Research into Cognitive Load Theory and Instructional Design at UNSW

Dr. Graham Cooper
University of New South Wales, Australia

(December 1998)

E-mail address: g.cooper@unsw.edu.au

1: Overview

In recent years there has been an increased focus on the role of education and training, and on the effectiveness and efficiency of various instructional design strategies. Some of the most important breakthroughs in this regard have come from the discipline of Cognitive Science, which deals with the mental processes of learning, memory and problem solving.

Cognitive load theory (e.g. Sweller, 1988; 1994) is an instructional theory generated by this field of research. It describes learning structures in terms of an information processing system involving long term memory, which effectively stores all of our knowledge and skills on a more-or-less permanent basis and working memory, which performs the intellectual tasks associated with consciousness. Information may only be stored in long term memory after first being attended to, and processed by, working memory. Working memory, however, is extremely limited in both capacity and duration. These limitations will, under some conditions, impede learning.

The fundamental tenet of cognitive load theory is that the quality of instructional design will be raised if greater consideration is given to the role and limitations, of working memory. Since its conception in the early 1980's, cognitive load theory has been used to develop several instructional strategies which have been demonstrated empirically to be superior to those used conventionally.

This paper outlines some of the basic principles of cognitive load theory. Examples of the instructional design strategies generated by cognitive load theory are also provided.

2: Memory

2.1 Remembering information

Some people believe that we remember information by 'capturing' it on something like a video tape in our minds. This is not the case. What we see and remember depends more on what we already know, than on what is actually presented.

Look at each of the following, and note what you see.

In the first example most people read 'THE CAT', even though the centre symbol in each word is the same. The context of reading provides information which we use to help interpret the symbols.

In the second example most people will read each symbol as an example of the letter "a", even though no two symbols are identical. We can read an infinite range of symbols as the letter "a", even most peoples' hand writing, although we have never seen *their* handwriting before. We are able to do so because of our knowledge of what constitutes the letter "a".

Similarly, we are also able to recognise literally millions of different trees, as trees, even though no two are identical.

These examples demonstrate that we cannot help but to impose meaning on things that we sense. Humans are able to behave and think in 'intelligent' ways because of their ability to quickly identify meaning in presented stimuli.

Our knowledge and skills in activities as diverse as reading, driving, mathematics and gardening all derive from the knowledge base which we hold more-or-less permanently in long-term memory.

2.2 Chunking information

When presented a "large" set of elements to remember, it is often helpful to combine the elements to form a smaller number of groups. Each of the groups is referred to as a "chunk" of information.

For example, it is common practice to combine the digits of a phone number into two or three chunks of several digits each, rather than listing all digits in one long sequence. The phone number 3476 - 2980 may be easier to remember than the sequence 3 4 7 6 2 9 8 0.

Chunking does not need to be based upon any underlying meaning or logic that can be identified within the elements of the to-be-learned information. However, if an underlying meaning or logic *can* be identified and is used to define the chunks, then remembering is greatly enhanced.

For example, remembering a shopping list where elements are chunked into like groups, such as:

```
apples      bread     soap         pens
oranges     butter    shampoo      paper
bananas     jam       toothpaste   envelopes
```

is much easier to remember than a list of identical elements which are chunked into groups without any underlying structure, such as:

```
apples      soap      bread        bananas
shampoo     paper     oranges      toothpaste
pens        jam       butter       envelopes
```

Look at each of the following statements in turn for just a few seconds, and try to memorise the sequence of letters and spaces.

1. Allfi shenj oycle anwa ter

2. All fish enjoy clean water

The first statement is difficult to memorise. The series of letters and spaces appears to be random. If we are unable to identify any form of pattern or meaning then we are reduced to a strategy of memorising individual letters in turn. If, however, we are able to identify the "scrambled" meaning, then our strategy for remembering becomes one of trying to remember the location of the spaces.

The second statement is easy to memorise because the spaces are located in a way that promotes meaning. Consequently we need only memorise a few ideas (All fish, enjoy, clean water).

When what we already know enables us to identify or impose *meaning* on a new piece of information because it connects with information held in long-term memory, then it is relatively easy for us to remember it because we can "build it into" our existing knowledge base in a way that makes sense for us. The new information becomes an integral part of our overall knowledge, held in long-term memory.
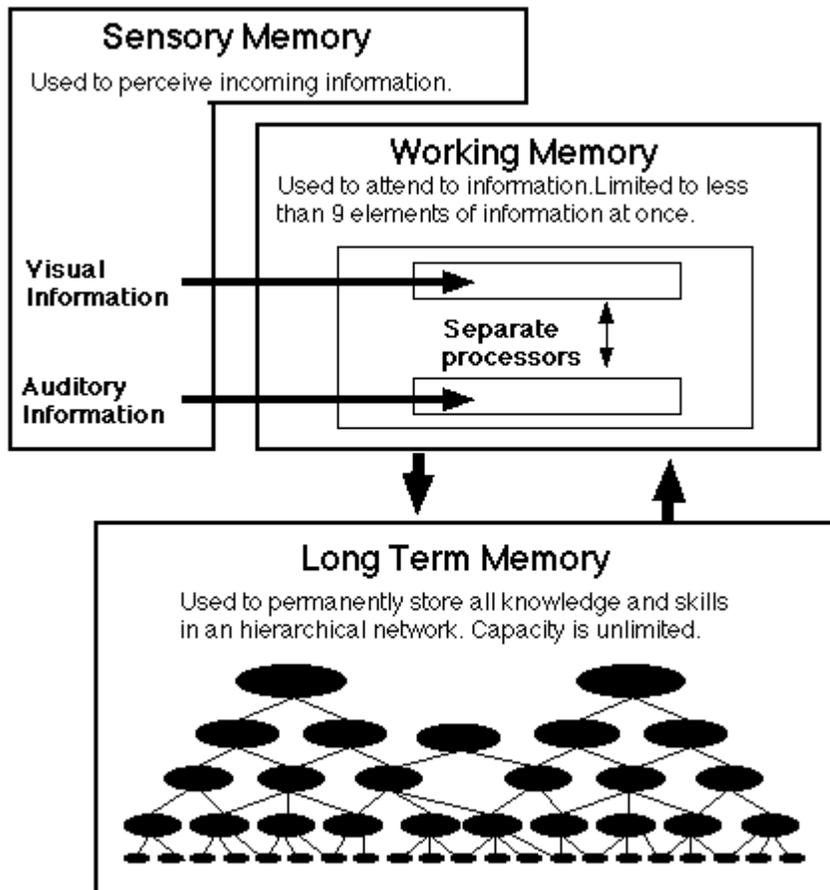
2.3 The modal model of memory

It is now widely accepted that we have, and use, more than one type of memory.

A modal model of memory distinguishes between three distinct memory types (modes). These are sensory memory, working memory and long term memory.

Each mode has its own characteristics and limitations.

These three modes are integrated to define an information processing model of human cognitive architecture.

Sensory Memory
Used to perceive incoming information.

Working Memory
Used to attend to information. Limited to less than 9 elements of information at once.

Visual Information

Separate processors

Auditory Information

Long Term Memory
Used to permanently store all knowledge and skills in an hierarchical network. Capacity is unlimited.

2.4 Sensory memory

Sensory memory deals with incoming stimuli from our senses. These are sights, sounds, smells, tastes and touches. A separate partition of sensory memory exists for each of the senses.

Sensory memories extinguish extremely quickly. (About half a second for visual information, 3 seconds for auditory information). In that time, we must identify, classify and assign meaning to the new information or it will be gone forever.

While looking at the picture below, quickly shut your eyes, and keep them shut for a few seconds. Repeat this several times.

As soon as you shut your eyes you may have noticed an image of the picture remaining for a split second "somewhere in your mind". This demonstrates the operation of the partition of your sensory memory that deals with visual perceptions. This is not restricted to blinking at pictures. Look at anything around you and it will still work.

2.5 Long term memory

Long term memory refers to the immense body of knowledge and skills that we hold in a more-or-less permanently accessible form.

Our name, date of birth, the letters of the alphabet, how to read, how to write, how to drive, swim, play chess, catch a ball and everything else that we "know" is all held in our long term memory awaiting activation.

Activation will occur as a direct result of our working memory querying long term memory for specific factual information (through our consciousness). Once a query has been made activation (and the 'answer') is effectively instantaneous.

Knowledge and skills that are activated with extremely high regularity, such as walking and talking, may be activated 'automatically' without the need for high levels of conscious attention, even though the task itself may be a complex one. (Automation is discussed further in Section 3.3.)

Consider each of the following questions.

Question 1: What is your name?

Question 2: What are the letters of the alphabet?

Question 3: Who won the lottery in 1992 at Wattle St., Sydney, Australia?

Question 1: What is your name?

You will be able to answer this quickly. It's no surprise since it is referred to frequently and consists of only a few words. Note how quickly you can provide the answer.

Question 2: What are the letters of the alphabet?

