Pergamon

0959–4752(94)00010–7

COGNITIVE LOAD THEORY, LEARNING DIFFICULTY, AND INSTRUCTIONAL DESIGN

JOHN SWELLER

University of NSW, Australia

Abstract

This paper is concerned with some of the factors that determine the difficulty of material that needs to be learned. It is suggested that when considering intellectual activities, schema acquisition and automation are the primary mechanisms of learning. The consequences of cognitive load theory for the structuring of information in order to reduce difficulty by focusing cognitive activity on schema acquisition is briefly summarized. It is pointed out that cognitive load theory deals with learning and problem solving difficulty that is artificial in that it can be manipulated by instructional design. Intrinsic cognitive load in contrast, is constant for a given area because it is a basic component of the material. Intrinsic cognitive load is characterized in terms of element interactivity. The elements of most schemas must be learned simultaneously because they interact and it is the interaction that is critical. If, as in some areas, interactions between many elements must be learned, then intrinsic cognitive load will be high. In contrast, in different areas, if elements can be learned successively rather than simultaneously because they do not interact, intrinsic cognitive load will be low. It is suggested that extraneous cognitive load that interferes with learning only is a problem under conditions of high cognitive load caused by high element interactivity. Under conditions of low element interactivity, re-designing instruction to reduce extraneous cognitive load may have no appreciable consequences. In addition, the concept of element interactivity can be used to explain not only why some material is difficult to learn but also, why it can be difficult to understand. Understanding becomes relevant when high element interactivity material with a naturally high cognitive load must be learned.

Introduction

The difficulties we face when learning new intellectual tasks can fluctuate dramatically. Learning can vary from being trivially easy to impossibly hard. Some of the reasons for variations in ease of acquisition, such as changes in amount of information, are obvious. In other cases, two tasks may appear to have roughly similar amounts of information but differ enormously in the effort required to achiev ` mastery. Students can find the concepts and procedures discussed in some curriculum areas notoriously intractable

---

while other areas may contain copious quantities of information that nevertheless, can be assimilated readily.

This paper is concerned with the features that make some material hard to learn. Since questions concerning learning difficulty are likely to be unanswerable without first establishing mechanisms of learning, in the first and second sections I will indicate what I believe to be the major, relevant learning processes and their place in our cognitive architecture. In the third section, the instructional consequences of these mechanisms will be summarized. The fourth and major section will be concerned with some structural differences in categories of information and the consequences of these structural features for the instructional modes discussed in the third section. The fifth section discusses some of the empirical and theoretical implications of the analysis.

What is Learned?

There are two critical learning mechanisms: schema acquisition and the transfer of learned procedures from controlled to automatic processing. It will be argued that intellectual mastery of any subject matter is overwhelmingly dependent on these two processes.

Schemas

A schema is a cognitive construct that organizes the elements of information according to the manner with which they will be dealt. An early discussion of schemas was presented by Bartlett (1932). He demonstrated that what is remembered is only partly dependent on the information itself. Newly presented information is altered so that it is congruent with knowledge of the subject matter. Knowledge of subject matter is organized into schemas and it is these schemas that determine how new information is dealt with. For example, consider schemas that deal with common objects such as trees. No two trees have identical elements but each tree seen can be instantly incorporated into a tree schema. As a consequence, if asked to describe a particular tree from memory, a person's description will be heavily influenced by a tree schema rather than entirely by the particular tree elements (leaves, branches, colour etc.) actually seen. Tree schemas allow people to deal effortlessly with the potentially infinite variety of objects called trees.

In a similar manner, there are schemas for dealing with problems. These schemas allow the classification of problems into categories according to how they will be dealt with, i.e., according to solution mode (e.g., see Chi, Glaser & Rees, 1982). Most people who have completed algebra courses, if faced with an algebraic problem such as $(a + b)/c = d$, solve for a, will be able to solve it immediately irrespective of the actual pro-numerals used. If, for example, the expression on the right side of the equation is long and complex, a schema will indicate that complexity at this location is irrelevant and the problem will be no more difficult to solve than with a simple expression. Schemas for this category of algebra problems allow the infinite variety of expressions incorporated in the category to be dealt with.

Schemas can be used to explain most of the learned, intellectual skills that people exhibit. People are able to read the infinite variety of the printed and handwritten versions of text that they can potentially encounter because they have acquired schemas for each letter, many words and probably even many word combinations. Learning to solve problems occurs by learning problem categories defined by the moves required for solution. These schemas permit people to readily solve problems that otherwise they would have immense difficulty solving if they had to rely solely on constructing a solution based on first principles.

Interest in schema theory has waxed and waned over many years with alternative terminology frequently being employed. Miller's (1956) concept of a chunk could be used as readily as the term schema, as could Schank and Abelson's (1977) scripts. In more recent times, Koedinger and Anderson (1990) provided an excellent formal analysis of schema-based problem solving. While their model is restricted to geometry problem solving, there seems little reason to suppose that the basic principles they employ should not be generalizable to a wide range of problem solving materials. Low and Over (1990) provide techniques for assessing schema acquisition for word problems that may be generalizable to other types of material.

In summary, knowledge and intellectual skill based on knowledge is heavily dependent on schema acquisition. Schemas provide the basic unit of knowledge and through their operation can explain a substantial proportion of our learning-mediated intellectual performance.

