Chapter 2
Amassing Information: The Information Store Principle

Natural environments tend to be both highly complex and highly variable. The frequently immeasurable number of variables associated with natural environments tends to be in constant flux. Most animals and plants must survive in complex, variable environments dealing with day and night, summer and winter, drought and flood. Any natural information processing system must find a way to handle this complexity and variability. Nevertheless, despite the complex, variable environment in which a natural information processing system must function, it must be able to treat its environment as familiar and predictable. It must be able to ignore variability that does not matter to its functioning while responding to variability that does matter. In one sense, the manner in which this complexity is handled is straightforward. Immense complexity is handled by immense information stores. Natural information processing systems build sufficiently large information stores to handle most of the vagaries inherent in their environments.

How Natural Information Processing Systems Store Information

Evolutionary Biology

The manner in which the need for a large information store is met by evolutionary biology is well known. All genomes include a huge amount of DNA-based information that determines most biological activity (Portin, 2002; Stotz & Griffiths, 2004). The size of any genome must be large because organisms survive in complex, information-rich environments. A simple, small information store is unlikely to be able to deal with the complexity of any natural environment. A large and sophisticated store is needed to deal with the inevitable environmental variations it will face.

There is no agreed measure of genomic complexity or size. Nevertheless, all genomes consists of, at a minimum, thousands, or in many organisms, billions of base pairs that can be considered units of information. (There is no consensus on

J. Sweller et al., *Cognitive Load Theory*, Explorations in the Learning Sciences, Instructional Systems and Performance Technologies 1,
DOI 10.1007/978-1-4419-8126-4_2,

what should be used as a measure of complexity but all measures yield very large numbers of units of information.) The genetic functioning of organisms and species rely on that large store of information. If all natural information processing systems require a large store of information, it follows that human cognition also must rely on an equivalently large store of information.

Human Cognition: Long-Term Memory

The role of long-term memory in cognition provides an analogical equivalent to a genome in evolutionary biology. Like a genome in biology, long-term memory acts as a very large store of information. The bulk of our normal, everyday activities are familiar. When we say something is 'familiar' what we really mean is that it is based on information in long-term memory. That information permits us to engage in activities from automatically recognising the huge number of objects we see from minute to minute to planning our routine daily activities. All depend on a huge, organised knowledge base held in long-term memory.

Storing biologically primary and biologically secondary information. It can be argued that much of the information stored in long-term memory consists of biologically primary knowledge. We have evolved to acquire enormous amounts of primary knowledge in order to survive and function in our world. For example, when we listen and speak, much of both the physical and social aspects of our activities are based on a massive store of primary knowledge held in long-term memory. Similarly, our ability to effortlessly navigate our physical world provides an indicator of knowledge stored in long-term memory. We see and recognise a large number of objects and faces and can engage in a wide range of physical activities. Many of these activities can be learned without lengthy training and so are indicative of primary knowledge held in long-term memory.

Our primary knowledge base enables us to engage in many of the activities that we frequently consider to be easy and simple. In fact, in information processing terms, many 'simple' activities are anything but simple. They appear simple solely because we have acquired them as biologically primary skills and all biologically primary skills are seen as simple and easy. The enormous amount of stored information required to engage in most biologically primary activities can be seen from the difficulty we have in readily programming computers to mimic primary skills such as recognising voices. That difficulty arises from the large knowledge base required by these skills. We have evolved to rapidly acquire a large knowledge base but because of its size and complexity, an equivalent knowledge base can be very difficult to readily program into a computer.

As an example, consider a simple task such as going outside to pick a flower. Learning to go outside and pick a flower is not a task that requires a long period of training. Despite its simplicity, programming a computer to engage in a similar task would require an enormous expenditure of time and effort and indeed, as far as we are aware, such a task is beyond today's robot-connected computers. In reality, it is an immensely complex task that only appears simple to us because it is based on a huge, primary knowledge base held in long-term memory. We may contrast other tasks that computers are programmed to do. Compared to going outside and accomplishing a simple task such as picking a flower, it is much easier to program a computer to play chess at grandmaster level or to carry out complex mathematical operations. We see these tasks as immensely complex because they are based on secondary knowledge that we have not evolved to acquire.

