Chunking mechanisms and learning

Gobet, F., & Lane, P. C. R. (2012). Chunking mechanisms and learning. In N. M. Seel (Ed.), Encyclopedia of the sciences of learning. New York, NY: Springer.

CHUNKING MECHANISMS AND LEARNING

Fernand Gobet

Department of Psychology, Brunel University

Uxbridge

United Kingdom

fernand.gobet@brunel.ac.uk

Peter C. R. Lane

School of Computer Science, University of Hertfordshire

Hatfield

United Kingdom

p.c.lane@herts.ac.uk

Synonyms

Definition

A *chunk* is meaningful unit of information built from smaller pieces of information, and *chunking* is the process of creating a new chunk. Thus, a chunk can be seen as a collection of elements that have strong associations with one another, but weak associations with elements belonging to other chunks. Chunks, which can be of different sizes, are used by memory systems and more generally by the cognitive system. Within this broad definition, two further meanings can be differentiated. First, chunking can be seen as a deliberate, conscious process. Here, we talk about *goal-oriented chunking*. Second, chunking can be seen as a more automatic and continuous process that occurs during perception. Here, we talk about *perceptual chunking*.

Theoretical Background

Chunking as a mechanism was initially proposed by De Groot (1946/1978) in his study of chess experts' << link to Development of expertise>> perception, memory and problem solving, to explain their ability to recall briefly presented positions with a high level of precision. It was also a central ingredient of Miller's (1956) classical article about the limits on human information-processing capacity. Miller proposed that chunk are the correct measure for the information in the human cognitive system, and that 7 ± 2 chunks can be held in short-term memory. Chase and Simon (1973) proposed a general theory of processes underpinning chunking. It is interesting to note that the approaches of De Groot as well as Chase and Simon emphasize the implicit nature of chunks, that are seen as the product of automatic learning processing sometimes called *perceptual chunking*. Miller's view emphasizes a type of strategic, *goal-oriented chunking*, where chunking is essentially re-coding of the information in a more efficient way. For example, the 9-digit binary number 101000111 can be re-coded as the 3-digit decimal number 327, making it easier to process and memorize for humans. The presence of chunks explains how humans, in spite of strict cognitive limitations in memory capacity, attention, and learning rate, can cope efficiently with the demands of the environment. Chunking has been established as one of the key mechanisms of human cognition, and plays an important role in showing how internal cognitive processes are linked to the external environment.

There is considerable empirical evidence supporting the notion of a chunk, for example in our ability to perceive words, sentences or even paragraphs as single units, bypassing their representation as collections of letters or phonemes; this explains, for example, how skilled readers may be insensitive to word repetition or deletion. Particularly strong evidence is found in those studies which use information about the timing of responses to infer the presence of chunks. The use of response times assumes that the output of elements within a chunk will be faster than the output of elements across different chunks. This is because the elements within a chunk belong to the same structure, as well as sharing a number of relations. There is good empirical evidence confirming that subjects' pauses are shorter within chunks than between chunks. For example, timing information shows that when the alphabet is recited back, letters are grouped in clusters, and clusters grouped in super-clusters. When trained to learn alphabets using scrambled letter orders, subjects also recall letters in a burst of activity followed by a pause, and therefore show evidence for clusters.

The strongest empirical evidence for chunks is based on their inference from several converging methods. For example, studies on chess have shown that chunks identified by latencies in recall or placement of chess pieces correlate highly with chunks identified by the number of relations shared between successively placed pieces. By analyzing the patterns picked out by chess players within a position for various natural relations (including proximity, color, and relations of attack or defense), it is evident that within-chunk relations are much stronger than between-chunk relations. This pattern was found whether the subjects were asked to place pieces on the board from memory (using timings to separate the groups), or to copy a board (using the presence of glances between the two boards to separate the groups). Further empirical evidence for chunking has been uncovered in a number of other areas including artificial-grammar learning, problem solving, and animal research.

The *chunking theory,* developed by Chase and Simon (1973) was an important attempt to formalize the mechanisms linked to chunking. It postulated that attention is serial and short-term memory limited to about seven items (Miller's magical number). When individuals acquire information about a domain with practice and study, they acquire an increasingly larger number of chunks, which themselves tend to become larger, up to a limit of four or five items. While learning is assumed to be slow (10 seconds per chunk), recognition of the information stored in a chunk occurs in a matter of hundreds of milliseconds. Another important assumption is that chunks are linked to possible information. For example, in chess, the domain in which the theory was first applied, a chunk could provide information about potentially useful moves (see Figure 1). Chunks help in a recall task, because groups of pieces rather than individual pieces can be stored in short-term memory. They also help in a problem-solving task, because some of the chunks, being linked to potentially useful information, provide clues about what kind of action should be taken.