Automation of Intellectual Operations

Schemas tend to be discussed as though schema acquisition results in dichotomous states: a person either has or has not acquired schemas. In fact, few intellectual skills are acquired in this manner. When something is first learned, the ability to use it is likely to be severely constrained. A student who has just learned how to multiply out the denominator of an equation cannot do so easily or fluently. He or she can do so only with considerable thought and effort. Similarly, an educated adult can read text without conscious effort whereas a child who has been learning for only a few years, while being able to read, will only be able to do so with considerable effort.

While intellectual skill through schema acquisition is acquired gradually and incrementally rather than in the all-or-none fashion that it is sometimes conveniently thought of, it also has been convenient to treat one of the underlying cognitive mechanisms in a dichotomous manner. We assume that the way in which information is processed can be either controlled or automatic (Schnieder & Shiffrin, 1977; Shiffrin & Schnieder, 1977; Kotovsky, Hayes & Simon, 1985). Controlled processing occurs when the information at hand is consciously attended to. Any cognitive activity that requires deliberate thought is being processed in a controlled fashion. Readers thinking about the contents of this paper are engaging in controlled processing. In contrast, automatic processing occurs without conscious control. Well learned material can be processed automatically without conscious effort allowing attention to be directed elsewhere. Readers of this paper can read the words on the page without conscious effort. There is no need to deliberate about the meaning of individual letters or words because processing at this level switched from conscious to automatic long ago. In contrast, someone who

is still learning to read may need to devote close and constant attention to individual letters and words rather than to deeper meaning. The consequences for understanding are, of course, inevitable.

While we treat controlled and automatic processing as dichotomous, the switch from one to the other is probably always continuous and slow. As familiarity with a domain is gained, the need to devote attention to the required processes is reduced. Gradually, they become more automated, freeing cognitive resources for other activities. This process of automation is the second major learning mechanism after schema acquisition and affects everything learned, including schemas themselves. Consider what needs to be automated in order to fluently solve problems such as $(a + b)/c = d$, solve for a. Some of the basic rules of algebra need to be learned and then automated. For example, when students first learn to multiply out a denominator, they may know and understand the rule, but they cannot use it without reminding themselves of the mechanics and conditions under which it is used (see Cooper & Sweller, 1987). It is only after considerable practice that they can multiply out a denominator automatically while thinking about some other aspect of the problem such as whether the move makes sense. Furthermore, before even considering multiplying c, it may be recognized that this problem configuration requires multiplying out the denominator as the first move. In other words, the student may have an appropriate schema. But this schema may be usable under conscious or automated control. The student may need to carefully study the expression before realizing that it is amendable to multiplying out the denominator or alternatively, he or she may glance at it briefly and be immediately aware of the category to which it belongs without engaging in any conscious thought at all. This schema that can be used to classify the problem may be fully automated, only usable under conscious control or fall anywhere in between.

In summary, when a complex intellectual skill is first acquired, it may be usable only by devoting considerable cognitive effort to the process. With time and practice, the skill may become automatic to the point where it may require minimal thought for its operation. It is only then that intellectual performance can attain its full potential. Without automation, performance is slow, clumsy and prone to error. It is an essential mechanism of learning.

What is the Function of Learning?

From the above analysis, one function of learning is self-evident: to store automated schemas in long-term memory. The ability to store huge numbers of schemas may be a primary intellectual characteristic. Evidence for the importance of schemas comes from work on novice–expert differences that suggests that differential access to a large store of schemas is a critical characteristic of skilled performance. Beginning with De Groot's (1965) work on novice–expert differences in chess, many studies in a wide variety of areas have established that experts are better able to recognize and reproduce briefly seen problem states than novices (e.g., Egan & Schwartz, 1979; Jeffries, Turner, Polson & Atwood, 1981; Sweller & Cooper, 1985). It can be assumed that experts are better able to remember problem configurations because their schemas permit them to see the configuration as a single entity rather than as, for example, the large number of chess pieces that novices must attempt to remember after briefly seeing a chess board configuration. Simon and Gilmartin (1973) have estimated that in intellectually complex

areas experts have acquired tens of thousands of schemas which are the building blocks of intellectual skill.

While storing information in long-term memory is an obvious function of learning, it may not be the only one. The two learning mechanisms discussed above, schema acquisition and automation, share one intriguing characteristic. Both have the effect of substantially reducing working memory load. It has been known since Miller (1956) that in contrast to a huge long-term memory, working memory is very limited. Working memory can store and process no more than a few discrete items at any given time. A major function of schema acquisition and automation may be to ameliorate or even by-pass this restriction.
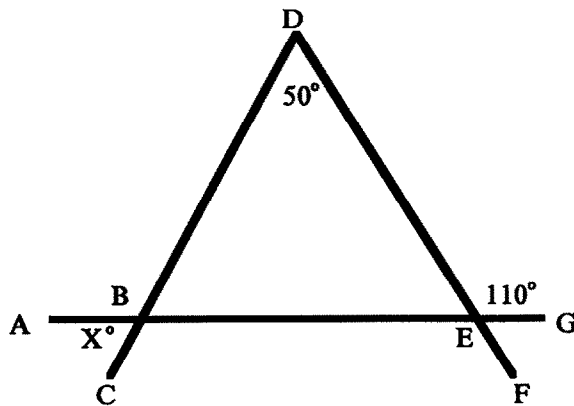
Schemas effectively increase the amount of information that can be held in working memory by chunking individual elements into a single element. A single tree, not thousands of leaves and branches needs to be remembered; a single word, not the individual letters or marks on a piece of paper need be remembered; the number of words on a page may exceed working memory but the number of ideas or concepts may not. In this sense, while the number of items held in working memory may be very limited, thanks to schemas, the amount of information held in working memory may be quite large and this may be one of the functions of schema acquisition. A schema not only permits long-term memory storage but also ameliorates working memory limitations.

