It does not apply to biologically primary knowledge. Thus, CLT is relevant to those aspects of knowledge dealt with in educational institutions (secondary knowledge) rather than the possibly far larger body of primary knowledge that we have specifically evolved to acquire.

When dealing with biologically secondary knowledge, human cognition can be characterised by five basic principles that govern its functions and processes. These principles apply equally to the processes that govern biological evolution (Sweller, 2003, 2004; Sweller & Sweller, 2006) and as such constitute a natural information processing system. They will be discussed in more detail subsequently but can be summarised as follows. The information store principle states that human cognition includes a large store of information that governs the bulk of its activity. Long-term memory provides this function. The borrowing and reorganising principle states that almost all of the information held in long-term memory has been borrowed from other long-term memory stores. Information obtained by imitation, listening, or reading exemplifies this process.

The randomness as genesis principle indicates that random generation followed by tests of effectiveness provide the initial source for the generation of all information held in long-term memory. When faced with a problem for which solution knowledge is not available or only partly available, the random generation of moves followed by tests of the effectiveness of those moves is an example. The narrow limits of change principle indicates that all effective changes to long-term memory occur slowly and incrementally. The capacity limitations of working memory when dealing with novel information exemplifies this principle. The environment organising and linking principle states that unlimited amounts of organised information from long-term memory can be used by working memory to determine interactions with the external world. These principles constitute a natural information processing system and derive from evolutionary theory. Although the derivation will not be discussed in this chapter (see Sweller, 2003, 2004; Sweller & Sweller, 2006), the link with biological evolution establishes an essential connection between the human cognitive system and human biology. A detailed description of the five principles follows.

Long-Term Memory and the Information Store Principle

Functioning in a complex environment requires a complex store of information to govern activity in that environment. The primary driver of activity of human cognition is its large store of information held in long-term memory. The realisation that long-term memory is not simply a repository of isolated, near-random facts but, rather, the central structure of human cognition, developed slowly. Its origins can probably be traced to the early work on expertise in the game of chess. When De Groot (1965) followed by Chase and Simon (1973) found that the only difference between chess masters and less able players was in memory of chess-board configurations taken from real games, it established the central importance of long-term memory to cognition. Chess has long been seen, appropriately, as a game that required the most sophisticated of human cognitive processes, a game of problem solving and thought. The discovery that the major difference between people who differed in ability was in terms of what they held in long-term memory changed our view of cognition. (It should be noted that what is held in long-term memory includes problem-solving strategies.) Long-term memory was not just used by humans to reminisce about the past but, rather, was a central component of problem solving and thought.

If long-term memory is essential to problem solving, we might expect results similar to those obtained using the game of chess to also be obtained in educationally more relevant areas. Increasing levels of expertise should be associated with increasing ability to reproduce relevant problem states and, indeed, the same expert–novice differences obtained in chess have been established in a variety of educationally relevant areas (e.g. Egan & Schwartz, 1979; Jeffries, Turner, Polson, & Atwood, 1981).

The amount of information required by the human cognitive system is huge. Although we have no metric for measuring the amount of information held in long-term memory, it might be noted that Simon and Gilmartin (1973) estimated that chess grand masters have learned to recognise many tens of thousands of the board configurations that are required for their level of competence. It is reasonable to assume that similar numbers of knowledge elements are required for skilled performance in areas more relevant to everyday life, including areas covered in educational contexts. If so, long-term memory holds massive amounts of information to permit adequate levels of performance in the various areas in which an individual is competent.

Although a large amount of information is held in long-term memory, where does the information come from? The next two sections discuss this issue.

Schema Theory and the Borrowing and Reorganising Principle

Based on the information store principle, human cognition includes a large store of information that governs most activity. What is the immediate source of that information? The borrowing and reorganising principle explains how most of the information found in any individual's long-term memory is acquired. Almost all information in long-term memory is obtained by imitating other people's actions or hearing or reading what others have said. In effect, our knowledge base is borrowed almost entirely from the long-term memory of other people. Nevertheless, the information borrowed is almost invariably altered and constructed. We do not remember exactly what we have heard or seen but, rather, construct a representation based on knowledge already held in long-term memory. Schema theory reflects that constructive process. A schema permits multiple elements of information to be treated as a single element according to the manner in which it will be used. Thus, a problem-solving schema permits us to classify problems according to their solution mode. A chess master has schemas that allow the classifying of chess-board configurations according to the moves required.