Again, you will be able to answer this quickly but this is a more interesting question than the first. Here there are 26 items in the answer and virtually everyone presents the 26 items in the same order. Our long term memory holds the letters of the alphabet in alphabetical sequence. If you try to say the letters of the alphabet in a random order, then you will find it an extremely difficult, probably impossible task.

Question 3: Who won the lottery in 1992 at Wattle St., Sydney, Australia?

Most people will quickly realise that they do not know the answer to this question. They recognise almost immediately that this is information that is not currently held in their long term memory. Generally, people "know that they don't know".

2.6 Working memory

Working memory is the part of our mind that provides our consciousness. It is the vehicle which enables us to think (both logically and creatively), to solve problems and to be expressive.

Working memory is intimately related to where and how we direct our attention to "think about something", or to process information.

The biggest limitation of working memory is its capacity to deal with no more than about seven elements of information simultaneously (Miller, 1956).

Working memory capacity may be expanded slightly by mixing the senses used to present information. That is, it is easier to attend to a body of information when some of the information is presented visually and the remainder of the information is presented auditorily than it is when all of the information is presented through a single sense (either all visually or all auditorily).

If the capacity of working memory is exceeded while processing a body of information then some, if not all, of that information will be lost.

Consider answering both of the following questions without using pencil and paper.

```
Question 1:          46+
                     37

Question 2:    83 468 446+
               93 849 937
               58 493 900
```

For most people Question 1 is quick and easy to solve as an example of mental arithmetic.

In many ways Question 2 is nothing more than a 'larger' version of Question 1, yet it is almost impossible to solve mentally.

The role of long term memory is effectively the same for these two questions (to recall the rules of addition).

The difference is that in Question 2 our working memory capacity is exceeded. It cannot cope with the large number of elements (in this case the numerals) that need to be attended to simultaneously in order

to solve this problem.

The use of pen and paper aids solution to Question 2 because it effectively relieves the burden placed upon working memory by giving us a means of recording elements in a 'permanent' form once we have finished processing them.

3: Learning

3.1 Definition of learning

Learning may be defined as the encoding (storage) of knowledge and/or skills into long term memory in such a way that the knowledge and skills may be recalled and applied at a later time on demand.

Humans have a great capacity for learning and tend to spend their lives doing so. They learn not only how to walk upright, but also how to talk, read and write. Many people today learn how to drive a car, operate a microwave oven, and use a computer. Some even learn how to perform a heart transplant operation. For all of these tasks (and just about every other task you care to mention) the role, capacity and qualities of sensory memory and working memory remain effectively unchanged. The driving force behind all skilled performance is the knowledge base that has been acquired within long term memory.

The capacity of our long term memory to acquire knowledge appears to be unlimited. No-one ever "runs out of space", although with age there may be an overall deterioration in the performance of our memory system.

It should also be noted that virtually everyone *can* learn how to drive a car, operate a microwave oven, use a computer or even perform a heart transplant operation, provided that they are given sufficient time and training to enable them to acquire the necessary knowledge and skills.

The next diagram presents part of an information network for cars for you to complete, or at least think about. There are no right or wrong answers.

Spend a few minutes writing down in point form some information about cars. Some ideas have been included for you to work from but you are free to add anything you like. For example, details about their use, cost, construction, road rules, impact on the environment, history of development, principles of combustion engines, how to change gears, how to replace spark plugs.....and so on.

Work quickly, writing down ideas as soon as they come to you. If you spend more than a few seconds "stuck", then begin another branch.
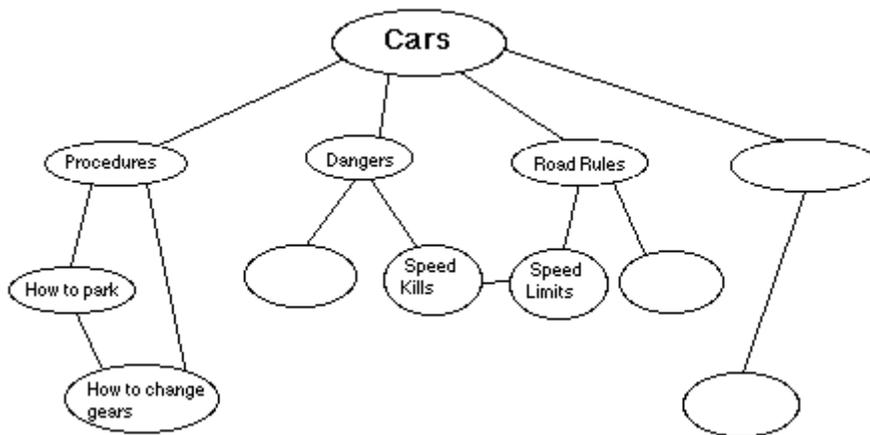
Everyone living in modern society holds an enormous amount of knowledge regarding cars, their use, road rules, and so on. This knowledge base is held in a well structured information network which is itself connected to other networks. Networks such as those for 'transport' or 'modern society' are higher order concepts, while networks for 'seat belt', 'spark plugs' and 'accelerator' are lower order concepts. Knowledge about procedures is also held (for example, how to park and how to change gears).

These hierarchical information networks are referred to as "schemas". Schemas build in detail and complexity as more extensive knowledge is acquired in a content area. The network in the diagram above is part of your schema for cars held in your long-term memory.

Individual differences exist in the nature and details of schemas. Someone who is employed as a mechanic and spends their pastime rebuilding vintage cars will have more detailed and complex schemas for cars than most people.

Schemas that are well learnt may be recalled and applied with relative ease. For example, someone learning to drive a manual car needs to concentrate intently on the knowledge and skills required to coordinate the movements of the clutch, gear stick and accelerator, in order to change gears smoothly. After several years of driving, however, most people are able to change gears "automatically". As automation develops, there is a reduction in the need for concentration.

3.2 Process of learning

The previous section (Section 3.1) argued that when we say that "something has been learnt", we mean that is has been successfully encoded into long term memory and can later be recalled on demand.

The next question to be considered is 'how does information become encoded into long term memory?' While the factors which contribute to encoding may vary from one situation to another, there is one factor that is always present. *To be encoded, information must first be attended to, and processed by, working memory* . If for any reason, working memory is unable to attend to a body of to-be-learnt information, then learning will be ineffective.

This has important implications for instructional design because the limitations of working memory may impede the learning process. This forms the basis of cognitive load theory.

Try to learn the following rhyme without paying attention to it.

..........Twinkle twinkle little star, how I wonder what you are.

In all likelihood you had not proceeded past the word 'star' before you became aware that you already knew this rhyme. Indeed, you could probably add a few more lines of the rhyme without difficulty.

You were instructed to learn this rhyme *without paying attention to it* . The fact that you "recognised" the rhyme as one that you already know, shows, however, that you did attend to the information. Perhaps you feel that this is due to the fact that a well known rhyme was used. Try the next one.

Try to learn the following rhyme without paying attention to it:

..........Emus and elephants into the stew, rub turns your until it tummy blue.

If you were aware of any grammatical problems with this statement, then you have again been paying attention to it. Once again, you could not help yourself.

Working memory, as the embodiment of our consciousness, cannot be "turned off" or "by passed" while we *are* conscious.

3.3 What a novice needs to learn to become an expert

For any given cognitive domain (algebra, crosswords, astrophysics, chess, electronics) we think of novices in that area as not knowing much and for their performance to be slow and error prone. In contrast we view experts as knowing almost everything and assume their performance to be quick and error free.

Contrary to popular belief, expertise does not appear to be due to anything as robust as "intelligence". Nor does it appear that experts are more "thoughtful" than novices.

The only two distinguishing features of expertise are:

1 . the expansive schemas (information networks) that experts hold, and

2 . the high level of automation (ability to perform tasks without concentrating) that experts exhibit.

Schemas and automation appear to explain *all* other expert/ novice differences.

Experts, because of their expansive set of schemas, have effectively seen almost every possible situation in the content domain before. Moreover, they have learnt what response is required for each situation and can carry out the required responses automatically, without the need for high levels of concentration. Experts are effectively just going through a set of routine exercises. It is no surprise then that experts are so fast and accurate in their performances.

Novices, on the other hand, have relatively few schemas. They have trouble recognising anything but the most basic and common situations as ones that they have encountered previously. Novices are presented with a "problem" almost every time they venture into the content domain (problem being defined as not knowing what to do or how to do it). Novices must "solve" almost every situation presented to them. To make matters worse, even when they realise what response is required, they may have difficulties in

performing the response. They need to concentrate intently if they are to avoid making errors.

Consider the task of reading this page of printed material. Presumably you may do so with little effort. While you need to concentrate on the *arguments* being presented, it is likely that you do not need to concentrate on the actual task of reading, that is, on the interpretation of all these squiggles which represent letters of the alphabet, which are sequenced to form words, which are sequenced to form sentences, and so on.

The task of reading is incredibly complex yet we use written documents as an easy and efficient way to communicate ideas.

All this changes, of course, if the reader is young (say five years old).

And what chance would a typical three year old have of reading even one line of this page? None.