Automation also has a significant effect on working memory. It permits working memory to be by-passed. Processing that occurs automatically requires less working memory space and as a consequence, capacity is freed for other functions. In this sense, automation, like schema acquisition, may have a primary function of circumventing limited processing capacity. Both schema acquisition and automation may occur precisely because of the characteristics of long-term and working memories. Given a superb long-term memory and relatively ineffective working memory, schema acquisition and automation are precisely the learning mechanisms that might be expected to occur.

Facilitating Learning and Problem Solving

If schemas are critical to learning and problem solving, what conditions are most likely to facilitate acquisition? Over the last decade or so, cognitive load theory (Sweller, 1988, 1989) has been used to investigate several instructional techniques. The theory suggests that instructional techniques that require students to engage in activities that are not directed at schema acquisition and automation, frequently assume a processing capacity greater than our limits and so are likely to be defective. In fact, a considerable array of commonly used techniques seem to incidentally incorporate just such an assumption of a processing capacity far in excess of most human beings.

When students are given relatively novel problems to solve, they will not be able to use previously acquired schemas to generate solutions. Nevertheless, they still may be able to find a solution. Most frequently, the strategy of choice for novice problem solvers in a given area is means-ends analysis (see Chi, Glaser, & Rees, 1982; Larkin, McDermott, Simon, & Simon, 1980). A means-ends strategy involves attempting to extract differences between each problem state encountered and the goal state and then finding problem solving operators that can be used to reduce or eliminate those differences. For example, assume a student is faced with the problem of finding a value

Figure 1. Conventional geometry problem and solution.

for Angle X of Figure 1. The initial problem state is the givens of the diagram. The goal state is a value for Angle X. The problem solving operators are the theorems of geometry. Using a means-ends strategy, a problem solver may attempt to find a series of theorems connecting Angle X to the knowns of the problem. For example, he or she may notice that if a value for Angle DBE could be found, the problem could be solved because Angles X and DBE, being vertically opposite, are equal. Angle DBE can become a subgoal. The next step is to discover that a value can be found for Angle DBE because Angle DBE = Angle DEG − Angle BDE. (The external angles of a triangle equal the sum of the vertically opposite internal angles.) Once a value for Angle DBE is obtained, a value for Angle X can be obtained and the problem is solved. (Most readers, of course, will have schemas for the solution to this problem involving supplementary angles and the angles of a triangle adding to 180 degrees. The above solution is merely used for convenience.)

This means-ends procedure is a highly efficient technique for attaining the problem goal. It is designed solely for this purpose. It is not intended as a learning technique and bears little relation to schemas or schema acquisition. In order to acquire an appropriate problem solving schema, students must learn to recognize each problem state according to its relevant moves. Using a means-ends strategy, much more must be done. Relations between a problem state and the goal state must be established; differences between them must be extracted; problem operators that impact favourably on those differences must be found. All this must be done essentially simultaneously and repeated for each move keeping in mind any subgoals. Furthermore, for novices, none of the problem

states or operators are likely to be automated and so must be carefully considered. According to cognitive load theory, engaging in complex activities such as these that impose a heavy cognitive load and are irrelevant to schema acquisition will interfere with learning. Students solving a series of practice geometry problems similar to Figure 1 do so with the ultimate intention of learning. The strategy they use is efficient in attaining the problem goal but is not efficient in attaining their real goal: schema acquisition and automation.

What procedures might better facilitate learning? A very long series of experiments generated by cognitive load theory over the last decade has indicated some instructional techniques that can be used as alternatives to conventional procedures. The use of reduced goal-specificity or goal-free problems was the first technique investigated (Owen & Sweller, 1985; Sweller, Mawer, & Ward, 1983; Tarmizi & Sweller, 1988). A goal-free equivalent of the above geometry problem asks problem solvers to "find the value of as many angles as possible" rather than to specifically "find a value for Angle X." It was reasoned that goal-free problems would eliminate the use of a means-ends strategy and its attendant misdirection of attention and imposition of a heavy cognitive load. Furthermore, a goal-free strategy should direct attention only to those aspects of a problem essential to schema acquisition: problem states and their associated moves.

Many experiments demonstrated repeatedly that goal-free problems facilitated learning. Sweller (1988) provided additional evidence for a reduced cognitive load associated with goal-free problems using production system models. Ayres and Sweller (1990) used cognitive load theory to predict major sources and locations of errors during geometry problem solving.

A goal-free strategy is not the only way to reduce extraneous cognitive load and direct attention to those aspects of a problem that should assist in schema acquisition and automation and indeed, under conditions where a very large number of moves can be generated, the strategy may be quite inappropriate if many of the moves are trivial. Cooper and Sweller (1987) and Sweller and Cooper (1985) suggested that worked examples could have the same effect as goal-free problems. They used algebra worked examples of the following type:

$$(a + b)/c = d \quad \text{Solve for } a$$
$$a + b = dc$$
$$a = dc - b$$

In order to follow this example, it is only necessary to attend to each line (or problem state) and the algebraic rule (or move) needed for the transformation to the next line. As was the case for goal-free problems, this activity corresponds closely to that required for schema acquisition. It might be expected that studying such worked examples should result in more rapid schema acquisition than solving the equivalent problems by means-ends analysis. Again, many experiments confirmed that studying algebra worked examples facilitated learning compared to solving the equivalent problems.