The modern origins of schema theory can be found in Piaget (1928) and Bartlett (1932), although the theory was largely ignored for several decades during the Behaviourist era. The relevance of schemas to problem solving was emphasised by Larkin, McDermott, Simon, and Simon (1980) and Chi, Glaser, and Rees (1982), who provided theory and data indicating that the possession of domain-specific schemas differentiated novices from experts in a particular area. The extent to which one is skilful in an area depends on the number and sophistication of one's schemas stored in long-term memory.

Schema construction, by indicating the form in which information is stored in long-term memory, provides us with a learning mechanism; therefore, learning includes the construction of schemas. Learning also includes the automation of schemas. Automation has occurred when knowledge is processed unconsciously (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977) rather than consciously in working memory. Problem solving using automated knowledge is much easier and more rapid than when basic processes must be consciously considered (Kotovsky, Hayes, & Simon, 1985). The automation of lower level schemas is frequently essential for the construction of higher level schemas. For example, without the automatic processing of the letters of the alphabet, because of the automation of schemas associated with recognising those letters, it would be difficult to combine those letters into words and sentences to permit reading.

A well-known study of Bartlett (1932), in providing a graphic example of the process of schema acquisition, also indicates how the borrowing and reorganising principle functions. One person read several paragraphs of a Native American legend and then wrote down from memory as much as possible of the passage. That remembered passage was then given to another person who again wrote down as much as possible from memory, with the process repeated with ten people. There were many alterations to the passage as it passed from person to person, and those alterations provided a window into memory. The alterations were not random, with Bartlett identifying two major categories. First, there was a levelling or flattening of aspects of the passage that were novel to the participants, resulting in a reduced emphasis or disappearance of these aspects entirely. Second, there was a sharpening of those aspects that accorded with knowledge in long-term memory, resulting in those aspects being emphasised. Thus, participants did not remember the passage as it was presented but rather, remembered a construction that consisted of a combination of the passage and previous information held in long-term memory. What we hold in memory consists of schematised constructions – schemas. While being constructed, the information held by schemas is essentially borrowed, via the borrowing and reorganising principle, from schemas held by others.

The various cognitive load effects shown in Table 2.1 provide strong evidence for the borrowing and reorganising principle. Each of the CLT effects listed in the table is concerned with techniques for presenting information to learners rather than having them generate information. The narrow limits of change principle, discussed later in the chapter, indicates why generating information is ineffective; therefore, since its inception, CLT has been concerned with techniques for presenting information to learners rather than having learners attempt to generate information. All cognitive load effects are intended to indicate how to provide auditory and visual information in a manner that best facilitates learning. In other words, CLT is concerned with how information held in the long-term memory of instructors can be borrowed for use by learners via schema acquisition.

Based on the borrowing and reorganising principle, learned information can be maintained by transmitting it among people indefinitely. But the transmission is rarely exact. It normally includes a constructive element that, over time, can result in substantial changes to the store. There is an inevitable random component to this process, and whether any changes are retained or jettisoned depends on their adaptive value, with beneficial changes retained and non-beneficial changes jettisoned. Adaptive value also is critical to the next principle discussed.

Problem Solving and the Randomness as Genesis Principle

Despite its constructive elements, the borrowing and reorganising principle is basically a device for combining and communicating information. It does not generate new information, which begs the question, how is the information that is transmitted via the borrowing and reorganising principle created in the first instance? A likely answer seems to be random generation followed by tests of effectiveness.

Consider a person dealing with a novel set of circumstances, for example, solving a problem. Most of the activity of that person will be based on knowledge held in long-term memory acquired through the borrowing and reorganising principle. Nevertheless, on many occasions, a problem solver will be faced with two or more possible moves and have no knowledge indicating which move should be made. Under these circumstances, random generation of novel problem solving moves followed by tests of effectiveness are needed. Random generation could be expected to lead to many dead-ends that will only be discovered after the event and, of course, when faced with a difficult novel problem, most problem solvers will, indeed, reach many dead-ends. A difficult problem may result in many more dead-ends than appropriate moves.

