Working Memory Capacity, Chunking and Attentional Control

Chunking

Those of us who work with individuals with working memory difficulties need to thoroughly understand the process of chunking, as it involves the inter-relationship between long-term memory and new in-coming information. Cowan (1999; 2005) sees working memory as a set of processes that jointly sustain a limited data set in a highly accessible state in order to carry out specific tasks competently (as most tasks we do involve the integration of new information with knowledge that has already been stored away). The idea that working memory is a long-term memory system surprises many practitioners, as the assumption is that working memory is a type of short-term memory system. But the notion of an overlap between short-term and long-term memory activations is not new (for example, Atkinson & Shiffrin, 1971).

Serial information is dealt with as separate chunks of information because, at least initially, the pieces of information cannot be linked or chunked or dealt with in larger units i.e. they lack semantic ties for the person processing the information. As Cowan puts it, "In situations in which stimuli are familiar and grouping and rehearsal of these stimuli is prevented, each stimulus can be assumed to be represented as a separate, single chunk in working memory" (Cowan, 2005, p. 40). For example, if someone said 'a, b, c, d' to you many of you will be able to remember this information with very little effort. This is because for most of you this will be a familiar sequence – the opening of the English alphabet. If you know the English alphabet, then 'a, b, c, d' would be held in memory as one memory chunk. But how many chunks would it represent for someone else? Of course, we can't speculate or make assumptions for everyone, but a good guess might be four memory chunks. But is this necessarily the case? There do indeed seem to be four chunks of information, but is something hidden? There is something hidden in two respects:

1. As already mentioned, for those of us who are familiar with the English alphabet, these four letters constitute one memory chunk – so there is an association that reduces the apparently four separate chunks to one. (See Figure 1). But this

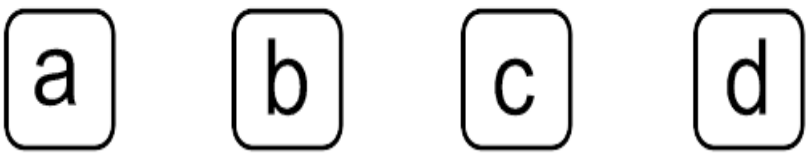

Figure 1

association (or unifying framework) is hidden. Think how this affects what we teach or what we expect people to retain?

2. But, running in the opposite direction, if we do not have a unifying framework there are different and more numerous associations that are hidden, so that rather than four pieces of information, it could be argued that there are actually seven. (See Figure 2). This is because the order in which the letters are placed needs to be stored in addition to the letters themselves. Think how this affects what we teach or what we expect people to retain?

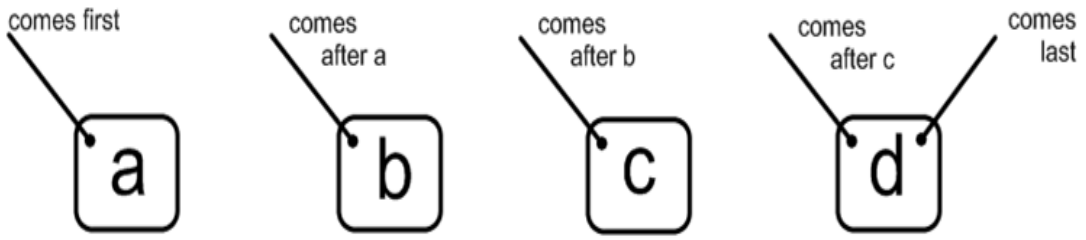

Figure 2

Chunking and Working Memory Capacity

Cowan suggests that the number of chunks that can be held in memory at one time, is somewhere between three and five (Cowan, 2005; see also Broadbent, 1975). He suggests that it is the focus of attention that creates this processing limit. Cowan, Chen & Rouder, 2004 found that this processing limit remained even after training – each chunk just contained a larger number of associations. As Miller says in his 1953 article "We can increase the number of bits of information that [each chunk] contains simply by building larger and larger chunks, each chunk containing more information than before" (p. 93). Miller is describing what we now know as 'superchunks' or 'super-categories' (see Cowan, 2005, p.80 ff.). To go back to our example of the beginning of the English alphabet if we know the sequence of information, this sequence would be held in memory as one chunk. By promoting 'grouping' and 'rehearsal' of items we can promote chunking and enhancing the ability to chunk leads to automatic access to larger and larger pools of information in long-term memory. If facts can be connected and linked to understood information held in long-term memory structures then memory load is decreased (Cowan, 2005; Skemp, 1976; Mellin-Olsen, 1981).