There are other demonstrations of the worked example effect. Zhu and Simon (1987) found a three year mathematics course was completed in 2 years by emphasizing worked examples rather than conventional instruction. Paas (1992) and Paas and Van Merrienboer (1994) found that worked out statistical or geometrical problems were superior to conventional problems. These latter two studies are particularly important

because they incorporated subjective measures of cognitive load that provided direct evidence that the worked example effect is caused by cognitive load factors.

Contrary to what might be expected, the above results do not indicate that worked examples should necessarily replace conventional problems: they indicate that extraneous cognitive load should be eliminated. It cannot be assumed that all worked examples under all circumstances will have beneficial consequences. Consider the conventional geometry worked example of Figure 1. The diagram alone tells us nothing of the solution. In turn, the solution steps below the diagram are quite unintelligible in isolation. Before the worked example can be understood, the diagram and the solution steps must be mentally integrated. The act of mental integration requires cognitive resources. These cognitive resources are required purely because it is conventional to present geometry diagrams and their associated statements as discrete, physically independent entities. Because they are not cognitively independent, we must make a cognitive effort to overcome the physical independence. This cognitive effort, while essential given the design of the worked example, is not intrinsically required to understand the relevant geometry. It is only required because of the format used and as such, an extraneous cognitive load is imposed.

The cognitive effort required to mentally integrate disparate sources of information can be reduced or eliminated by physically integrating the various entities. Figure 2 provides a physically integrated variant of the worked example of Figure 1. As can be seen, the solution presented in both figures is identical. The major difference is that Figure 2 has the statements physically integrated within the diagram. A large number of experiments using a wide variety of curriculum materials has demonstrated that both worked examples and other instructional materials are assimilated much more rapidly when presented in integrated rather than conventional format with much higher subsequent test performance levels (Chandler & Sweller, 1991; Chandler & Sweller, 1992; Purnell, Solman, & Sweller, 1991; Sweller, Chandler, Tierney, & Cooper, 1990; Tarmizi & Sweller, 1988; Ward & Sweller, 1990). These results provide evidence of the split-attention effect. The most obvious explanation for this effect is in terms of the imposition of an extraneous cognitive load.
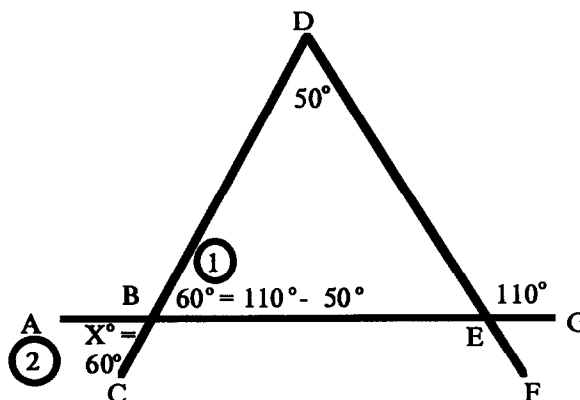

Figure 2. Integrated geometry problem and solution.

Just as not all worked examples are effective if cognitive load principles are ignored, so the integration of disparate sources of information can be ineffective if no reference is made to cognitive load effects. We should not conclude from the preceding findings that, for example, all diagrams and their associated texts should be integrated. Consider the example used by Chandler and Sweller (1991). They presented students with a fully labelled and descriptive diagram depicting the flow of blood through the heart, lungs and body. This diagram was associated with a series of statements describing aspects of the diagram such as "Blood from the lungs flows into the left atrium." Similar examples are common in biology and other texts. For most students, the diagram is self-explanatory and the text redundant. The self-contained nature of the diagram contrasts markedly with the materials discussed above that lead to the split-attention effect. Those materials are unintelligible in isolation and must be integrated, either physically or mentally, before they can be processed. In the case of the materials used by Chandler and Sweller (1991), integration is not necessary. The material can be learned fully from the diagram alone. If the text is redundant, processing it imposes an extraneous cognitive load. Furthermore, integrating the diagram and text is likely to unnecessarily force students to process the text leading to integration having negative rather than positive effects. Under these circumstances, extraneous cognitive load can be reduced by eliminating the text rather than integrating it with the diagram. This redundancy effect has been obtained by Chandler and Sweller (1991) and Bobis, Sweller and Cooper (1993) using a variety of students and materials.

Other instructional techniques also have been devised based on cognitive load theory. For example, Paas (1992) and Van Merrienboer and De Croock (1992) have used cognitive load theory to predict that partially completed problems that students had to complete themselves would reduce cognitive load compared to solving the entire problem. Results supported this hypothesis using mathematical and computer programming problems.

This section has described several techniques for facilitating learning by reducing extraneous cognitive load. There are bound to be many more undiscovered procedures for reducing cognitive load. With respect to the procedures already discovered, should cognitive load theory and the techniques described above be applied to the design of all learning and problem solving materials? Almost certainly not. If the materials themselves do not impose a heavy cognitive load, the extraneous cognitive load imposed by instructional techniques may not be important because the total cognitive load may not exceed the processing capacity of the individual. The next section discusses the characteristics of material to which cognitive load theory should be applied.

Element Interactivity

The findings summarized in the previous section suggest that extraneous cognitive load should be an important consideration when designing instruction. Extraneous cognitive load, by definition, is entirely under instructional control. It can be varied by varying the manner in which information is presented and the activities required of students. Nevertheless, the cognitive load that is imposed by material that needs to be learned is not just a function of instructional design. Cognitive load imposed by instructional material can be partitioned into that which is due to the intrinsic complexity of the

core information and that which is a function of the cognitive activities required of students because of the manner in which the information is presented. A study of intrinsic complexity requires techniques for comparing different types of information. The next section provides one potential framework.