The extent to which we are able to go outside and pick a flower or carry out mathematical operations are both determined by the amount of knowledge held in long-term memory. Nevertheless, the category of knowledge required by the two tasks is quite different. Picking a flower requires biologically primary knowledge that is quantitatively immense but easily acquired because we have evolved to acquire that knowledge. Complex mathematics requires secondary knowledge that we have not specifically evolved to acquire and so is much more difficult. The knowledge base is probably greater in the case of the flower-picking exercise even though we acquire a flower-picking knowledge base much more readily than mathematical skills. As indicated above, the evidence that going outside to pick a flower may require a larger knowledge base than even complex mathematical operations comes from the differential ease of programming a computer to accomplish both tasks. It is probable that the information associated with biologically primary activities may constitute the bulk of the knowledge we hold in long-term memory.

Evidence for the size and function of stored biologically secondary information. While most knowledge held in long-term memory probably can be categorised as primary, in absolute terms our secondary knowledge base is still immeasurably large. De Groot's (1965) and Chase and Simon's (1973) work on chess can be used to indicate to us the immense amount of secondary information held in the long-term memory store. Furthermore, that seminal work indicated for the first time that many higher-level cognitive activities that were assumed to rely minimally or not at all on long-term memory were largely driven by the contents of the long-term store.

The initial impetus for de Groot's work was the fact that chess grandmasters almost always defeat weekend players. He was concerned with finding the factors that almost invariably result in this outcome. What knowledge does a grandmaster have, in what activities does a grandmaster engage, to enable such dominance? There is a range of plausible hypotheses that, if valid, could provide an answer to this question. Indeed, several possible answers associated with problem-solving skill seem to have been intuitively assumed by most people who either played chess or thought about the factors that lead to chess skill.

One easily accepted plausible hypothesis is that grandmasters engage in a greater level of problem-solving search for suitable moves than weekend players. They may be particularly skilled at using means–ends analysis, for example. When using a means–ends strategy, grandmasters may search to a greater 'depth' than weekend players. That means instead of only considering the consequences of a small number of moves ahead they may consider a much longer series of moves. Considering the consequences of a longer rather than shorter series of moves should result in being able to choose better moves. Alternatively, chess grandmasters may engage in a greater search ‘in breadth’. Whenever they had to make a move, they might consider a large range of alternative moves while a weekend player may only consider a small number of alternative moves. If a larger number of alternatives are considered we might expect a better move to be found than if only a smaller number of alternatives are considered.

In fact, de Groot found no evidence that grandmasters’ superiority derived from either a greater search in depth or a greater search in breadth than weekend players. He found only one difference between different levels of players and that difference seemed to be quite unrelated to problem-solving skill. Rather, it was concerned with memory. De Groot showed chess players board configurations taken from real games for about 5 s before removing the pieces and asking the players to attempt to replace them in the configuration that they had just seen. The results indicated that masters and grandmasters were able to reproduce the board configurations that they were shown with a high degree of accuracy. In contrast, less-able players were far less accurate in replacing the pieces (see also, De Groot & Gobet, 1996).

De Groot obtained this result for chess board configurations taken from real games. He did not attempt to investigate if the same result could be obtained for pieces placed on a board in a random configuration rather than a real game configuration. Instead, Chase and Simon (1973) replicated de Groot’s result but in addition ran exactly the same experiment placing the pieces on the board in a random configuration. The results were much different. For random configurations, there were no differences between more- and less-expert chess players. All players performed equally poorly on random configurations compared to grandmasters reproducing real game configurations. Only expert players faced with real game configurations performed well on this memory test. Less-expert players performed poorly on both real game and random configurations while expert players performed poorly on random configurations only.