A key aspect of chunking is the presence of meaningful frameworks and how they contribute to understanding and the ability to apply knowledge. The teacher and academic R.R. Skemp

wrote about these matters and it is well worth quoting at length from an article published back in the 1970s as he clearly explains the problems with the word 'understanding' within the context of mathematics (but his ideas have a far wider application to all types of learning). He writes:

> *It was brought to my attention some years ago by Stieg Mellin-Olsen, of Bergen University, that there are in current use two meanings of this word. These he distinguishes by calling them 'relational understanding' and 'instrumental understanding'. By the former is meant what I have always meant by understanding, and probably most readers of this article: knowing both what to do and why. Instrumental understanding, I would until recently not have regarded as understanding at all. It is what I have in the past described as 'rules without reasons', without realising that for many pupils and their teachers the possession of such a rule, and ability to use it, was what they meant by 'understanding'. (1976, p. 21)*

Skemp's point was that if we give those we work with a framework within which to slot information, then we are giving them the opportunity to gain relational understanding. This is best illustrated by an analogy. Imagine you are new to a town. You are driven from the station to your hotel on Saturday evening. On Sunday you want to go to church. You ask and find out that the church is near the station – so being a stranger in the town you walk back to the station and find the church from there… This is not the most direct route and takes far longer – but as you do not know the layout of the town you are dependent on the few bits of knowledge you have. Figure 3 is a resource created to help illustrate this point – more direct ways can be found, if you have 'a map'. This analogy allows us to see the importance of creating meaningful frameworks when learning.

Having more connections in memory triggered by a stimulus, might be thought of as an undesirable state of affairs that would slow thought down or lead to memory being overwhelmed, but as Wolfgang Klimesch puts it "... the speed of search processes in human memory increases as the complexity of interconnected knowledge increases" (p. viii, 1994). It is the 'interconnectedness' that is the key. This also begins to give us an understanding of how expertise is gained (see Ericsson, 2008, for example).

In the case of people with dyslexia, there is often a requirement to understand the information before it can be remembered. Rote learning is therefore very often an unsuccessful approach as the shallow level of learning engagement does not produce enough semantic purchase to promote the consolidation of the information in memory. Consider how important this makes work on categorisation and on different types of association (on synonyms, antonyms, etymology etc.).

Summary: In creating meaningful frameworks we enhance the ability to chunk information; this can circumvent our working memory capacity limit.

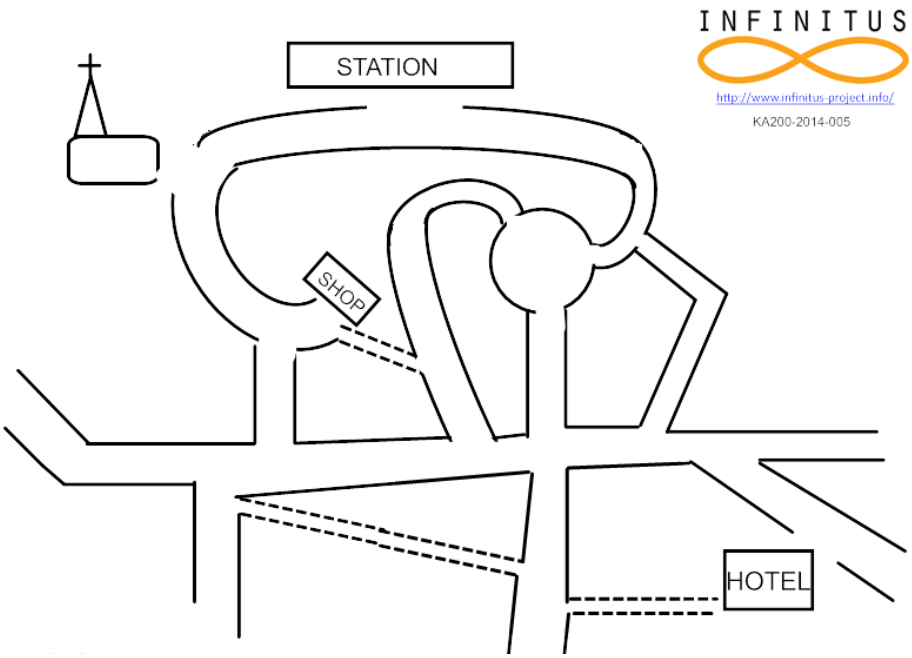

Figure-3

Attentional Control

Research done by Miyake and colleagues (2000) also contributes to a clearer understanding of why certain sorts of tasks are more difficult for learners with dyslexia. This team examined three possible attentional subsystems that could help us to understand the burden of complex cognitive tasks. They found evidence to support these three subsystems:

1. **Mental set shifting** – shifting attention from one task or one set of information to another.
2. **Information updating and monitoring** – maintaining the context of a task whilst processing the information itself.
3. **Inhibition** of what are known as prepotent (that is powerful, automatic) responses – blocking information that is not relevant to the task in hand.