The relation between knowledge-based move generation and randomly generated moves can be described in terms of a central executive (Sweller, 2003). A central executive must know what the consequences of a move are prior to it occurring and then arrange for it to occur. An executive system such as this is possible for familiar moves generated by knowledge in long-term memory but impossible for novel moves. The moves made to solve a problem can originate from long-term memory but only to the extent that information is available in long-term memory indicating potential solution moves. Knowledge held in long-term memory can indicate what type of problem we are faced with and what types of moves are appropriate for that problem. That knowledge is acquired from previous experience – it is learned. Such knowledge provides the characteristics expected of a central executive and, in that sense, it is a learned central executive.

If information concerning potential solution moves is not available in long-term memory, the problem solver can select a move randomly and test the effectiveness of that move. Failing knowledge held in long-term memory, there may be no logical alternative. The potential effectiveness of a move cannot be determined prior to its selection, so random selection is needed. Prior knowledge of either the correct move or knowledge of the potential usefulness of a range of moves that permits a hierarchy of moves to be established eliminates the need for random selection. Failing that knowledge held in long-term memory, random selection can be used as a substitute. It should be noted that for the reasons outlined here, computational models of human problem solving require random generation of moves as a last resort when knowledge is unavailable (e.g., see Sweller, 1988). Of course, if knowledge is available to us, we are highly likely to use it.

On this argument, human cognition does include an executive but it is a learned executive held in the information store. That information acts as an executive indicating what should be done, when it should be done, and how it should be done. To the extent that this executive is unavailable because of a lack of relevant information held by an information store, random generation followed by tests of effectiveness are necessary procedures to alter the store. Because the procedures are required and because there is no substitute for them unless another store can be found from which the relevant information can be communicated, random generation followed by effectiveness testing provides the genesis of all information held by long-term memory. Indeed, it can be argued that it is the source of all human creativity (Sweller, 2009).

Novice Working Memory and the Narrow Limits of Change Principle

A major consequence that flows from the randomness as genesis principle is that all random alterations to the information store must be incremental and slow. A large, rapid, random change to long-term memory is unlikely to be adaptive in the sense that it is likely to destroy critical aspects of the store. Small, incremental changes will leave most of the store intact and functioning and are therefore unlikely to be fatal. The larger the change, the larger the probability that previous structures that have been established as effective have been compromised. Furthermore, small changes can be reasonably tested for effectiveness. Assume a small working memory that must deal with three elements that must be combined in some manner. Assume further that the various combinations are tested using the logic of permutations. There are $4! = 4 \times 3 \times 2 \times 1 = 24$ possible permutations. It is within the bounds of possibility to test 24 permutations. In contrast, assume a somewhat larger working memory that can handle ten elements. This working memory must test $10! = 10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 3{,}628{,}800$ permutations, which in most situations is unlikely to be possible. Thus, working memory is the structure that ensures that alterations to long-term memory are limited, with working memory unable to handle large amounts of novel information. We are unable to hold more than about seven items of novel information in working memory (Miller, 1956) and can probably process no more than about four items (Cowen, 2001).

The narrow limits of change principle provides a central plank of CLT. Competence derives from a large store of information held in long-term memory and largely borrowed from the long-term memories of others. The narrow limits of change principle suggests that that information must be carefully structured to ensure working memory is not overloaded and that schemas are effectively constructed and transferred to long-term memory. The CLT effects listed in Table 2.1 are intended to meet this objective. The narrow limits of change principle indicates why CLT, with its emphasis on a limited working memory, argued that encouraging learners to generate information that could readily be presented to them was likely to be an ineffective instructional technique. A working memory that can process no more than two to four elements of information is unlikely to be capable of rapidly generating the knowledge required. That knowledge is available via the borrowing and reorganising principle. Each of the effects listed in Table 2.1 is based on the assumption that instruction should emphasise the borrowing and reorganising of knowledge from knowledgeable instructors and should be structured in a manner that reduces unnecessary cognitive load. According to CLT, the more that is borrowed and the less that learners need to generate themselves, the more effective the instruction is likely to be (Kirschner, Sweller, & Clark, 2006).

It must be emphasised that the limitations imposed by the narrow limits of change principle only apply to novices who, by definition, are dealing with information with which they are unfamiliar. As indicated next, information that has been organised by schemas in long-term memory does not suffer these limitations.