Many three year olds can recite the letters of the alphabet. They can also identify many, if not all, of the letters in written form. However, lower case letters may present some difficulties, and running writing may be virtually impossible for them to handle. If a three year old can spell his or her name (or simple words like 'cat' or 'dog'), then the adults around the child express praise and encouragement.

By the age of five a child is likely to have developed more refined schemas for letter recognition, and perhaps even for recognition of some words (their name for example). However, there is a general absence of automation in reading. Reading is slow, error prone, and needs high levels of concentration (mental effort). It is likely that in "reading" the child will sometimes sound out the letters of each word in a sentence, but not actually comprehend the sentence as a whole. This is because their attention needs to be fully focussed on each word in isolation.

Contrast this to your reading skills. You no longer need to attend to individual letters or individual words. It is likely that you can process the text as quickly, if not faster, than you can say the material aloud. The only times that you need to slow your reading speed will be when reading becomes "difficult".

One source of difficulty lies in physical factors such as tiredness (low levels of attention), loud music (distractions), tiny text or poor lighting (inability to discriminate).

The more interesting difficulties, however, arise when something presented on paper fails to fit into your schemas and/or level of automation. Uncommon or technical words such as 'einstellung' or 'xanthoma', or misspelt words such es thiis werrd may cause your attention to be directed to the individual word, perhaps even the individual letters.

The irony about tasks such as walking, talking and reading is that they are among the most difficult that humans ever master, yet we are able to perform each of these with extremely low levels of mental effort. Our schemas in these areas have become so complete, and our level of automation so high, that we now find each of these tasks to be almost trivially easy.

A well known proverb states that "familiarity breeds contempt". In the context of education and training this should perhaps be modified to read "familiarity breeds expertise".

4: Cognitive Load Theory

4.1 Definition of cognitive load

Cognitive load refers to the total amount of mental activity imposed on working memory at an instance in time.

The major factor that contributes to cognitive load is the number of elements that need to be attended to.

Look at each of the following statements in turn for just a few seconds, and try to memorise the sequence of digits. Note that you do not need to remember all statements at once. Give all of your attention to each statement in turn.

Statement 1.   9 2

Statement 2.   7 9 5 3

Statement 3.   3 9 2 4 6 7 1 5

Statement 4.   3 9 2 5 9 7 4 6 2 8 4 6 7 1 5 6

For this activity we may use the number of digits (the elements) to be remembered as a simple measure of cognitive load. Consequently:

Statement 1: 9 2   has a cognitive load of 2

Statement 2: 7 9 5 3   has a cognitive load of 4

Statement 3   3 9 2 4 6 7 1 5   has a cognitive load of 8

Statement 4: 3 9 2 5 9 7 4 6 2 8 4 6 7 1 5 6 has a cognitive load of 16

Note that the measure used for cognitive load does not equate mathematically to task difficulty. That is, even though statement 2 has twice the number of digits as statement 1, it is almost as easy to remember.

In contrast, statement 4 has twice the number of digits as statement 3, yet seems more than twice as difficult to remember. While statement 3 can be remembered with effort, statement 4 is impossible for most people to remember without some form of practice or memory aid.

4.2 Reasons why some material is difficult to learn

The previous activity (Activity 4.1) used a digit span task to demonstrate that human working memory has a *threshold* of somewhere between 4 and 10 elements.

For the previous activity this means that:

1.....when the total number of digits to be remembered is four or less then the task is trivially easy for most people.

2..... when the total number of digits to be remembered is between five and nine then the task is achievable for most people if they exert 'some' mental effort.

3.....when the total number of digits to be remembered is ten or more then the task is difficult for most people.

In many ways, however, this task is artificial. People are rarely required to memorise sequences of random digits. After all, even telephone numbers and post codes may have an underlying logic.

Most of the information that we are required to learn in our lifetime is far more complex than a simple sequence of objects (whether they be digits in a telephone number, or items on a shopping list). Content areas such as mathematical calculus, biochemistry and computer programming are considered to be "difficult" to master. One of the reasons for this is undoubtedly the sheer volume of information that must be acquired (and built into schemas) before an expert knowledge base is held in the area. But there is another critically important quality that is evident in these content areas: that of 'high element interactivity'.

Element interactivity is defined as the degree to which the elements of some to-be-learned information can, or cannot, be understood in isolation. While the nature of element interactivity is difficult (and often subtle) to comprehend, a simple example may assist in describing this concept.

Example 4.2 - *Element Interactivity*

Consider the task of learning a foreign language. Most people can quickly learn some simple, everyday words, but will have difficulty in generating grammatically correct sentences, even when all of the words used in the sentence are known.

Vocabulary is an example of low element interactive material. Although there may be literally thousands of words to be learnt, most words may be learnt in isolation to all of the other words.

To build sentences that are grammatically correct, however, one must attend to *all of the words within the sentence at once* while also considering syntax, tense, verb endings and so on. Grammar is an example of high element interactive material because to learn it, many elements must be considered simultaneously.

Determine if either of the following statements could be true.

.....1. My fathers' brothers' grandfather is my grandfathers' brothers' son.

.....2. My fathers' brothers' grandfather is my grandfathers' brothers' father.

Although each of these statements requires only a few elements (people) to be considered, the activity is extremely difficult because there is a need to also attend to the relationships *between* the elements. This is an example of "complex" information where elements interact with each other. As a consequence of the high element interactivity, the cognitive load induced exceeds the resources of working memory.

The cognitive load associated with this material can be greatly reduced if the information is presented pictorially. Elements which interact with each other often have the potential to be presented in pictorial form, where the picture itself holds (and conveys) some of the information, reducing the need for it to be held in working memory.

The partial family tree presented below shows that statement 2 is logically possible.
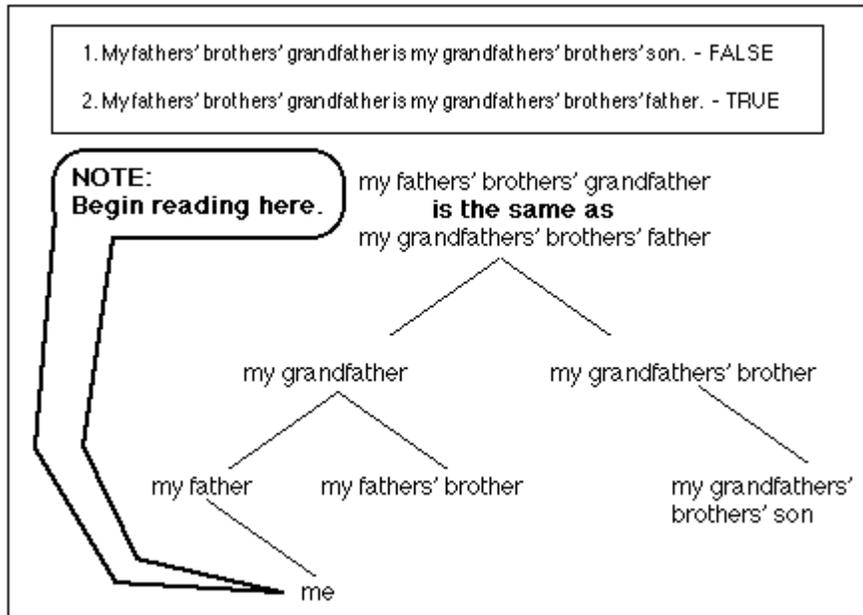
4.3 Elements held in working memory are schemas

This paper has argued that the limited resources of working memory mean that only a few elements of information may be attended to at any given time.

The previous section (Section 4.2) demonstrated that to-be-learned information which has a high level of element interactivity imposes a cognitive load over and above that imposed by the elements themselves, due to the need to attend also to the relationships between elements. Consequently, high element interactive material exacerbates the difficulties which result from working memory limitations.

All of this begs the question "what is an element?" The short answer is "that it depends". It depends on the schemas held by the person who is required to attend to some body of to-be-learned information because generally, elements are schemas. What is a single element consisting of a single schema for an expert may be several elements consisting of sub-schemas for a novice.

Consider again the contrast between statements of the type represented by 1. and 2. below.

    1.   hea cto egttch. edh sda

    2.   The dog chased the cat.

The first statement presents itself as a random sequence of letters and spaces. It is without meaning and consequently each letter and each space is a separate element which working memory needs to attend to.

In contrast, the second statement contains obvious meaning. Each cluster of letters forms a meaningful word, and the words combine to form a meaningful sentence. Here the number of elements for an expert reader, who knows a little about the behaviour of dogs and cats, may be as few as one. After all, it is a grammatically correct sentence, and it is well known that dogs do chase cats.

Schemas not only provide the ability to combine 'many elements' into a single element. They also have the capacity to incorporate the interactions between elements. This means that information which consists of several elements, all of which interact with one another, may be embodied into a single schema.

For example, a professional fibre glasser holds a schema for 'mixing resin' which takes into account not only the ideal ratio of resin and catalyst that need to be mixed, but also, automatically, considers interacting factors such as the air temperature, air moisture, and purpose of the mixture. It is likely that a novice in this area would not even know that if environmental factors such as temperature and moisture are not taken into account, then a defective mixture may result.