Informational Complexity

Assume people are presented with a simple paired associate task in which pairs of words must be memorized so that the second word of each pair can be stated on presentation of the first word. While paired-associate learning is artificial, some real tasks do bear a degree of similarity to paired associate lists. Having to learn a second language vocabulary without concentrating on its syntactic or complex semantic aspects provides one example.

While the difficulty of learning paired associates can be varied by using a memory strategy such as the use of imagery, or by using nonsense syllables instead of real words, nevertheless, difficulty is closely related to the number of items on the list. For present purposes, the important points are (a) that this simple task can be very difficult if the list is long enough and (b) that each element is simple to learn and largely independent of every other element. In this paper, an element is defined as any material that needs to be learned, in this case a paired associate. While there may be some unintended interference between paired-associates, each pair can be learned in isolation and furthermore, considered in isolation, each pair presents a trivially easy task.

When the elements of a task can be learned in isolation, they will be described as having low element interactivity. The level of element interactivity or connectedness refers to the extent to which the elements of a task can be meaningfully learned without having to learn the relations between any other elements. Elements interact if they are related in a manner that requires them to be assimilated simultaneously. In other words, the structure of the task is such that it would be meaningless to attempt to learn elements one at a time. In contrast, elements do not interact if they can be assimilated serially. Paired associate learning is probably the ultimate in low element interactivity because the paired associates can be learned one at a time without reference to any other paired associate.

High element interactivity or connectedness occurs when a task cannot be learned without simultaneously learning the connections between a large number of elements. While learning some aspects of a second language vocabulary was used as an example of low element interactivity, learning syntactic and semantic elements tends to have a higher level of interactivity. Learning appropriate word orders in English provides an example. It is appropriate to say *when learning English* but not appropriate to use any other combination such as *English when learning*. Learning the appropriate word order of this phrase requires the relative position of all three categories of words to be learned simultaneously. The elements, which consist of the relative word positions, cannot be learned serially because they interact.

Much of mathematics seems to involve relatively high element interactivity. Learning a simple mathematical procedure such as how to multiply out a denominator involves a large number of interacting elements. Assume a student is learning to multiply out the b in the equation, $a/b = c$. In order to learn this process, the student must simultaneously

learn that the numerator on the left side and the denominator which is not shown on the right side, remain unchanged. The denominator on the left side is eliminated and appears on the right side as cb. Furthermore, if the student is to have any understanding of the logic of the manipulation, the full intermediate steps, $ab/b = cb$ followed by cancellation of the b's need to be understood and learned. All of these elements must be processed in an essentially simultaneous rather than serial fashion. When learning to multiply out a denominator, it makes little sense to learn what happens to the left side denominator without simultaneously learning what happens to the rest of the equation. If a student does learn the process as a series of steps, we are likely to feel that understanding has not been attained. Learning how to multiply out a denominator involves processing all of the elements and relations between them simultaneously. The elements have a very high degree of interactivity.

It must be emphasized that initially, the individual steps required to multiply out a denominator can, and in most circumstances, are learned serially. A student can learn that a/b can be multiplied by b giving ab/b. Independently, they can learn that the b's can be cancelled out in ab/b giving a. They also can learn that anything done to one side of an equation must be done to the other. In the normal course of events, a student may be taught and learn each of these procedures independently and without reference to the other procedures. These tasks do not interact and so are low in element interactivity at this point. The irreducible interaction occurs when students must learn to multiply out a denominator in order to isolate a pronumeral on one side of an equation. No matter how well automated the individual elements are, at this point they and their relations must be consciously considered simultaneously.

A similar analysis can be made of a wide variety of curriculum materials. Students learning to move on an (X, Y) coordinate system, first will learn to move on the X and Y axes separately. Subsequently, when they must learn to move on both axes simultaneously, the complex interactions of the elements associated with the two axes must be considered simultaneously because of high element interactivity.

In contrast to high element interactivity materials, for other areas, the degree of interaction of the various elements learned may be limited. Learning the anatomy and associated terminology of a biological specimen provides an example. While some interaction exists, much can be learned individually without ever considering the rest of the anatomy. The task may be difficult and lengthy because of the amount of information, not because of element interactivity.

Schemas and Elements

An element was defined above as any information that needs to be learned. It follows, that we cannot determine beforehand, merely by analysing the materials, what constitutes an element. The knowledge of the learner as well as the characteristics of the material must be taken into account. The more sophisticated and knowledgeable the learner, the more complex will be the elements he or she is dealing with. For instance, the algebra example above was analysed from the perspective of a student who is just beginning to learn elementary algebra. That example, for most of the readers of this paper, may itself act as a single element if it needs to be used in a novel way in a different context: perhaps as part of an algebra word problem. The schema associated

with multiplying out a denominator may be a single element when more expert problem solvers deal with more complex procedures such as algebra word problems. When learning to use basic algebra to solve algebra word problems, the schemas of basic algebra are some of the elements of algebra word problems. Learning to solve algebra word problems involves learning the interactions between these schema/elements.

From this analysis, it may be seen that schemas organize elements and can act as elements themselves in higher order schemas. We develop schemas used to solve some mathematics problems. These schemas can then act as elements in more complex tasks that must be learned. Once a schema has been acquired and automated in the more complex task, it too can act as an element in further tasks. In effect, when dealing with high interactivity tasks requiring the learning of multiple elements, we are dealing with schema acquisition. The schemas being acquired may be considered higher or lower level. The elements involved in higher order schema acquisition may be lower level schemas.