Expert Working Memory and the Environment Organising and Linking Principle

In contrast to working memory limitations when dealing with novel information, there are no known limits to the amount of information that working memory can process if it has been organised and tested for effectiveness, that is, if it comes from long-term memory. The environment organising and linking principle explains how we are able to transfer massive amounts of organised information from long-term to working memory to effect the complex actions required of the human cognition. To account for the huge amount of organised information from long-term memory that can be handled by working memory compared with the small amount of novel information that can be handled, Ericsson and Kintsch (1995) suggested a new construct, long-term working memory. Unlike short-term working memory, long-term working memory allows the rapid processing of large amounts of information providing that information has previously been organised in long-term memory.

The environment organising and linking principle provides the ultimate justification for human cognition: the ability to function in a complex external environment. Experts can transfer large amounts of organised, schematic information from long-term to working memory to perform appropriately in their environment. CLT assumes that novice–expert differences (see next chapter) primarily result from differences in schematic information held in long-term memory that can be transferred as a single entity to working memory to generate actions appropriate to an environment. Furthermore, the environment organising and linking principle requires the preceding principles. According to the environment organising and linking principle, we need a large information store and learning mechanisms to build that store. The learning mechanisms are the borrowing and reorganising and randomness as genesis principles. In turn, the narrow limits of change principle permits the learning principles to function without destroying the information store. Once information is stored in the information store, the environment organising and linking principle allows that information to be used to guide appropriate action.

INSTRUCTIONAL IMPLICATIONS

These five principles provide a base for human cognitive architecture. Together, they result in a self-perpetuating, integrated, information processing system. As indicated earlier, it is a natural information processing system that is also used during biological evolution (Sweller, 2003, 2004; Sweller & Sweller, 2006), thus effecting a necessary connection between cognition and the biological world. The elimination of any one of the five principles will eliminate the functionality of the system. Its ability to generate action and accommodate to changing circumstances to continue to generate action requires all five principles.

The centrality of the five principles to human cognitive architecture mirrors their centrality to CLT and to instructional design. Each principle has instructional design consequences. From the information store principle, we know that the major function of instruction is to alter the information held in long-term memory. Based on the information store principle, learning can be defined as a change in long-term memory. According to this definition, we have no grounds for assuming the effectiveness of proposed instructional techniques that cannot specify the changes in long-term memory that follow from use of the techniques (Kirschner et al., 2006). In contrast, the accumulation of knowledge in long-term memory is central to CLT.

The borrowing and reorganising principle suggests that the bulk of human knowledge is learned from others rather than discovered by problem solving or a similar process. We acquire some of the knowledge held in the long-term memory of other people by copying what they do, listening to what they say, or reading what they have written. We must engage in the difficult task of constructing knowledge in our own long-term memories by these processes of copying, listening, or reading (Kirschner et al., 2006; Mayer, 2004). Accordingly, CLT emphasizes ways of communicating knowledge in educational contexts through observing, listening, and reading, including reading diagrams.

Discovery occurs over generations, with its guiding assumption being the randomness as genesis principle. The random components of true discovery ensure that the process may be difficult or impossible to teach. In contrast, teaching can assist in the accumulation of knowledge held in long-term memory. What instructional designs are likely to assist in the goal of changing the information held in long-term memory? Those designs that take into account the narrow limits of change principle. In human cognition, that assumption is expressed through the limitations of working memory when dealing with novel information. In contrast, there are not only no limits to the amount of organised information that can be held in long-term memory, there are no known limits to the amount of organised information from long-term memory that can be used by working memory. This alteration in the characteristics of working memory from a limited capacity, limited duration structure when dealing with novel information to an unlimited capacity, unlimited duration structure when dealing with familiar, organised information from long-term memory is expressed through the environment organising and linking principle and is central to CLT. The primary purpose of CLT has been to indicate how to present novel information structured according to the narrow limits of change principle to reduce unnecessary working memory load and facilitate change in long-term memory. In turn, changes in long-term memory permit complex actions through the environment organising and linking principle. There are three categories of cognitive load that affect working memory and these will be discussed next.

CATEGORIES OF COGNITIVE LOAD

Whereas the function of instruction is to increase schematic knowledge in long-term memory, novel information must first be processed by working memory, and when dealing with novel information, working memory is limited in capacity (e.g., Miller, 1956) and duration (e.g. Peterson and Peterson, 1959). All instructional material imposes a working memory or cognitive load, and that cognitive load can be divided into two independent categories – intrinsic and extraneous – with a third category, germane cognitive load, dependent on intrinsic cognitive load. Intrinsic and extraneous cognitive load are additive, and together, they determine the total cognitive load. If that cognitive load exceeds working memory capacity, information processing, including learning, will be compromised. In other words, if total working memory load is excessive, the probability of useful changes to long-term memory is reduced. Each of the three categories of cognitive load will be discussed separately before a discussion of how they interact.