In principle, these results are able to provide a full explanation of chess expertise without recourse to any other factors. Skill at chess is not based on an ability to think through a series of unique and ingenious chess moves. Expertise derives from learning to recognise many thousands of the board configurations that are found in chess games as well as learning the moves that are the most likely to be successful for the various configurations. This skill is acquired slowly over many years of consistent, continuous practice. That practice needs to be carried out with the explicit intention of improving performance, called ‘deliberate practice’ (Ericsson, Krampe, & Tesch-Romer, 1993). A chess grandmaster typically requires 10 years of deliberate practice before acquiring a high level of expertise.

Until de Groot’s and Chase and Simon’s work, the cognitive changes that occurred due to practice were essentially unknown. We now know what is learned during practice. According to Simon and Gilmartin (1973), chess grandmasters have stored in long-term memory tens of thousands of board configurations along with the best moves associated with those configurations. The source of chess-playing skill derives from that stored information rather than some mysterious thinking skill. Paradoxically, it is more likely that a less-skilled player must engage in complex thought because a less-skilled player does not have large numbers of board configurations and their associated moves stored in long-term memory. In the absence of stored knowledge, moves must be generated by problem-solving search. With the development of expertise, the need for problem-solving search activities is reduced. Instead, the best move at each choice-point becomes apparent without having to engage in search because that best move can be retrieved from long-term memory. Novices need to use thinking skills. Experts use knowledge.

This account of chess skill can be used to explain the phenomenon of simultaneous chess. In demonstration games, a chess grandmaster can simultaneously play and defeat a dozen weekend players. In the absence of a long-term memory explanation for chess skill, we would need to ask how anyone could possibly simultaneously devise multiple strategies for playing a dozen, complex, different games. The answer, of course, is that it is not possible, but nor is it required. Only the grandmaster's opponents must attempt to devise a strategy for their single game. The grandmaster can arrive at a board and irrespective of the progress of the game, look at the board configuration, immediately recognise it and recall the best move for that configuration. That process then can be repeated for the remaining boards. A novel game strategy does not need to be devised for each board. In contrast, the grandmaster's opponents do need to devise a novel strategy for their single game. In the absence of relevant information stored in long-term memory indicating the best move for each configuration, either a strategy or more probably random moves (see below) will be needed by less-knowledgeable players.

The findings associated with the game of chess are not, of course, unique. The cognitive processes associated with chess skill can be expected to apply to every area requiring biologically secondary knowledge. In particular, topics taught in educational institutions can be expected to have similar cognitive profiles as found in chess. In any biologically secondary area, we can expect the major, possibly sole difference between novices and experts to consist of differential knowledge held in long-term memory. Increased problem-solving skill should be directly caused by increased knowledge of relevant problem states and their associated moves rather than due to the acquisition of unspecified, general problem-solving strategies.

All of the readers of this book have skills similar to those exhibited by chess grandmasters. The only difference is that those skills are, for most people, in fields other than chess. If readers were asked to look at the last sentence for about 5 s and then replicate the very large number of letters that constitute that sentence, most could do so easily and accurately. Similar to the chess results, that skill disappears for randomly ordered letters. Replicating the letters of a sentence is in principle, no different to replicating the pieces from a chess board taken from a real game. The only difference is that educated people spend many years practicing reading while chess grandmasters spend many years practicing and studying chess. The cognitive consequences are identical.

As might be expected, findings similar to those obtained in chess have been obtained in many areas including understanding and remembering text (Chiesi, Spilich, & Voss, 1979), electronic engineering (Egan & Schwartz, 1979), programming (Jeffries, Turner, Polson, & Atwood, 1981) and algebra (Sweller & Cooper, 1985). These findings have important instructional implications that in conjunction with other aspects of human cognitive architecture discussed below have guided the instructional processes discussed in this book.

De Groot's findings not only have implications for instruction, they also provide us with vital information concerning the nature of human cognition and in the process have the potential to change our view of ourselves. De Groot's results provide us with some of the most important findings to be found in the field of human cognition. Humans may have a natural tendency to consider long-term memory as little more than fairly limited sets of isolated facts and incidents. Long-term memory can easily be considered to have a quite peripheral role in really important aspects of human cognition such as problem solving and thinking. De Groot's findings turn this view on its head. The function of long-term memory is vastly more important than simply enabling us to recall events, meaningful or otherwise, from our past. Instead, long-term memory is not only central to human cognition, but central to those aspects of cognition that are seen as representing the apex of the human mind. We require long-term memory for all higher-level cognitive activities such as problem solving and thinking. Expertise in such high-level cognitive processes is entirely dependent on the content of long-term memory. We are competent in an area because of information held in long-term memory. Furthermore, that information held in long-term memory may be the major reason for competence and skill.