When dealing with very low level interactivity tasks such as paired associate learning, it is inappropriate to use the term schemas because most theorists have applied the term *schema* to complex materials that involve multiple, interacting elements. When dealing with the learning of simpler tasks such as paired associate lists, each paired associate can best be thought of as an element rather than a schema. Nevertheless, when we are concerned with second language vocabulary learning, which bears some relation to paired-associate learning, it needs to be recognized that the elements that need to be learned must be used subsequently in the higher level interactivity tasks associated with syntax and semantics. At this level, using accepted definitions, learning involves schema acquisition.

Estimating the Extent of Element Interactivity

A precise measure of element interactivity that is independent of the learner is unobtainable because, as indicated above, what constitutes an element is affected by the knowledge of the individual. For example, for readers of this paper, previously acquired schemas permit words or combinations of words to act as single elements. For someone who has just learned to read, individual letters act as schemas and so reading a word may involve several interacting elements rather than the single element of an experienced reader. Nevertheless, by assuming the knowledge level of a learner, it is possible to estimate the number of interacting elements that must be acquired simultaneously in order to learn a particular task or procedure.

Assume a person is learning how to multiply out the denominator on one side of an equation in order to make the numerator the subject of the equation. The person is learning how to transform $a/b = c$ into $a = cb$. The number of elements that must be learned simultaneously can be estimated by listing and counting as follows:

(1)   Multiply the left side by b giving ab/b.
(2)   Because the left side has been multiplied by b, the same operation must be carried out on the right side, giving cb, in order to maintain equality.
(3)   The new equation is $ab/b = cb$.

(4)   The b's in the numerator and denominator on the left side can cancel giving a.

(5)   The new equation is $a = cb$.

These 5 elements interact in the sense that there is little function, purpose or meaning in any of them in isolation. Each element is meaningful only in conjunction with the other four elements. To learn how to multiply out a denominator from one side to the other side of an equation requires consideration of all the elements simultaneously. While in isolation, each element is simple and easily learned, one cannot learn, for example, the third element without at least learning the first two and in order to see its function, probably the last two as well. All the elements interact.

The five interacting elements of the above example may be contrasted numerically with the single elements of some other subject matter. The example of learning the nouns of a foreign language has been used above. In most cases, because the elements do not interact, they can be learned in isolation giving an element interactivity count of one.

It must be emphasized that the five elements that must be considered simultaneously in the algebra example above only provide an estimate based on the assumed knowledge of the learner. For most readers of this paper, an automated schema incorporating all five elements will have been acquired long ago and so the element count is one, rather than five. In contrast to people for whom multiplying out a denominator is a single rather than five elements, for some algebra novices the five elements may require expansion. As an example, Element 1 above is assumed to be a single element because most algebra students will be aware that multiplying a/b by b results in ab/b. If a student attempts to learn the above procedure without a schema for the first element, it would need to be divided into two elements, with the first indicating that the left side of the equation needs to be multiplied by b and the second that the consequence is the expression ab/b.

Element Interactivity and Cognitive Load

We might expect element interactivity to have cognitive load consequences. If both element interactivity and instructional formats have cognitive load consequences, relations between these factors need to be considered. I would like to suggest that total cognitive load is an amalgam of at least two quite separate factors: extraneous cognitive load which is artificial because it is imposed by instructional methods and intrinsic cognitive load over which instructors have no control. The primary determinant of intrinsic cognitive load is element interactivity. If the number of interacting elements in a content area is low it will have a low cognitive load with a high cognitive load generated by materials with a high level of element interactivity. On this analysis, intrinsic cognitive load is determined largely by element interactivity.

Halford, Maybery and Bain (1986) and Maybery, Bain and Halford (1986) provided evidence for the importance of element interactivity as a source of cognitive load. Using transitive inference problems (e.g., a is larger than b; b is larger than c; which is the largest?) they hypothesized that integrating the two premises should generate the heaviest cognitive load because element interactivity is at its highest at this point. Evidence was provided for this hypothesis using secondary task analysis.

While there is a clear distinction between intrinsic and extraneous cognitive load, from the point of view of a student required to assimilate some new material, the distinction is irrelevant. Learning will be difficult if cognitive load is high, irrespective of its source. In contrast, from the point of view of an instructor, the distinction between intrinsic and extraneous cognitive load is important. Intrinsic cognitive load is fixed and cannot be reduced. On the other hand, extraneous cognitive load caused by inappropriate instructional designs can be reduced using the techniques discussed previously. Nevertheless, while intrinsic cognitive load cannot be altered, it does have important implications for instructional design. The implications are discussed in the next section.

Some Instructional Implications of Intrinsic Cognitive Load

We know, from previous work, discussed above, that inappropriate instructional designs can impose a heavy extraneous cognitive load that interferes with learning. In addition, it was suggested in the previous section, that element interactivity also imposes a cognitive load. If cognitive load is caused by a combination of design features and element interactivity, then the extent to which it is important to design instruction to reduce extraneous cognitive load, may be determined by the level of element interactivity. While extraneous cognitive load can severely reduce instructional effectiveness, it may do so only when coupled with a high intrinsic cognitive load. If the total cognitive load is not excessive due to a relatively low intrinsic cognitive load, then a high extraneous cognitive load may be irrelevant because students are readily able to handle low element interactivity material with almost any form of presentation. In contrast, if intrinsic cognitive load is high because of high element interactivity, adding a high extraneous cognitive load may result in a total load that substantially exceeds cognitive resources, leading to learning failure.