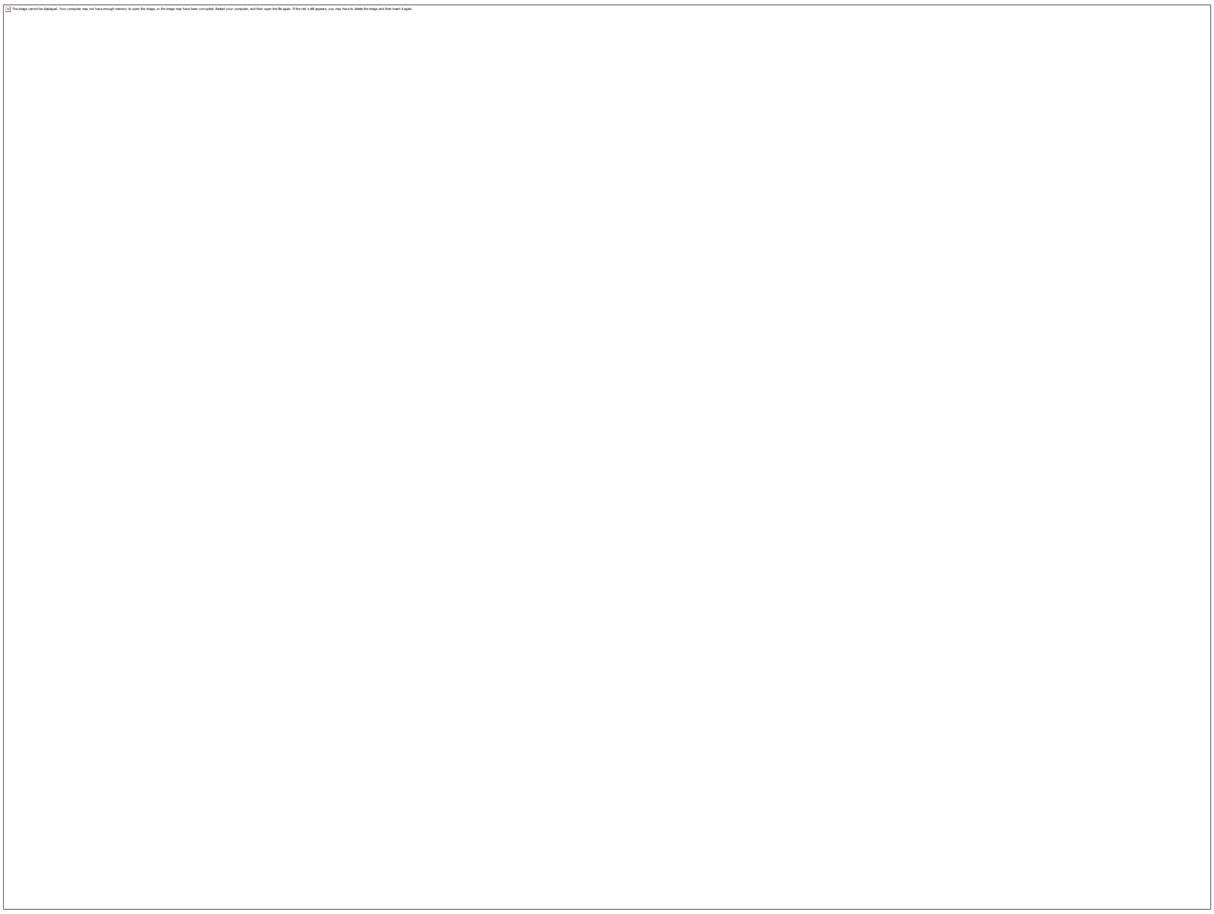

Figure 1. Top panel: examples of chunks in a chess position. Bottom panel: one of the chunks elicits a possible move (retreating the white bishop).

There is also evidence that people, in particular experts in a domain, use higher-level representations than chunks. For example, data from chess research indicate that sometimes the entire position, up to 32 pieces, is handled as a single unit by grandmasters. In addition, evidence from expertise research indicates that information can sometimes be encoded in long-term memory faster than the 10 seconds proposed by chunking theory. Together, these results led to a revision of the chunking theory with the *template theory* (Gobet & Simon, 1996). The template theory proposes that frequently used chunks become "templates", a type of schema. A template consists of a *core*, which contains constant information, and *slots*, where variable information can be stored. The presence of templates considerably expands experts' memory capability.

A methodological difficulty with research on chunking has been to precisely identify the boundaries between chunks. For example, the most direct explanation for observing a set of actions as a chunk is for the actions to be represented internally as a single unit, i.e. a chunk, and so retrieved and output together. However, it is also possible for a subject to plan output actions ahead, and so either break long sequences into sub-parts (e.g., to take a breath when reciting the alphabet) or else compose short sequences into what appear as longer ones (e.g., where a second chunk begins naturally from where the first one finished). Distinguishing between these types is only possible with the aid of a computational model, where the precise items of information known by the subject at a given point in time can be ascertained (Gobet et al., 2001). The advantage of using computer models is discussed in more detail in the entry on Learning in the CHREST cognitive architecture, a model based on the template theory.

Chunk-based theories, such as the chunking and template theories, not only provide a powerful explanation of learning and expert behavior, but also offer useful information as to how learning occurs in the classroom and how it could be improved (Gobet, 2005). We briefly discuss some of the implications for education (further principles are listed in Table 1).