Schema theory. Given the importance of information held in long-term memory, it is appropriate to analyse the form in which that information is held. Schema theory provides an answer. The theory became important with the work of Piaget (1928) and Bartlett (1932). Bartlett described an experiment that clearly indicates the nature and function of schemas. He asked one person to read a passage describing events from a foreign culture. That person then wrote as much of the passage as could be remembered. The remembered passage then was presented to a second person with the same instructions; the second person's written passage then was given to a third person etc. This process was repeated until a total of 10 people had read and recorded their memory of the previous person's passage.

Bartlett analysed the changes that occurred from passage to passage. He found two effects: levelling or flattening according to which unusual descriptions of events, such as descriptions of ghosts that appeared foreign to the readers tended to disappear; and sharpening, according to which descriptions of events that were familiar were emphasised. Descriptions of battles that appear commonly in Western literature and culture provide an example of features that were sharpened. What was remembered were not the events depicted in the original passage but rather, schematic representations of those events. Long-term memory holds countless numbers of schemas and those schemas determine how we process incoming information. What we see and hear is not determined solely by the information that impinges on our senses but to a very large extent by the schemas stored in long-term memory.

Schema theory became increasingly important in the 1980s, providing an explanation of aspects of problem-solving performance. A schema can be defined as a cognitive construct that permits us to classify multiple elements of information into a single element according to the manner in which the multiple elements are used (Chi, Glaser, & Rees, 1982). Most people who have completed junior high school algebra, for example, are likely to have a problem-solving schema that indicates that all problems of the form $a/b = c$, solve for a, should be solved in the same manner, by multiplying both sides of the equation by the denominator on the left side (b). Anyone who has acquired this schema will treat this and all similar problems requiring a similar solution as the same entity to be treated in the same way despite any differences between examples. With sufficient levels of expertise, all of the individual elements such as the pro-numerals and the mathematical symbols that constitute this and similar problems are treated as a single element by the relevant schema. As a consequence, any information that appears to correspond to this schema will be treated in an essentially identical manner. We will attempt to solve all problems such as the above algebra problem in a similar manner. In effect, the schema provides a template that permits us to effortlessly solve the problem.

The manner in which problem-solving schemas function provides us with immense benefits. We are able to solve problems that otherwise would be difficult or impossible to solve. Unfortunately, the same processes also have negative consequences that appear unavoidable. Sometimes, a problem will appear to be relevant for a particular schema but, in fact, is not relevant. Schemas held in long-term memory not only can render difficult problems easy to solve but can render simple problems very difficult to solve if the schema is erroneously assumed to provide an appropriate template. When we attempt to solve a problem by using an inappropriate schema because the problem looks as though it belongs to a particular category of problems but does not belong to that category, we have an example of *einstellung* or mental set (Luchins, 1942; Sweller, 1980; Sweller & Gee, 1978). Schemas stored in long-term memory may be essential for us to function but they also can prevent us from seeing what would otherwise be obvious.

Automation. Newly acquired schemas must be processed consciously and sometimes with considerable effort. With increasing practice schemas can be used with less and less conscious processing. Instead, they can be used automatically and largely without effort (Kotovsky, Hayes, & Simon, 1985; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). Our ability to read provides a clear example. Initially, when we are learning to read, we must consciously process each individual letter. With increasing practice, we acquire automated schemas for individual letters but still may need to consciously process the groups of letters that constitute words. With additional practice, word recognition becomes automated and even groups of familiar words can be read without conscious control. With high degrees of automation, conscious effort may only need to be expended on the meaning of text. Thus, competent English readers of this text not only have schemas for individual letters that permit the recognition of an infinite number of shapes, including hand-written letters, a large combinations of letters that form words, phrases and sentences also can be recognised automatically. The lower-level schemas for letters and words become increasingly automated with increasing competence and no longer need to be consciously processed because they have been automated. In contrast, beginning English readers are unlikely to have the same level of automation and will need much more effortful processing to fully understand the text.