Because of the predilections of the investigators, the goal-free, worked example, split-attention and redundancy effects (discussed above) were all tested using high element interactivity materials with a high intrinsic cognitive load. Associating such materials with high extraneous cognitive load presentation modes may result in overwhelmingly high cognitive loads. As a consequence, it is to be expected that reducing extraneous cognitive by the various techniques associated with each effect results in substantial performance increments. Nevertheless, the advantages found may be available only with high element interactivity materials. All the effects may disappear using low element interactivity materials because total cognitive load levels may not exceed available capacity.

Consider the split-attention effect. Sweller *et al.* (1990) demonstrated this effect teaching students numerical control programming. This language requires students, among other things, to learn how to move an object using a co-ordinate system with a very high level of element interactivity. In common with other co-ordinate systems, it is difficult, if not impossible, to learn how the system works without learning the entire system. To move an object from one position to another, one must learn, for example, that a diagonal movement can be represented by simultaneous movements on both the X and Y axes, in addition to learning the codes for moving on these two axes. Basically, proficiency can be obtained only by learning how each of the elements of the coordinate system interact. Simply learning one element such as moving up the

X-axis will not provide an essential understanding of the system. All elements and their relations must be learned. Sweller *et al.* (1990) found that integrating diagrams of the coordinate system with explanatory text was far superior to the conventional split-source format of diagrams and separate text.

In contrast to numerical control programming, consider another computer application such as learning to use a word processor. This application may be taught by separately explaining the meaning of each command and diagrammatically demonstrating its screen output and/or consequences or by integrating the explanation with the output and consequences to eliminate split-attention. In this case, eliminating split-attention may have no positive consequences. This result would not follow because word processor procedures involve less information or less time to learn than numerical control programming. Indeed, it may take longer to learn how to use a word processor than to learn elementary aspects of numerical control programming. The word processing task appears easier because each element is relatively independent of other elements and can be learned readily without reference to other elements. Learning how to insert text can be learned quite independently of learning how to delete text or how to move the cursor about the screen or how to format a document for printing. Each command can be learned in isolation with minimal interaction between them. As a consequence, intrinsic cognitive load is low and integrating command meaning with diagrams of its screen consequences may have minimal effects on learning efficiency. Sweller and Chandler (1994) found that the split-attention effect could be obtained when learning a numerical control programming language but not when learning word-processing procedures.

Similar arguments apply to the other effects generated by cognitive load theory. The redundancy effect is not likely to occur if we are dealing with low element interactivity materials and a low intrinsic cognitive load. If each redundant segment of material can easily and readily be assimilated, its inclusion may not have negative consequences. Again, Sweller and Chandler (1994) obtained the redundancy effect using numerical control programming but not word processing.

As other examples, both the goal-free and worked example effects occur because goal-free problems and worked examples are compared to solving conventional problems by means-ends analysis. A means-ends strategy invariably involves high element interactivity because it requires problem solvers to simultaneously consider the goal, the current problem state, differences between them, problem solving operators and relations between these various entities. (Relations between element interactivity and means-ends analysis were pointed out to me by Paul Chandler.) If problem solving strategies other than means-ends analysis with reduced element interactivity are employed, the goal-free and worked example effects may not occur. Comparing worked examples with a problem solving strategy that does not require the problem solver to simultaneously process several elements is not likely to result in a worked example advantage. Indeed, goal-free problem solving is just such a strategy. Compared to a means-ends strategy, a goal-free strategy requires problem solvers to process only a very limited number of elements at any given time. To solve a goal-free problem one merely needs to consider a problem state and any operator that can be used at that point (see Sweller, 1988). It is reasonable to assume that any problem solving strategy used by subjects that reduces element interactivity compared to means-ends analysis should reduce cognitive load and reduce or eliminate the goal-free or worked example effects. (It needs to be recognized

that when we are discussing problem solving strategies, normally we are concerned with extraneous rather than intrinsic cognitive load because the load can be altered by altering the strategy used by students. If a change in strategy affects cognitive load then we are dealing with extraneous rather than intrinsic cognitive load.)

In summary, the instructional consequences of extraneous cognitive load may be heavily determined by intrinsic cognitive load caused by element interactivity. An extraneous cognitive load may have minimal consequences when dealing with material that has low element interactivity because the total cognitive load may be relatively low. The effects of extraneous cognitive load may manifest themselves primarily when dealing with high element interactivity materials because the combined consequences of a high extraneous and high intrinsic cognitive load may overwhelm limited processing capacity. Thus, we should not expect to demonstrate those effects reliant on cognitive load using low element interactivity materials.

Some Theoretical and Instructional Consequences of Element Interactivity

Our limited processing capacity is one of the most important and well known of our cognitive characteristics. The consequences of this limitation on the manner in which information is presented and received is not nearly as well known. Despite the minimal attention paid to cognitive load characteristics of information until recently, this aspect of the materials with which students must interact may be the most important factor that instructional designers must consider. In this context, element interactivity of the information being assimilated can be a vital aspect of the design process.

Cognitive load theory now has been used to generate novel instructional designs in a variety of contexts using a very wide variety of materials. Nevertheless, despite the range of materials used, it turns out that they all had one characteristic in common. All the materials seem to have had a high degree of element interactivity resulting in a high intrinsic cognitive load. A high degree of element interactivity may be an essential condition for the generation of the effects associated with cognitive load theory. Without a high degree of element interactivity, extraneous cognitive load may have no discernible consequences. In fact, it may be useful to consider element interactivity as an effect in its own right. Just as the worked example effect will not occur if worked examples are presented in split-attention format, so none of the cognitive load effects may occur if element interactivity is low. Initial data collected strongly support this hypothesis.