4.4 Intrinsic and extraneous cognitive load
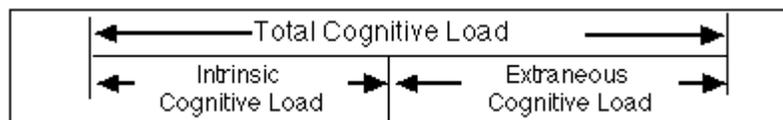
Intrinsic cognitive load
Intrinsic cognitive load is due solely to the intrinsic nature (difficulty) of some to-be-learned content. Intrinsic cognitive load cannot be modified by instructional design. For example, content which is high in element interactivity remains high in element interactivity regardless of how it is presented.
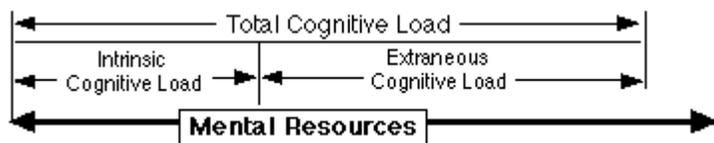
Extraneous cognitive load
Extraneous cognitive load is due to the instructional materials used to present information to students. Teaching materials addressing a concept such as continental drift, for example, will be more effective if it makes an appropriate use of graphics rather than a text only presentation.
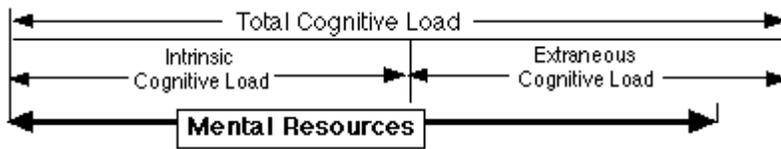
By changing the instructional materials presented to students, the level of extraneous cognitive load may be modified. This may facilitate learning.

Demonstration 4.4

1. When intrinsic cognitive load is low (simple content) sufficient mental resources may remain to enable a learner to learn from "any" type of instructional material, even that which imposes a high level of extraneous cognitive load.

2. If the intrinsic cognitive load is high (difficult content) and the extraneous cognitive load is also high, then total cognitive load will exceed mental resources and learning may fail to occur.

3. Modifying the instructional materials to engineer a lower level of extraneous cognitive load will facilitate learning if the resulting total cognitive load falls to a level that is within the bounds of mental resources.
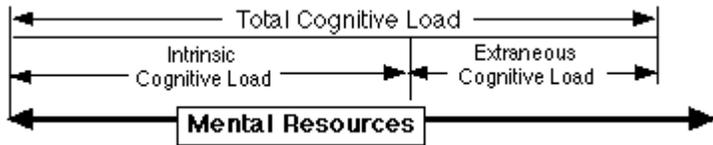
4.5 Principles of cognitive load theory

Cognitive load theory focuses on the role of working memory in the learning process.

The fundamental principles of cognitive load theory rest upon the following argument.

1. Working memory is extremely limited.

2. Long term memory is essentially unlimited.

3. The process of learning requires working memory to be actively engaged in the comprehension (and processing) of instructional material to encode to-be-learned information into long term memory.

4. If the resources of working memory are exceeded then learning will be ineffective.

4.6 Applying cognitive load theory to instructional design

The fundamental principles of applying cognitive load theory to instructional design rest upon the following argument.

1. Excessively high levels of cognitive load may result directly from the instructional materials presented to students.

2. Redesigning instructional materials to reduce the levels of extraneous cognitive load may enhance learning.

3. Content areas that are most likely to demonstrate beneficial results from improved instructional design are those that deal with "complex" information where the elements of to-be-learned information interact with one another (therefore imposing a high level of intrinsic cognitive load).

Summary 4.6 - *Applying cognitive load theory to instructional design*

Cognitive load theory states that learning will be maximised by ensuring that as much of a learners' working memory as possible is free to attend solely to encoding to-be-learned information.

5: Effects Generated by Cognitive Load Theory

5.1 The different effects

Cognitive load theory has been used successfully to develop several instructional techniques which facilitate learning.

These include:

.....the goal free effect

.....the worked example and problem completion effect

.....the split attention effect

.....the redundancy effect

.....the modality effect.

5.2 Benefits for learning

Each of the effects listed above in Section 5.1 has been shown empirically to provide strong benefits to learners when used appropriately. In each case the benefits include *all* of the following:

.....reduced training time

.....enhanced performance* on test problems (similar to those seen during training)

.....enhanced performance* on transfer problems (those which are dissimilar to problems seen during training but requiring the same rules for solution).

Note :
Enhanced performance* means both shorter times to complete problems, and fewer errors.

The fact that students spend less time learning, yet return superior performances when tested, is a powerful finding that has considerable implications for education and training.

Of special importance is the increased performance on transfer problems. This shows that the learning which results from each of these effects is at a level of true understanding that enables students to solve

a wider range of problems than those students taught using "conventional" instructional materials.

5.3 Generating a measurable effect

Each effect has been developed by the argument that engineering a cognitive load which falls within the limitations of working memory facilitates learning.

The specifics which determine when and how each of the effects operate for a given set of learners on a given set of to-be-learned content, may be found in the original research papers.

Of particular importance to the successful generation of these effects is the expertise of the learner relative to the to-be-learned information.

When learners hold high levels of expertise in the content area then the elements which their working memory may attend to are each, *in and of themselves* , large complex knowledge networks (high level schemas). Consequently, their working memory need only consider a few elements in order to hold all of the to-be-learned information in mind. Ample cognitive resources thus remain for the process of learning. Instructional design manipulations for this group of learners will be ineffective because their working memory capacity is not being exceeded.

In contrast, when learners hold a low level of expertise in the content area then only simple elements (low level schemas) have been acquired (perhaps almost none in the case of true novices). Consequently, working memory needs to attend to many elements in order to hold all of the to-be-learned information in mind. Here, cognitive resources are stretched beyond their capacity and insufficient cognitive resources remain for the process of learning. Instructional design manipulations for this group of learners will be effective if the reduction of cognitive load results in a level that is within the capacity of working memory.

The dynamics of generating any of the effects thus depends on obtaining a group of students whose relative level of expertise to content difficulty is ripe for instructional design manipulations. (See Cooper & Sweller, 1987, for details on how student ability impacts upon the generation of a measurable effect.)

5.4 Conventional problems

Before presenting information detailing the effects generated by cognitive load theory a brief overview describing conventional problems, and the process by which novices solve conventional problems, will be presented. This is because both the goal free effect and the worked example effect are based upon the finding that the method employed by novices to solve conventional problems (means-ends analysis, which is discussed in the next section) imposes a relatively high level of cognitive load (Sweller, 1988).

Conventional problems are those which present students with a set of given data (the known information) and a well defined goal (specifies what needs to be found). Moreover, the answer may be objectively determined to be correct or incorrect by applying rules (such as formulae) in an algorithm based sequence.

Conventional problems are typically found in all topic areas of mathematics and science, and in all

subject areas that make use of mathematical principles (for example engineering, accountancy and computer programming).
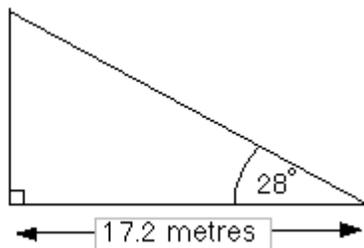
Example 5.4.a

If y = x + 6, x = z + 3, and z = 6, *find the value of y* .

Example 5.4.b

A particle starts from rest and is accelerated at 12 m/s2 for 4.5 seconds.
What is its terminal velocity?

Example 5.4.c

For the right triangle shown, *determine the length of the hypotenuse* .

5.5 Using means-ends analysis to solve problems

Means-ends analysis is a problem solving heuristic (strategy) which is widely used to solve conventional problems by people who are not highly familiar with the specific problem type (Larkin, McDermott, Simon & Simon, 1980; Simon & Simon, 1978).

Means-ends analysis is based upon the principle of reducing differences between the current problem state (which begins at the problem givens) and the goal state. In practice, this procedure often results in a problem solver working backwards from the goal to the problem givens, before then working forwards from the givens to the goal.

While this strategy is very effective in obtaining answers (assigning a value to a goal state) it has a necessary consequence of inducing very high levels of cognitive load. This is because the nature of the strategy requires attention to be directed simultaneously to the current state, the goal state, differences between them, procedures to reduce those differences and any possible subgoals that may lead to solution. Full details of how means-ends analysis operates, and its consequences for working memory, are presented in Sweller (1988).

Example 5.5

If y = x + 6, x = z + 3, and z = 6, *find the value of y* .

A novice problem solver (using means-ends analysis) would first focus on the goal state (find the value of y).

Rereading the question s/he would note that the value of "y" is provided by the equation "y = x + 6", so finding the value of "x" becomes a subgoal.

Similarly, a further rereading of the question would show that the value of "x" is provided by the equation "x = z + 3", so finding the value of "z" becomes a subgoal also.