Instructional Implications

The role of long-term memory in learning and problem solving provides us with a purpose and function for instruction. The purpose of instruction is to increase the store of knowledge in long-term memory. If nothing has changed in long-term memory, nothing has been learned. Instructional procedures that cannot describe what has changed in long-term memory run the risk that nothing has changed and nothing has been learned.

Not only do we know that an increase in knowledge held in long-term memory is central to learning, we now have a much better idea of what is learned and stored as a consequence of instruction. While we do store very domain-general concepts and procedures in long-term memory, these do not usually provide appropriate subject matter for instruction. General problem-solving strategies such as means–ends analysis must be stored in long-term memory but they cannot easily be the subject of instruction because they are acquired automatically and unconsciously as part of our biologically primary knowledge. Knowledge that can be the subject of instruction tends to be much more narrow. It is domain specific rather than domain general.

De Groot's work on chess indicated the astonishing extent of that specificity. We can learn the rules of chess in about 30 min and using those rules, we can theoretically generate every game that has ever been played and that ever will be played. Learning those rules is essential to chess skill but in another sense, it is trivial. Real chess skill comes from acquiring automated schemas. Good chess players must learn to recognise countless numbers of board configurations and the best moves associated with each configuration. Without that knowledge, knowing the rules of chess is largely useless. Exactly the same principle applies to learning in every curriculum area. For competence, we must acquire domain-specific schemas in curriculum areas that we wish to learn. While we need to learn the well-defined rules of mathematics and science or the more ill-defined rules associated with language-based disciplines such as literature or history, that knowledge will not take us very far. For real competence, we also must learn to recognise large numbers of problem states and situations and what actions we should take when faced with those states and situations.

De Groot's lesson should not be forgotten when designing instruction. In areas where we have not evolved to acquire knowledge, covered by most curriculum areas, knowledge consists of large numbers of domain-specific schemas that must be acquired. That knowledge provides a complete description of expertise. De Groot did not find that chess grandmasters had vastly superior repertoires of general problem-solving strategies, or indeed, cognitive, meta-cognitive, or thinking strategies. Furthermore, there is no body of literature demonstrating enhanced levels of these skills in expert chess players. We argue that such strategies are likely to be biologically primary and so do not need to be taught. Instead of general strategies, chess grandmasters had knowledge of chess board configurations and the best moves associated with those configurations. Differing levels of that knowledge are sufficient to fully explain differing levels of chess expertise. Domain-specific knowledge is also sufficient to fully explain expertise in curriculum areas. We must carefully consider whether many recently popular instructional techniques associated with inquiry, problem-based or constructivist learning procedures that do not emphasise domain-specific knowledge have any base in our cognitive architecture (Kirschner, Sweller, & Clark, 2006). Such techniques appear to proceed without reference to long-term memory or any of the other aspects of human cognition discussed in the next two chapters.

Conclusions

The information held in the long-term memory store is central to all facets of human cognition just as the information held in a genomic store is central to the information processes necessary for evolution by natural selection. The immense size of these stores is necessary to enable them to function in complex natural environments. A natural information store must be sufficiently large to enable it to respond flexibly and appropriately to a very large range of conditions. In the case of human cognition, our long-term memory store is sufficiently large to enable the variety of cognitive activities, both biologically primary and secondary, engaged in by humans. The next issue concerns how natural information stores are acquired. That issue is covered by the *borrowing and reorganising principle* and the *randomness as genesis principle*, covered in Chapter 3.

http://www.springer.com/978-1-4419-8125-7

Cognitive Load Theory
Sweller, J.; Ayres, P.; Kalyuga, S.
2011, XVI, 274 p. 20 illus., Hardcover
ISBN: 978-1-4419-8125-7