The concept of element interactivity may have explanatory significance in other contexts. Understanding plays an important role in both theoretical and practical treatments of higher level cognition. Nevertheless, the concept of understanding has been difficult to explain or even to define. What are the processes of understanding and why is some information difficult to understand? Why, on some difficult tasks such as learning lengthy paired associate lists does the concept of understanding not even apply while it is critical on other, easier tasks containing apparently little information such as "understanding" a simple mathematical procedure? Element interactivity may provide an answer to these questions.

Material may be difficult to understand if it incorporates a high level of element interactivity. If material cannot be learned without the simultaneous assimilation of

multiple interacting elements, it is likely to be assumed that the material contains difficult concepts that are hard to understand. If students manage to assimilate some but not all of the elements and their relations, there is a tendency to say that they have failed to understand the concept or only partially understood it. Thus, if a student, in multiplying out the denominator of the equation, $a/b = c$, ends up with the equation, $a/b = cb$, it will be assumed that the procedure has not been understood. In the terminology of this paper, not all of the elements and their relations have been learned. In contrast, if the material consists of elements that interact minimally, failure to learn some of the elements tends to be interpreted as nothing more than learning failure. The concept of understanding is not invoked. If a language student is unable to indicate the translation of the word *cat*, it normally would not be interpreted as a failure of understanding. Rather, it is a failure of learning or memory.

From this analysis, it can be seen that the concept of understanding is only applied to some but not other material. The perspective taken in this paper suggests that information that needs to be "understood," rather than merely learned, consists of material that has a high degree of element interactivity. Material that has a low level of interactivity only needs to be learned rather than both understood and learned. In this context, understanding can be defined as the learning of high element interactivity material. In fact, it can be suggested that all information falls on a continuum from low to high element interactivity and learning is the only cognitive factor operating. When the schemas associated with high element interactivity material have been acquired, people feel they have understood the material. When the schemas have become automated, it is understood very well.

The analysis presented in this paper has empirical consequences both for experimenters and for instructional designers. Experimenters who design experiments based on some aspect of cognitive load theory may not obtain any of the effects associated with the theory if they use relatively low element interactivity materials. Effects may be non-existent or weak compared to those obtainable using high element interactivity materials. Instructional designers, in turn, who base their designs on cognitive load theory but whose materials have low element interactivity, may be incorporating design features that have no useful effects. The effects generated by cognitive load theory may apply only to high element interactivity material. As a consequence, the theory may be irrelevant when dealing with low element interactivity materials.

References

Ayres, P., & Sweller, J. (1990). Locus of difficulty in multi-stage mathematics problems. *American Journal of Psychology*, **103**, 167–193.

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

Bobis, J., Sweller, J., & Cooper, M. (1993). The redundancy effect in an elementary school geometry task. *Learning and Instruction*, **3**, 1–21.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, **8**, 293–332.

Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, **62**, 233–246.

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.

Cooper, G., & Sweller, J. (1987). The effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology*, **79**, 347–362.

De Groot, A. (1965). *Thought and choice in chess*. The Hague: Mouton.

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

Halford, G., Maybery, M., & Bain, J. (1986). Capacity limitations in children's reasoning: A dual task approach. *Child Development*, **57**, 616–627.

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.

Koedinger, K., & Anderson, J. (1990). Abstract planning and perceptual chunks: Elements of expertise in geometry. *Cognitive Science* , **14**, 511–550.

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.

Low, R., & Over, R. (1990). Text editing of algebraic word problems. *Australian Journal of Psychology*, **42**, 63–73.

Maybery, M., Bain, J., & Halford, G. (1986). Information processing demands of transitive inference. *Journal of Experimental Psychology: Learning Memory and Cognition*, **12**, 600–613.

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

Owen, E., & Sweller, J. (1985). What do students learn while solving mathematics problems? *Journal of Educational Psychology*, **77**, 272–284.

Paas, F. (1992). Training strategies for attaining transfer of problem solving skill in statistics: A cognitive load approach. *Journal of Educational Psychology*, **84**, 429–434.

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

Purnell, K., Solman, R., & Sweller, J. (1991). The effects of technical instructions on cognitive load. *Instructional Science*, **1991**, 443–462.

Schank, R., & Abelson, R. (1977). *Scripts, plans, goals, and understanding*. Hillsdale, NJ: Erlbaum.

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 autmomatic 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. (1989). Cognitive technology: Some procedures for facilitating learning and problem solving in mathematics and science. *Journal of Educational Psychology*, **81**, 457–466.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction*, **12**, 185–233.

Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load and selective attention as factors in the structuring of technical material. *Journal of Experimental Psychology: General*, **119**, 176–192.

Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction*, **2**, 59–89.

Sweller, J., Mawer, R., & Ward, M. (1983). Development of expertise in mathematical problem solving. *Journal of Experimental Psychology; General*, **112**, 634–656.

Tarmizi, R., & Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology*, **80**, 424–436.

Van Merrienboer, J., & De Croock, M. (1992). Strategies for computer-based programming instruction: Program completion vs. program generation. *Journal of Educational Computing Research*, **8**, 365–394.

Ward, M., & Sweller, J. (1990). Structuring effective worked examples. *Cognition and Instruction*, **7**, 1–39.

Zhu, X., & Simon, H. (1987). Learning mathematics from examples and by doing. *Cognition and Instruction*, **4**, 137–166.