Intrinsic Cognitive Load

Some material is intrinsically difficult to understand and learn regardless of how it is taught. The critical factor is element interactivity, which refers to the number of elements that must be simultaneously processed in working memory to understand and learn material under instruction. For example, assume a student is learning the symbols for chemical elements. Each element can be learned independently of every other element. The task may be difficult because there are many elements that must be learned, but it does not impose a heavy working memory load. Because the working memory load is light, the issue of "understanding" does not arise. We may have failed to learn or forgotten the symbol for a particular element, but we are not likely to use the term "understanding" in this context.

In contrast, assume that the learner is learning how to balance a chemical equation or deal with an algebraic equation. The number of relevant elements may be far less than the number dealt with when learning chemical element symbols, but element interactivity is high. When learning how to solve the problem $(a + b)/c = d$, *solve for a*, one cannot just attend to one of the elements in the equation or one of the solution steps while ignoring the others if the solution is to be understood. All of the learning elements interact and unless all are considered simultaneously in working memory, the problem and its solution will not be understood. Element interactivity is high, so working memory load and intrinsic cognitive load are high.

In some senses, element interactivity is fixed because it is an intrinsic property of all material that must be learned and cannot be altered. Nevertheless, this statement needs to be modified by two points. First, regardless of element interactivity, it is always possible to learn material one element at a time. In the case of high element interactivity material, that means the interacting elements are treated as though they do not interact. Learning can proceed in this manner but understanding cannot. Until all of the elements are processed in working memory, understanding will not occur. For very complex material, the learning of interacting elements as isolated elements may be unavoidable, leading to the isolated/interacting elements effect (Pollock et al., 2002; van Merriënboer & Sweller, 2005). By presenting learners with high element interactivity material as isolated elements and only requiring them to learn the relevant interactions later, learning is enhanced compared with requiring learners to learn the interacting elements immediately when instruction commences.

Learning itself provides the second way in which the effects of high element interactivity can be reduced. Element interactivity cannot be determined merely by analysing the nature of the material that needs to be learned. Depending on the schemas that have been acquired, material that is complex for one individual may be very simple for another. If a set of interacting elements have become incorporated into a schema, only that schema needs to be processed in working memory, not the interacting elements. Accordingly, working memory load is low. For readers of this chapter, the word "accordingly" is treated as a single element because we all have schemas for this written word. For someone who is just beginning to learn to read written English, the individual letters may need to be processed simultaneously, and that is a high element interactivity task that may exceed the capacity of working memory. Once the word is learned, individual words can be treated as a single element rather than as multiple elements. In this way, a high intrinsic cognitive load due to element interactivity is altered by learning.

As can be seen, this alteration in the intrinsic cognitive load due to learning is a consequence of the alteration in the characteristics of working memory depending on whether information is organised or random. According to the environment organising and linking principle, there are no limits to the amount of organised information that can be used for action. This organised information is held in long-term memory, and large amounts of such information can be transferred to working memory. In contrast, according to the randomness as genesis and the narrow limits of change principles, there are severe limits to the extent to which the store can be changed to permit new actions. Accordingly, only a limited amount of novel information can be organised in a manner that alters long-term memory.

Extraneous Cognitive Load

This category of cognitive load also depends on element interactivity but unlike intrinsic cognitive load, the interacting elements are fully under instructional control, and CLT was devised primarily to provide principles for the reduction of extraneous cognitive load. Whether an instructional procedure imposes an extraneous cognitive load can be assessed by determining whether it is in accord with the cognitive principles outlined earlier. If an instructional procedure does not facilitate change to the information store (long-term memory), if the procedure attempts to alter that store by use of the randomness as genesis principle instead of the borrowing and reorganising principle or functions on the assumption that randomness as genesis can or should be taught, or if the instructional procedure ignores the narrow limits of change principle by either ignoring the limitations of working memory or proceeding on the assumption that working memory has no limitations, that instructional procedure is likely to be ineffective because it unnecessarily introduces interacting elements that should be eliminated. As an example, consider a person attempting to learn via a discovery learning technique. Rather than being told a scientific rule, the person is given minimal information and required to work out the rule from that information. The act of discovering a rule is highly likely to make heavy demands on working memory, thus violating the narrow limits of change principle. It depends minimally on the communication of knowledge, thus violating the borrowing principle. To the extent that knowledge is unavailable, discovery relies heavily on random generation followed by tests of effectiveness, which is an extremely slow, ineffective way of accumulating information because discovery introduces a large range of interacting elements unrelated to learning. As a consequence, there is minimal emphasis on building knowledge in the information store – long-term memory – which should be the primary goal of instruction. All discovery and problem-solving based teaching techniques follow this pattern, and as a consequence, all violate every one of the five cognitive principles outlined earlier.