Rereading the question yet again s/he would identify that the value of z is provided as given information (z = 6). This value may now be substituted into the equation "x = z + 3" to obtain the value "x = 9".

A true novice at this point may forget why the value of "x" was required. After all, their working memory has been heavily taxed attending to many elements of the problem.

Nevertheless, he or she will eventually identify that the value of "x" was calculated so that it could be substituted into the equation "y = x + 6". Doing so yields the value of "y = 15", which is the goal state.

As can be seen by this example (and this is just a simple problem), means-ends analysis is very cumbersome, and requires large amounts of cognitive resources for the strategy to be implemented successfully. Problem solvers using means-ends analysis may successfully solve "many" problems of an identical type, yet effectively learn nothing from the activity (Sweller & Levine, 1982)

5.6 The goal free effect

Means-ends analysis operates on the principle of reducing differences between the goal state and problem givens. Consequently, means-ends analysis may be rendered inoperable by redefining the problem goal so that no obvious goal exists (for example, "find what you can"). This is the principle behind the generation of goal free problems.

If problems are "goal free" then a problem solver has little option but to focus on the information provided (the given data) and to use it where ever possible. This automatically induces a forwards working solution path similar to that generated by expert problem solvers. Such forward working solutions impose very low levels of cognitive load and facilitate learning (Owen and Sweller, 1985; Ayres, 1993)

Example 5.6.a

If y = x + 6, x = z + 3, and z = 6, *find what you can.*

Attention would focus on "z = 6" as this is the only variable specified as a numerical value.

Rereading the question it would be identified that the value of "z = 6" can be substituted into the equation "x = z + 3". Doing so provides "x = 9".

Rereading the question it would now be identified that the value of "x =9" can be substituted into the equation "y = x + 6". Doing so provides "y = 15".
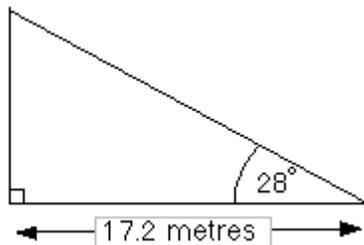
Nothing else remains to be found.

It can be seen that this solution path is far simpler than that generated by means-ends analysis in Example 5.5.

Example 5.6.b

A particle starts from rest and is accelerated at 12 m/s2 for 4.5 seconds.
Find what you can.

Example 5.6.c

For the right triangle shown,
Find what you can .

5.7 The worked example and problem completion effect

Historically subjects such as mathematics and science have been taught using the following general technique:

Step 1 :Introduce a new topic. Present background knowledge, principles and rules.

Step 2 :Demonstrate, using a *few* worked examples, how to apply the principles and rules.

Step 3 :Have the students "practice" how to apply the principles and rules by solving *many* , conventional, goal specific problems.

Section 5.5 described how the use of means-ends analysis to solve conventional problems imposes high levels of cognitive load, and thus impedes learning. It is therefore likely that the emphasis given to "practice problems" described above will not result in efficient learning.

While the use of goal free problems provides an effective alternative to conventional problem solving its application is limited to situations where the problem space is "small". As the size of the problem space becomes "large" the increasing number of alternatives faced at each step in a solution render the technique impractical for teaching purposes.

An alternative technique may be found in reconsidering the nature and purpose of worked examples. Worked examples are presented to students to show them directly, step by step, the procedures required to solve different problem types. Worked examples contain explicit information that equates to schemas and automation.

That is, worked examples promote the acquisition of knowledge and skills required to:
.....identify problems as being of a particular type,
.....recall the steps (in sequence) needed to solve each particular type, and
.....perform each step without error.

Studying worked examples imposes a low level of cognitive load because attention need only be given to two problem states at a time and the transformation (rule operator) that links them.

A successful method for placing emphasis on worked examples is to present them with conventional problems in an alternating sequence (example type A, problem type A, example type B, problem type B and so on). Students are informed of the paired nature of the material and instructed to study each example closely because they will not be allowed to look back at it once they begin the associated problem.

Students thus focus their attention on the problem type and the associated steps to solution (the schemas). In solving the associated conventional problem they are testing themselves to determine if they have learnt the procedure. This may be a more genuine form of "practice problem solving".

Example 5.7 - *A worked example format for teaching algebra.*

Following the numbered sequence, first study the worked example, then cover it, and attempt to solve the associated problem.

For each of the following, solve for 'a'.

1) $\quad c(a+b) = f$
$\quad\quad (a+b) = f/c$
$\quad\quad\quad a = f/c - b$

2) $\quad g(a+m) = k$

3) $\quad ac + e = g$
$\quad\quad ac = g - e$
$\quad\quad\ a = \dfrac{g - e}{c}$

4) $\quad ab + f = h$

The **problem completion** procedure has a similar rationale and effect to the use of worked examples (see Paas, 1992; Van Merrienboer and Krammer, 1987). Instead of providing an entire worked example followed by a problem, students are just provided with partially completed worked examples. For instance, in example 1 above, they may be provided with the first two lines and required to complete the third line themselves.

Discussion 5.7

The specific details regarding the number of example-problem pairings or completion problems to present, the range of examples to present, the rate at which the orbit of problem type is increased and so on, depends on the complexity of the material relative to the expertise of the learners. The greater the relative expertise, the quicker the pace of increase in problem types.

Worked example techniques have been demonstrated to be highly effective at facilitating learning across a wide range of mathematically based content (see Cooper and Sweller, 1987; Zhu and Simon,1987; Pass and Van Merrienboer, 1994).

5.8 The split attention effect

Many instructional materials require both a pictorial component and a textual component of information.

Conventionally a graphic has been presented with the associated text above, below, or at the side. Such instructional presentations introduce a split attention effect where the student needs to attend to both the graphic and the text. Neither the graphic, nor the text, alone, provide sufficient information to enable understanding. The instructional material can only be understood after the student has mentally integrated the multiple sources of information. The portion of working memory that needs to be used in integrating the graphic and text is unavailable for the learning process. Consequently learning is ineffective.

Consider this conventional mathematics based example, taken from Sweller, Chandler, Tierney & Cooper (1990).

Example 5.8.a *- Split instructional format for teaching co-ordinate geometry.*
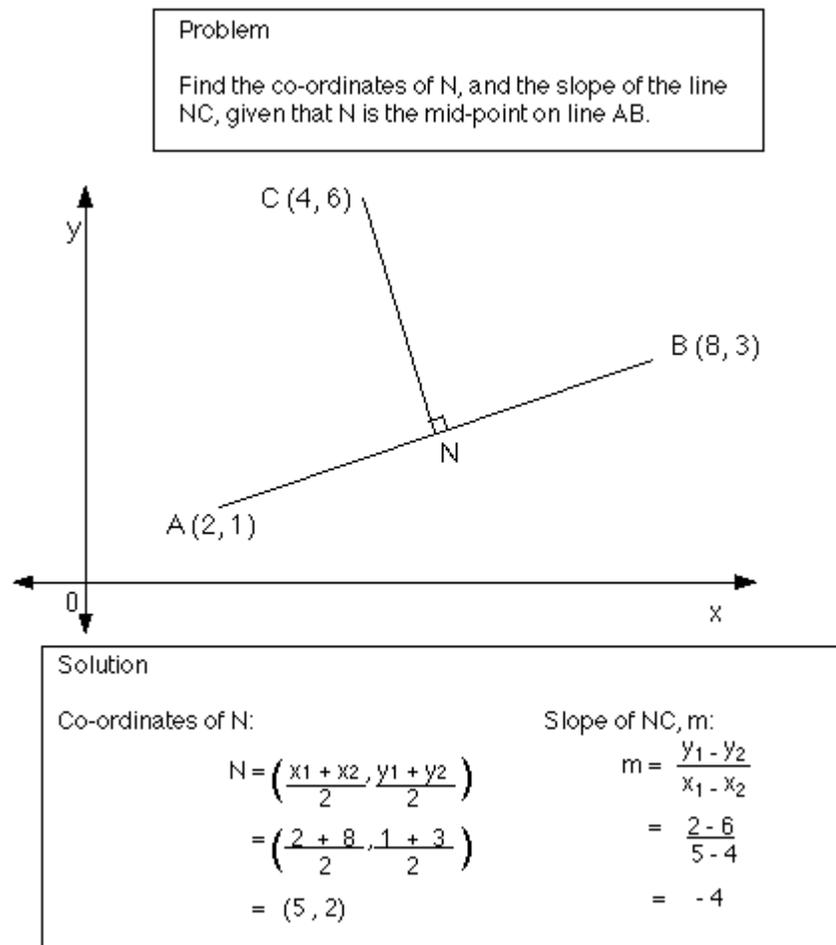
The presentation may be restructured to improve learning by physically integrating the solution *into the graphic* to produce a single source of instructional information. This eliminates the need to split attention between the graphic and the text. The association between the text and the graphic is clearly indicated.

Example 5.8.b *- Integrated instructional format for co-ordinate geometry.*

Problem

Find the co-ordinates of N, and the slope of the line NC, given that N is the mid-point on line AB.