When set out in this way we can see that each of these processes contains ample opportunity for the dropping of information (between shifts in the focus of attention), for the loss of on-task vigilance (maintaining concentration on the process of monitoring whilst properly scrutinising information) and for the inappropriate filtering of incoming information (because appraising for task relevance can be onerous). If we think carefully about a wide range of academic and work-based tasks, we can see that these sorts of activities are

integral to many of them. Consider workers on search and rescue missions, air traffic control officers, shepherds guiding sheep into a pen, chefs cooking a complex dish, integrating ideas to plan an essay etc.

The essence of strategic support is giving those you work with ways of minimising the disruption that these attention shifting activities can cause. This would be done in a range of ways: deliberate practice[1], promoting relational understanding, providing task structure prompts, enhancing metacognitive awareness so that task monitoring becomes explicit and habitual etc.

References

Atkinson, R.C. & Shiffrin, R.M. (1971). The control processes of short-term memory. Retrieved 01/11/18, from: http://suppescorpus.stanford.edu/techreports/IMSSS_173.pdf

Broadbent, D. E. (1975). The magic number seven after 15 years. In Kennedy, A. & Wilkes, A. (Eds.). Long-Term Memory Studies. London: Wiley.

Conway, A.R., Kane, M.J., Bunting, M.F., Hambrick, D.Z., Wilhelm, O. & Engle, R.W. (2005). Working memory span tasks: A methodological review and user's guide. Psychonomic Bulletin & Review, 12(5), 769-786

Cowan, N. (1999). An embedded-processes model of working memory. Models of working memory: Mechanisms of active maintenance and executive control, 20, 506.

Cowan, N. (2005) Working memory capacity. Hove: Psychology Press

Cowan, N., Chen, Z., & Rouder, J.N. (2004). Constant capacity in an immediate serial-recall task: A logical sequel to Miller (1956). Psychological Science, 15, 634-640

Cowan, N., Elliott, E.M., Saults, J.S., Morey, C.C., Mattox, S., Hismjatullina, A. & Conway, A.R.A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. Cognitive Psychology, 51(1), 42-100

Ericsson, (2008). Deliberate Practice and Acquisition of Expert Performance: A General Overview. Academic Emergency Medicine, Special Issue: Proceedings of The 2008 AEM Consensus Conference: The Science of Simulation in Healthcare: Defining and Developing Clinical Expertise, 15(11), 988–994

Klimesch, W. (1994). The structure of long-term memory: a connectivity model of semantic processing. London: Routledge.

[1] 'Based on a review of research on skill acquisition we identified a set of conditions where practice had been uniformly associated with improved performance. Significant improvements in performance were realized when individuals were 1) given a task with a well-defined goal, 2) motivated to improve, 3) provided with feedback, and 4) provided with ample opportunities for repetition and gradual refinements of their performance.' (Ericsson, 2008, p.991).

Mellin-Olson, S. (1981) Instrumentalism as an educational concept. Educational studies in Mathematics, 12, 351 – 67

Miller, G.A. (1956). The Magical Number Seven, Plus or minus two: some Limits on Our capacity from processing Information. The Psychological Review, 63, 81-97.

Miyake, A., Friedman, N.P., Emerson, M.J., Witzki, A.H., Howerter, A. & Wager, T.D. (2000). The unity and diversity of executive functions and their contributions to complex ‘‘frontal lobe’’ tasks: a latent variable analysis. Cognitive Psychology, 41, 49–100

Skemp, R. R. (1989). Mathematics in the primary school. London : Routledge.

Skemp, R.R. (1976). Relational understanding and instrumental understanding, Mathematics Teaching, 77, 20–26. Retrieved 01/11/18 from: https://alearningplace.com.au/wp-content/uploads/2016/01/Skemp-paper1.pdf