Table 1. Educational principles derived from chunk-based theories (after Gobet, 2005).

- Teach from the simple to the complex
- Teach from the known to the unknown
- The elements to be learned should be clearly identified
- Use an ‘improving spiral,’ where you come back to the same concepts and ideas and add increasingly more complex new information
- Focus on a limited number of types of standard problem situations, and teach the various methods in these situations thoroughly
- Repetition is necessary. Go over the same material several times, using varying points of view and a wide range of examples
- At the beginning, don’t encourage students to carry out their own analysis of well-known problem situations, as they do not possess the key concepts yet
- Encourage students to find a balance between rote learning and understanding

A first implication of chunk-based theories is that acquiring a new chunk has a time cost, and therefore time at the task is essential, be it in mathematics or dancing. As documented by research into deliberate practice <<link to Deliberate practice>>, practice must be tailored to the goal of improving performance. Chunk-based theories give attention a central role – see for example the CHREST model – and such theories are therefore suitable models of deliberate practice. In particular, conceptual knowledge is built on perceptual skills, which in turn must be anchored on concrete examples. Thus, curricula should provide means to acquire perceptual chunks in a given domain.

There are different useful ways to direct attention and to encourage the acquisition of perceptual chunks: to segment the curriculum into natural components, of the right size and difficulty; to present these components with an optimal ordering and suitable feedback; and to highlight the important features of a problem.

If perceptual chunking is an important way of storing knowledge, then a clear consequence is that transfer will be difficult. Unfortunately for learners, this prediction is correct, both for school knowledge and more specific skills such as sports and arts. More than 100 years of research have established that transfer is possible from one domain to another only when the components of the skills required in each domain overlap. Thus, it might be helpful to augment the teaching of specific knowledge with the teaching of metaheuristics – including strategies about how to learn, how to direct one's attention, and how to monitor and regulate one's limited cognitive resources.

As noted above, an important idea in Chase and Simon's (1973) theory is that perceptual chunks can be used as conditions to actions, thus leading to the acquisition of productions. Then, an important aspect of education is to balance the acquisition of the condition and action parts of productions. Another important aspect of education is to favor the acquisition of templates (schemata). Templates are created when the context offers both constant and variable information. As a consequence, and as is well established in the educational literature, it is essential to have variability during learning if templates are to be created.

Finally, chunk-based theories are fairly open to the possibility of large individual differences in people's cognitive abilities. In particular, while they postulate fixed parameters for short-term memory capacity and learning rates, it is plausible that these parameters vary between individuals. In addition, differences in knowledge will lead to individual differences in performance. A clear prediction of chunk-based theories is that individual differences play a large role in the early stages of learning, as is typical of classroom instruction, but tend to be less important after large amounts of knowledge have been acquired through practice and study.

Important Scientific Research and Open Questions

Chunk-based theories have spurred vigorous research is several aspects of learning and expertise. A first aspect is the acquisition of language, where recent research has shown that chunking plays an important role in the development of vocabulary and syntactic structures. A second aspect is related to the neurobiological basis of chunking. Recent results indicate that perceptual chunks are stored in the temporal lobe, and in particular the parahippocampal gyrus and fusiform gyrus.

Other issues being currently researched include the effect of order in learning, and in particular how curricula can be designed so that that they optimize the transmission of knowledge. A possible avenue for future research is the design of computer tutors that use chunking principles for teaching various materials, optimizing instruction for the abilities and level of each student by providing personalized curricula, providing judicious feedback, and teaching strategies.

Cross-References

→ Bounded rationality and learning

→ Decision making and learning

→ Deliberate practice

→ Development of expertise

→ Learning in the CHREST cognitive architecture

→ Schema

References

1. Gobet, F., & Simon, H. A. (1996). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology, 31*, 1-40.

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

3. De Groot, A. D. (1978). *Thought and choice in chess (first Dutch edition in 1946).* The Hague: Mouton Publishers.

4. Gobet, F. (2005). Chunking models of expertise: Implications for education. *Applied Cognitive Psychology, 19*, 183-204.

5. Gobet, F., Lane, P. C. R., Croker, S., Cheng, P. C-H., Jones, G., Oliver, I. & Pine, J. M. (2001). Chunking mechanisms in human learning. *TRENDS in Cognitive Sciences, 5*, 236-243.

6. 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.

Definitions

Chunk: A meaningful unit of information built from smaller pieces of information

Chunking: The creation of new chunks in long-term memory.

Short-term memory span: The largest amount of information that can be held in short-term memory at a given time.

Chunking theory: Theory developed by Chase and Simon in 1973, explaining how experts circumvent the limitations of cognitive processes through the acquisition of domain-specific knowledge, in particular small meaningful units of inter-connected elements (chunks).

Template theory: Theory of expertise, developed in 1996 by Gobet and Simon, building on the chunking theory and proposing that well-elaborated chunks lead to larger meaningful units (templates).