Step 2
Slope of NC, m:

$$m = \frac{y_1 - y_2}{x_1 - x_2}$$

$$= \frac{2 - 6}{5 - 4}$$

$$= -4$$

Step 1
Co-ordinates of N:

$$N = \left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2} \right)$$

$$= \left(\frac{2 + 8}{2}, \frac{1 + 3}{2} \right)$$

$$= (5, 2)$$

C (4, 6)

B (8, 3)

A (2, 1)
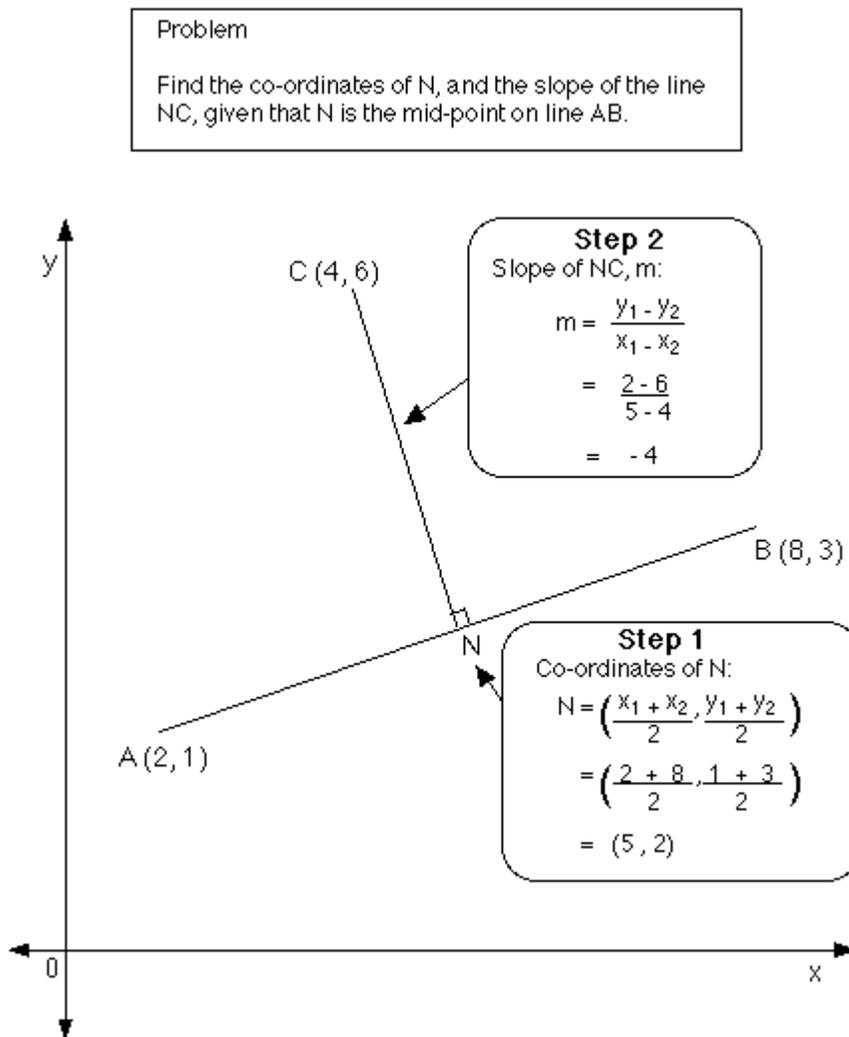
The split attention effect is not limited to worked examples in mathematics. It is demonstratable in all contexts where a graphical and a textual presentation are both necessary to impart meaning.

Consider the instructional material presented below dealing with electrical testing. (Taken from Chandler & Sweller, 1991)

Example 5.8.c -*Split instructional format for teaching a procedure.*

INSULATION RESISTANCE TESTS

a) CONDUCTORS IN PERMANENT WIRING

Test : To test Insulation Resistance from conductors to earth.

How conducted : i) Disconnect appliances and busways during these tests. Make sure mainswitch is "on" and all fuses are "in". Remove main earth from neutral bar and set meter to read insulation. Connect one lead to earth wire at MEN bar and take first measure by connecting the other lead to the active. Take next measure by connecting the lead to the neutral.ii) If resistance is not high enough in either of the two tests in i) then measure each circuit separately.

Results required :
i) At least One Megaohm
ii) Same result as i) above
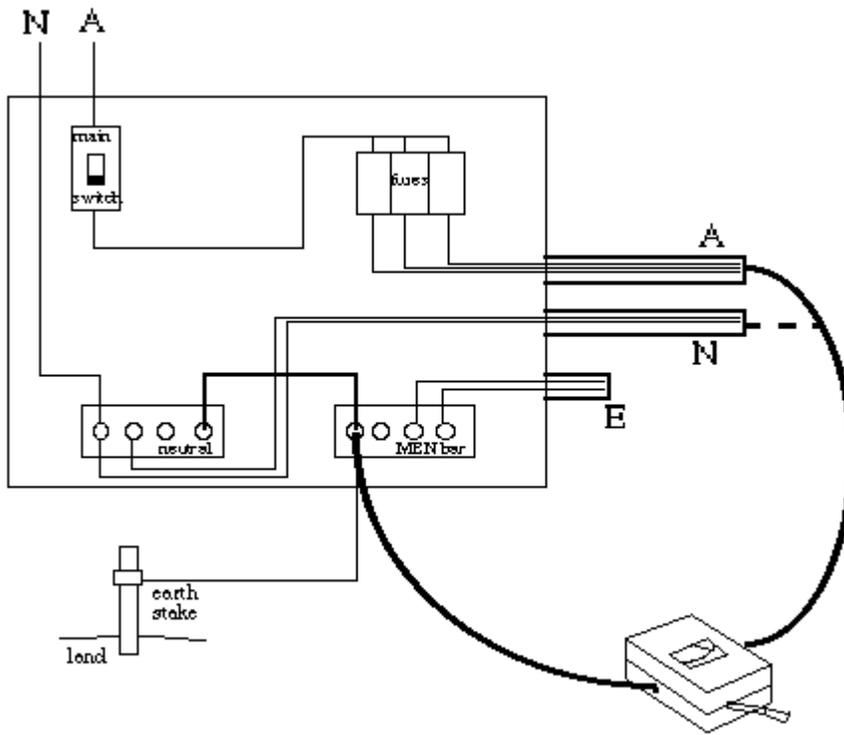
Again, by reformatting the material so that the instructions are integrated into the graphic, learning is enhanced. In fact, in this study, evidence indicated better performance resulted on both theoretical *and* practical tests.

Example 5.8.d - *Integrated instructional format for teaching a procedure.*

INSULATION RESISTANCE TESTS

a) CONDUCTORS IN PERMANENT WIRING
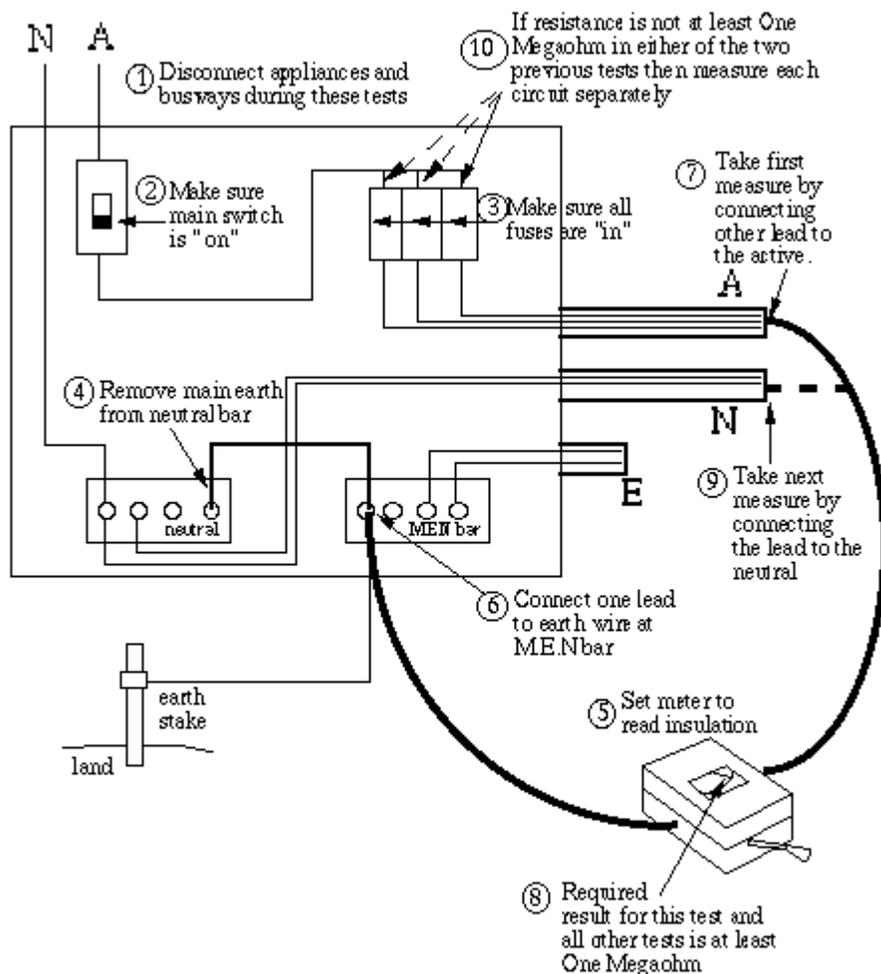
N   A

① Disconnect appliances and busways during these tests

⑩ If resistance is not at least One Megaohm in either of the two previous tests then measure each circuit separately

② Make sure main switch is "on"

③ Make sure all fuses are "in"

⑦ Take first measure by connecting other lead to the active.