Based on the previously described theory, we might expect there to be no systematically organised body of research demonstrating the effectiveness of discovery-based teaching techniques and, indeed, after almost a half-century of protagonism, the lack of evidence for these techniques is glaring (Kirschner et al., 2006). Evidence needs to consist of randomised, controlled studies altering one variable at a time. In contrast, there is extensive evidence for the advantages of providing learners with information rather than having them discover it themselves. The worked-example effect, according to which learners provided with worked examples learn more than learners provided with the equivalent problems to solve, flows directly from the cognitive architecture described earlier. The effect indicates in the clearest possible terms the advantages of providing learners with information rather than having them discover it for themselves.

There are many other CLT effects. All depend on one or more of the five cognitive principles outlined previously. Many are discussed in the chapters of this volume. All are summarised in Table 2.1. More detailed summaries may also be found in Sweller (2003, 2004).

Germane Cognitive Load

Reducing extraneous cognitive load would have little function if the working memory resources so freed were not used for productive learning. Instruction should be designed to ensure that the bulk of working memory resources is germane to the goal of schema acquisition and automation. In other words, working memory resources should be devoted to dealing with intrinsic cognitive load rather than extraneous cognitive load because schema acquisition is directed to the interacting elements associated with intrinsic cognitive load. Instructional designs that increase the use of working memory resources devoted to intrinsic cognitive load have the effect of increasing germane cognitive load, which should be increased to the limits of working memory capacity. Beyond that point, increases in germane cognitive load become counterproductive and can be categorised as extraneous cognitive load.

Much of the work on extraneous cognitive load assumed that as that category of cognitive load was reduced, germane cognitive load would automatically increase because learners would devote a similar effort to learning regardless of the effectiveness of the instruction. For that reason, there are relatively few germane cognitive load effects. There are some, nevertheless, and the example variability and imagination effects are summarised in Table 2.1.

Interactions among Sources of Cognitive Load

As indicated earlier, intrinsic and extraneous cognitive load are additive, and if they exceed available working memory capacity, learning (indeed, all information processing) will be compromised and is likely to cease. All of the cognitive load effects listed in Table 2.1 are caused by various interactions among these sources of cognitive load. Most of the effects occur because a reduction in extraneous cognitive load permits an increase in working memory resources devoted to intrinsic cognitive load, increasing germane cognitive load and enhancing learning. These effects only occur if intrinsic cognitive load is high. If intrinsic cognitive load is low, alterations in extraneous cognitive load may not matter because sufficient working memory resources are likely to be available to overcome a poor instructional design that imposes a heavy extraneous cognitive load, giving rise to the element interactivity effect (Table 2.1).

Not only is the type of material critical to CLT, so is learner knowledge. Information or learner activities that are important to novices may interfere with further learning by more expert learners, giving rise to the expertise reversal effect (Table 2.1). In other words, as expertise increases, procedures that were important for novices, and therefore part of germane cognitive load, contribute to extraneous cognitive load for more expert learners. Although complex, considerable information is now available concerning these various interactions between sources of cognitive load that give rise to the various cognitive load effects. Nevertheless, interactions between various categories of cognitive load constitute a major, current research area, and much still remains to be done.

CONCLUSIONS

There is likely to be widespread agreement that instructional design requires knowledge of cognition. If we do not understand the mechanisms of learning and problem solving, our chances of designing effective instruction are likely to be minimal. The success of CLT as an instructional theory is heavily dependent on its view of human cognition. The theory does more than merely pay lip service to the organization and function of the structures that constitute human cognitive architecture. Cognitive architecture is central to the theory. Unless we have a conception of the bases of human intelligence and thought, effective instructional procedures are likely to elude us. The suggested principles that constitute human cognition provide one such possible base. That cognitive base, in turn, can inform us of the types of instructional procedures that are likely to be effective. The success of CLT in generating the instructional effects shown in Table 2.1 provides some evidence for the validity of the underlying assumptions of the theory.