A

N

④ Remove main earth from neutral bar

neutral

MEN bar

E

⑨ Take next measure by connecting the lead to the neutral

⑥ Connect one lead to earth wire at M.E.N bar

earth stake

land

⑤ Set meter to read insulation

⑧ Required result for this test and all other tests is at least One Megaohm

5.9 Sources of split attention

The examples presented in Section 5.8 (the split attention effect) focussed on the need to eliminate split attention effects which result from separate textual and graphical components of instructional materials. Appropriately integrating the text into the graphic facilitates learning.

Split attention, however, will result *whenever* a learner needs to simultaneously attend to two or more sources of instruction or activities.

Multiple sources of purely text based instructional materials will induce a split attention effect if two or more sources must be considered simultaneously. For example, this is likely to occur when cross referencing documents, or even cross referencing within a single document.

Chandler and Sweller (1992) provided evidence that a split attention effect occurs when reading conventional experimental papers because the results section and the discussion section are reported separately, yet need to be considered simultaneously to understand the complex of results and their implications. Here the split attention effect may be eliminated and intelligibility increased, by restructuring experimental papers to integrate the results and discussion sections.

A split attention effect may also result from mixing activities. For example, when learning to use a

software package it is common practice for the learner to simultaneously refer to a hard copy tutorial (or manual) *and* the computer. The tutorial provides step-by-step instructions for performing each task and the learner attempts to carry out each step on the computer. While this may seem to be an obvious way of learning a software package, experimental investigations have shown that far more effective learning strategies are available.

The simplest modification is to eliminate the use of the computer in the learning phase and replace it by appropriate pictures and diagrams. Provided the manual contains all of the relevant information, then students who study the manual alone outperform students who perform each step in sequence on the computer based upon the manual instructions. The irony here is that the manual-only-group complete their "training" without ever having used the software package, yet in testing, on a computer with the real software, they perform better than the group who has already spent time using the software package. See Chandler and Sweller (1996) for details.

Another alternative is to develop a computer based training package which integrates text based instructions into a computer simulation of the target computer package. When this is done the manual may be eliminated from the training process, leaving students to focus their attention wholly on the computer screen. This eliminates split attention and facilitates learning. See Cerpa, Chandler & Sweller (1996).

Summary 5.9 - *Sources of split attention*

Split attention occurs whenever a learner needs to attend to more than one source of information, or more than one activity. A common source of split attention is the need for a learner to perform a search. Searching a graphic to locate a component, searching a document to find a reference and searching software pull-down menus to find a function referred to in a manual are all examples of split attention.

Redesigning instructional materials to eliminate search and other sources of split attention facilitates learning.

5.10 The redundancy effect

Sections 5.8 and 5.9 described the benefits which result from integrating mutually referring textual and graphical sources of instruction.

Caution needs to be exercised, however, to ensure that both sources of instruction truly are necessary for the to-be-learned information to be intelligible.

In situations where a source of textual instruction, or a source of graphical instruction *alone* provides full intelligibility then only one source of instruction should be used (either the textual or the graphical), and the other source, which is redundant, should be removed completely from the instructional materials. In these contexts a single source of instruction returns higher levels of learning than either an integrated format (text integrated into the graphic), or a dual format (both text and graphic presented in parallel).

Cognitive load theory explains this result by focussing on the levels of cognitive load imposed upon the learner who needs to process the varying instructional materials.
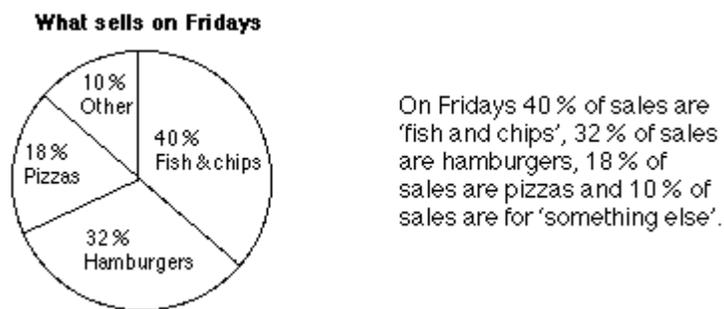
Attending to both textual and graphical sources of instruction requires more mental resources than

attending to a single source. Attending to both textual and graphical sources of instruction, therefore, results in a reduced portion of working memory being available for the process of learning.

Maps, whether their purpose is to locate countries (an atlas), indicate the steepness of terrain (a topographic map) or to show the way to get from A to B (a street directory) are examples of graphically based sources of instruction that are fully self contained. Provided the user has the skills to read and interpret a map, then there is no need for any associated body of textual information.
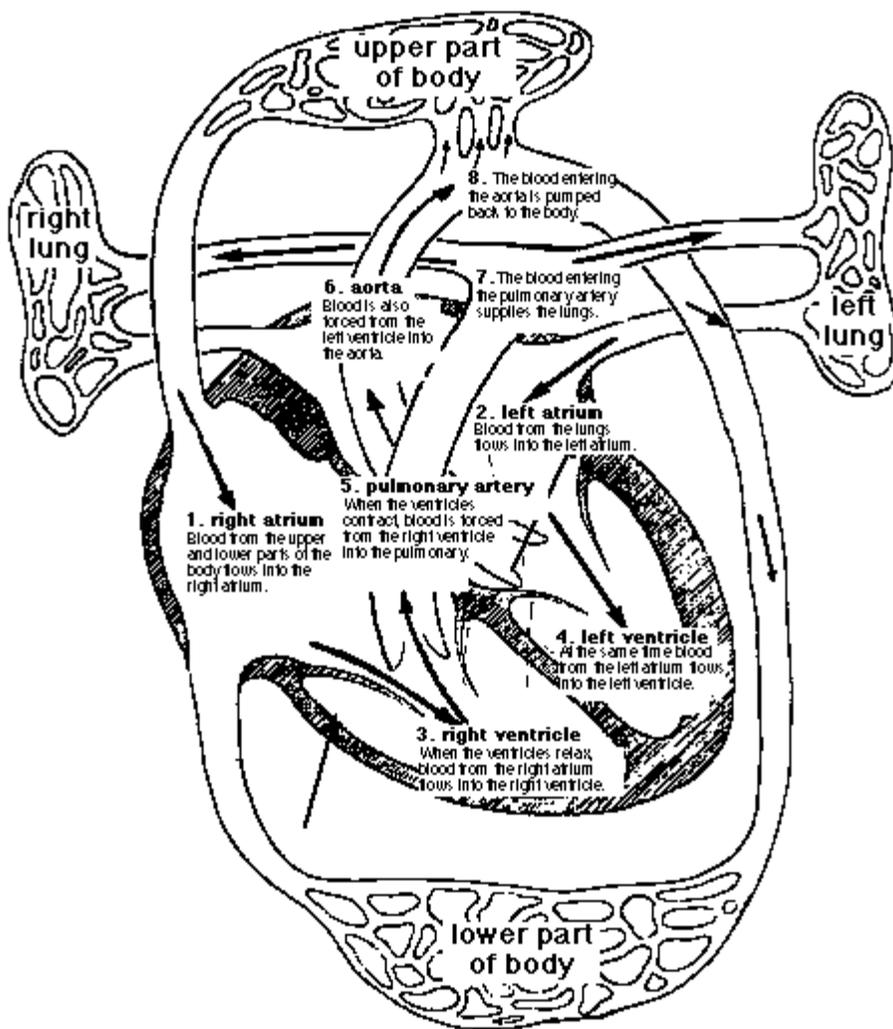
Similarly, many instances of textual instruction have no need for graphics. Arguments of litigation, analysis of history and the use of a dictionary or thesaurus are fully intelligible in a text-only format. The use of graphics in these situations actually reduces the level of learning that results from the use of these documents.

Example 5.10.a - *Redundant textual information in a dual format*

What sells on Fridays

On Fridays 40 % of sales are 'fish and chips', 32 % of sales are hamburgers, 18 % of sales are pizzas and 10 % of sales are for 'something else'.

Example 5.10.b - *Redundant textual information in an integrated format*

This example, dealing with the functions of the heart, is taken from Chandler and Sweller (1991).

upper part
of body

right lung

left lung

8. The blood entering
the aorta is pumped
back to the body.

6. aorta
Blood is also
forced from the
left ventricle into
the aorta.

7. The blood entering
the pulmonary artery
supplies the lungs.

2. left atrium
Blood from the lungs
flows into the left atrium.

5. pulmonary artery
When the ventricles
contract, blood is forced
from the right ventricle
into the pulmonary.

1. right atrium
Blood from the upper
and lower parts of the
body flows into the
right atrium.

4. left ventricle
At the same time blood
from the left atrium flows
into the left ventricle.

3. right ventricle
When the ventricles relax,
blood from the right atrium
flows into the right ventricle.

lower part
of body

Note

The graphic contains labels to indicate parts of the heart, and arrows to indicate the flow of blood.

The textual statements which are integrated into the graphic do nothing other than restate the parts and the flow of blood. On this basis the textual statements are redundant and should be deleted from the instructional material.

Students presented a graphic only instructional format learn more than students presented either an integrated format or a dual format.

5.11 The modality effect

All of the effects discussed so far in this paper have emphasised the need to reduce cognitive load because of the limitations of working memory.

While information processing models of learning have historically emphasised the "fixed" limits of working memory, there is evidence (Pavio, 1990; Baddeley, 1992) that under some conditions, an

expansion of working memory may be achieved.

Consequently, rather than attempting to reduce cognitive load, an alternative strategy, that of expanding working memory, may be pursued as a means of facilitating learning.

The work by Pavio and Baddeley indicates that at least some portions of working memory appear to be sensory mode specific. That is, some portion of working memory is dedicated to attending to visual information only (especially diagrammatic information) and some other portion of working memory is dedicated to attending to aural information only (especially verbal information). (Note, however, that the majority of working memory appears to be in the form of a central resource which may be allocated to any type of sensory information.)

Partitioning to-be-learned information so that some information, such as graphics, is presented visually, while other information, such as text, is presented auditorily enhances learning (see Mousavi, Low and Sweller, 1995; Jeung, Chandler and Sweller, 1997; Tindall-Ford, Chandler and Sweller, 1997). The modality effect holds the potential to impact upon the multi media industry.

Example 5.11 - *Mixed mode instructional format*

This example is taken from Jeung, Chandler and Sweller (1997).

Notes:
1. The graphic is presented visually but the text is *only* presented auditorily.
2. Screen highlights (flashing) were used to identify the components of the graphic referred to by each auditory statement to eliminate screen search.

When two parallel lines intersect with a third line, four pairs of corresponding angles are equal. In the diagram, two parallel lines, AB and CD, intersect with a third line, XY. The following four pairs of angles are corresponding angles:
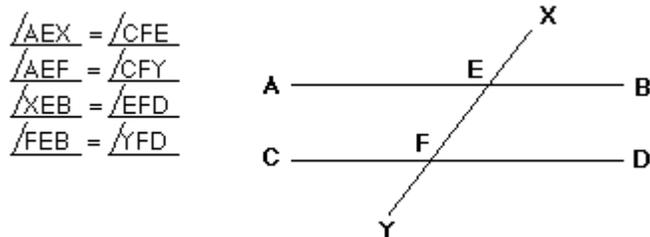
Section 6: Summary and Discussion

Cognitive load theory displays strong consistencies with current knowledge regarding memory, thought, learning and problem solving.

It is a theory which views the limitations of working memory to be the primary impediment to learning. Reducing total cognitive load imposed by a body of to-be-learned information increases the portion of working memory which is available to attend to the learning process. This may only be achieved by

engineering reduced levels of extraneous cognitive load through instructional design.

It is interesting (and important) to note that the effects generated by cognitive load theory often "fly in the face" of standard practices. This attests to the strength of the theory. The table below outlines this observation.

Standard practice	Cognitive load generated effect
1. Use conventional problems which specify the goal so that students 'know what they have to find'.	The goal free effect Use goal free problems.
2. Students need to solve many problems to learn because 'practice makes perfect'.	The worked example effect Students learn by studying worked examples. Problem solving is used to test if learning has been effective.
3. Instructional materials which require both textual and graphical sources of instruction should be presented in a 'neat and tidy' fashion where the text and graphics are located separately.	The split attention effect Instructional materials which require both textual and graphical sources of instruction should integrate the text into the graphic in such a way that the relationships between textual components and graphical components are clearly indicated.
4. The same information should be presented in several different ways at the same time.	The redundancy effect Simultaneous presentations of similar (redundant) content must be avoided.
5. Similar to-be-learned information should be presented using an identical media format to ensure consistency in the instructional presentation.	The modality effect Mix media, so that some to-be-learned information is presented visually, while the remainder is presented auditorily.

The effects generated by cognitive load theory should be viewed as "rules of thumb" rather than absolute "laws of instruction". The bottom line, according to cognitive load theory, will always be the need to reduce total cognitive load, and the need to maximise cognitive resources available to be utilised in the learning process. If for some reason cognitive load increases rather than decreases, then learning will be inhibited.

For example, the worked example effect will not occur if the examples used actually increase, rather than decrease, extraneous cognitive load. This is the case for examples which impose a split attention effect. Redesigning the format of the examples to eliminate split attention returns the educational benefit of the use of the worked examples.

The success of cognitive load theory in developing strategies and techniques which result in both reduced training times and enhanced performance is of paramount importance to the education and training industries.

Any fears that the application of instructional design techniques generated by cognitive load theory may result in a "poorer" quality of student or worker who is less able to think and act independently in

unusual or unforeseen situations are totally unfounded. Over the last ten years a large body of evidence has been acquired to show that students taught using cognitive load generated materials are actually more able to deal with such unusual or unforeseen situations as attested to by their superior performances on transfer problems (those that differ to problems seen during training, but requiring similar rules for their solution).

It should also be noted that current research projects have provided preliminary evidence for four additional effects generated by the application of cognitive load theory. These are (1) the procedural learning effect, (2) the imagination effect, (3) the colour coding effect and (4) the interaction effect. These effects are not discussed in the current paper as the results are not yet published (at December 1998). However, the effects appear to be real and promise to deliver further strategies for instructional design.

Suggested readings

If you wish to pursue further readings in cognitive load theory then try:
1. Sweller (1991), a very short, non technical description of how educational practice is often based upon myths rather than empirical research, and then
2. Sweller (1994) which presents a more detailed (though still non-technical) review of cognitive load theory and the effects generated.
3. Sweller, Van Merrienboer & Paas (1998) which presents a detailed (and partially technical) summary of human cognitive architecture and the implications for instructional design.

After that you may wish to go to the original journal papers. *Happy readings.*

References

Ayres, P. (1993).Why goal free problems can facilitate learning. *Contemporary Educational Psychology, 18* , 376-381.

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

Cerpa, N., Chandler, P., & Sweller, J. (1996). Some conditions under which integrated computer-based training software can facilitate learning. *Journal of Educational Computing Research, 15,* 345-367.

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 Education Psychology, 62,* 233-246.

Chandler, P., & Sweller, J. (1996). Cognitive load while learning to use a computer program. *Applied Cognitive Psychology, 10,* 151-170.

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

Jeung, H. J., Chandler, P., & Sweller, J. (1997). The role of visual indicators in dual sensory mode instruction. *Educational Psychology, 17,* 329-343.

Larkin, H., McDermott, J.,Simon, D., & Simon, H. (1980). Models of competence in solving physics problems. *Cognitive Science, 11 ,*65-99.

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.

Mousavi, S., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology, 87,* 319-334.

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., & Van Merrienboer, J. (1994). Variability of worked examples and transfer of geometric problem-solving skills: A cognitive load approach. *Journal of Educational Psychology, 86,* 122-133.

Pavio, A. (1990). *Mental representations : A dual coding approach.* New York : Oxford University Press.

Simon, D. P., & Simon, H. A. (1978). Individual differences in solving physics problems. In R.S. Seigler (Ed.), *Children's thinking : What develops ?* Hillsdale, NJ : Lawerence Erlbaum Associates.

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

Sweller, J. (1991). Some modern myths of cognition and instruction. In J. B. Biggs (Ed.), *Teaching for Learning: The view from cognitive psychology* : ACER, Radford House, Vic., Australia.

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

Sweller, J., Chandler, P., Tierner,P., & Cooper, M. (1990). Cognitive load in the structuring of technical material. *Journal of Experimental Psychology; General, 119,* 176-192.

Sweller, J., & Levine, M. (1982). Effects of sub-goal density on means-ends analysis and learning. *Journal of Experimental Psychology : Human Learning and Memory, 8,* 463-474.

Sweller, J., Van Merrienboer, J., & Paas, F. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review, 10 (3),* 251-296.

Tindall-Ford, S., Chandler, P., & Sweller, J. (1997). When two sensory modes are better than one. *Journal of Experimental Psychology: Applied, 3,* 257-287.

Van Merrienboer, J., & Krammer, H. (1987). Instructional strategies and tactics for the design of introductory computer programming courses in high school. *Instructional Science, 16 ,* 251-285.

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

[Education](#) Home Page