REFERENCES

Ayres, P., & Sweller, J. (2005). The split-attention principle. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 135–146). New York: Cambridge University Press.

Bartlett, F. (1932). *Remembering: A study in experimental and social psychology.* London: Cambridge University Press.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, 4, 55–81.

Chi, M., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 7–75). Hillsdale, NJ: Erlbaum.

Clark, R., Nguyen, F., & Sweller, J. (2006). *Efficiency in learning: Evidence-based guidelines to manage cognitive load.* San Francisco: Pfeiffer.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87–114.

De Groot, A. (1965). *Thought and choice in chess.* The Hague, Netherlands: Mouton. (Original work published 1946).

Egan, D. E., & Schwartz B. J. (1979). Chunking in recall of symbolic drawings. *Memory and Cognition*, 7, 149–158.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102, 211–245.

Geary, D. (2007). Educating the evolved mind: Conceptual foundations for an evolutionary educational psychology. In J. S. Carlson & J. R. Levin (Eds.), *Psychological perspectives on contemporary educational issues* (pp. 1–99). Greenwich, CT: Information Age Publishing.

Geary, D. (2008). An evolutionarily informed education science. *Educational Psychologist*, 43, 179–195.

Jeffries, R., Turner, A., Polson, P., & Atwood, M. (1981). Processes involved in designing software. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 255–283). Hillsdale, NJ: Erlbaum.

Kalyuga, S. (2005). Prior knowledge principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 325–337). New York: Cambridge University Press.

Kirschner, P., Sweller, J., & Clark, R. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential and inquiry-based teaching. *Educational Psychologist*, 41, 75–86.

Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problems hard? Evidence from Tower of Hanoi. *Cognitive Psychology*, 17, 248–294.

Larkin, J., McDermott, J., Simon, D., & Simon, H. (1980). Models of competence in solving physics problems. *Cognitive Science*, 4, 317–348.

Leahy, W., & Sweller, J. (2004). Cognitive load and the imagination effect. *Applied Cognitive Psychology*, 18, 857–875.

Low, R., & Sweller, J. (2005). The modality principle. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 147–158). New York: Cambridge University Press.

Mayer, R. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *American Psychologist*, 59, 14–19.

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

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

Paas, F., & van Merriënboer, J. (1994). Variability of worked examples and transfer of geometrical problem solving skills: A cognitive-load approach. *Journal of Educational Psychology*, 86, 122–133.

Peterson, L., & Peterson, M. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology*, 58, 193–198.

Piaget, J. (1928). *Judgement and reasoning in the child*. New York: Harcourt.

Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. *Learning and Instruction*, 12, 61–86.

Renkl, A. (2005). The worked out example principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 229–245). New York: Cambridge University Press.

Schneider, W., & Shiffrin, R. (1977). Controlled and automatic human information processing: I. Detection, search and attention. *Psychological Review*, 84, 1–66.

Shiffrin, R., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, 84, 127–190.

Simon, H., & Gilmartin, K. (1973). A simulation of memory for chess positions. *Cognitive Psychology*, 5, 29–46.

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

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

Sweller, J. (2003). Evolution of human cognitive architecture. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 43, pp. 215–266). San Diego, CA: Academic Press.

Sweller, J. (2004). Instructional design consequences of an analogy between evolution by natural selection and human cognitive architecture. *Instructional Science*, 32, 9–31.

Sweller, J. (2005). The redundancy principle. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 159–167). New York: Cambridge University Press.

Sweller, J. (2007). Evolutionary biology and educational psychology. In J. S. Carlson & J. R. Levin (Eds.), *Psychological perspectives on contemporary educational issues* (pp. 165–175). Greenwich, CT: Information Age Publishing.

Sweller, J. (2008). Instructional implications of David Geary's evolutionary educational psychology. *Educational Psychologist*, 43, 214–216.

Sweller, J. (2009). Cognitive bases of human creativity. *Educational Psychology Review*, 21, 11–19.

Sweller, J., & Sweller, S. (2006). Natural information processing systems. *Evolutionary Psychology*, 4, 434–458.

van Merriënboer, J., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, 17, 147–177.