Lane, P. C. R., Cheng, P. C-H., & Gobet, F. (2001). The CHREST model of active perception and its role in problem solving. [Commentary] *Behavioural and Brain Sciences, 24*, 892.

Commentary on: Bernhard Hommel, Jochen Musseler, Gisa Aschersleben, Wolfgang Prinz.

The CHREST Model of Active Perception and its Role in Problem Solving

Peter C. R. Lane {pcl@psychology.nottingham.ac.uk}
http://www.psychology.nottingham.ac.uk/staff/Peter.Lane

Peter C-H. Cheng {peter.cheng@nottingham.ac.uk}
http://www.psychology.nottingham.ac.uk/staff/Peter.Cheng

Fernand Gobet {frg@psychology.nottingham.ac.uk}
http://www.psychology.nottingham.ac.uk/staff/Fernand.Gobet

ESRC Centre for Research in Development, Instruction and Training, School of Psychology, University of Nottingham, University Park, Nottingham NG7 2RD, UK.

Abstract:

We discuss the relation of TEC to a computational model of expert perception, CHREST, based on the chunking theory. TEC's status as a verbal theory leaves several questions unanswerable, such as the precise nature of internal representations used, or the degree of learning required to obtain a particular level of competence: CHREST may help answer such questions.

Main text:

The target article presents a unifying framework for perception and action, assuming their representation within a common medium. We discuss the relation of TEC to a tradition of computational modelling based on the chunking theory, which also addresses many of TEC's theoretical concerns; in particular, the notion of active perception guided by action sequences. The basic principles of TEC include:

1. Common coding of perceptual content and action goals.
2. Feature-based coding of perceived and produced events.
3. Distal coding of event features.

Principles (1) and (2) argue that action goals should be represented as composite action-events, just as perceptual objects are composite perceptual-features, and that integration is required to make perception an active element in action planning. Principle (3) implies that the level of representation is abstracted from that of concrete actions or stimuli, facilitating transfer between perception and action.

The chunking theory (Chase & Simon, 1973; Gobet et al., 2001) includes many of these elements and, further, embodies them within an established chain of computational models. The basic elements of the chunking theory are that collections of primitive features or actions frequently seen or experienced together are learnt as an independently referenced chunk within long-term memory (LTM). Several chunks may be retrieved and referenced within short-term memory (STM), and their co-occurrence used to construct multiple internal references. One such internal reference is the association of perceived chunks with action chunks (Chase & Simon, 1973; Gobet & Jansen, 1994). These ideas have been extended into a computational model of diagrammatic reasoning, CHREST+ (Lane, Cheng & Gobet, 2000, 2001). One consequence of assuming that chunks underpin human memory is that output actions are driven by the perception of chunks in the stimulus; this leads to observable effects in the relative timing of the action events.

The CHREST (Chunk Hierarchy and Retrieval Structures) computational model of expert perception provides a unified architecture for perception, retrieval and action (e.g., see description in De Groot & Gobet, 1996, or Gobet &

Simon, 2000). The model employs a discrimination network LTM, a fixed-capacity STM, and input/output devices, with timing parameters for all cognitive processes. Perception in CHREST is a process involving multiple interactions between the constituent elements of its architecture. The process may be summarised as follows. An initial fixation is made of the target stimulus and the features within the eye's field-of-view retrieved. These features are then sorted through the discrimination network so as to retrieve a *chunk* from LTM. This chunk is then placed within STM. The model next attempts to locate further information relevant to the already identified chunk; it does this by using any expected elements or associated links of the retrieved chunk to guide further eye fixations. Information from subsequent fixations is combined with earlier chunks, within STM, building up an internal image of the external world. If the LTM suggestions for further fixations are not of use, then the eye resorts to bottom-up heuristics, relying on salience, proximal objects, or default movements to seek out useful further features.

How is this process related to problem solving? As with the earlier models of chess playing, chunks within LTM become associated with relevant action sequences. One application of the CHREST model is to problem solving with diagrammatic representations; for example, solving electric-circuit problems with AVOW diagrams (details may be found in Cheng, 1996, 1998); this version of CHREST is known as CHREST+). As CHREST+ acquires information about the two forms of stimuli, it constructs *equivalence links* between related chunks stored in LTM. During perception of a novel problem, CHREST+ identifies known chunks within the problem, and uses these to index chunks in the solution representation. These solution-chunks act as plans, forming the basis from which CHREST+ constructs an action-sequence composed of lines to draw. The level at which the CHREST+ model is operating is very much in concord with that assumed by TEC: the input is composed of features, representing the tail-end of primitive perceptual processing, and the output is a plan, representing the initiation of output processing. Both the input and output chunks are constructed from compositions of features, and both interact with the STM in a common representational format.

Although internal representations are difficult to investigate, a number of testable predictions may be derived from the chunking theory, relating to the relative timing of action sequences (Chase and Simon, 1973; Cowan, 2001; Gobet & Simon, 2000). An example of the process is found when drawing complex shapes, where the latencies between drawing actions are partly governed by planning activities, and these planning activities are mediated by the state of the drawer's memory. A typical series of latencies includes a limited number of isolated local maxima, corresponding to longer periods of reflection; although participants are often not aware of their presence, these maxima are evident when detailed timing information is gathered. Such patterns have been shown to correspond well with a theory that chunks underlie the planning and memory processes (Cheng, McFadzean & Copeland, 2001), and have also been modelled using CHREST+ (Lane, Cheng & Gobet, 2000, 2001).

Without detailed modelling of the lower-level empirical evidence presented in the target article, it is not fair to claim that computational models such as CHREST embody all of TEC's central ideas. However, what is interesting is that the aims of EPAM/CHREST, to capture the relatively high-level processes of expert memory in symbolic domains, have led to a model of active perception with a similar style to that proposed by TEC. CHREST may also be used to formalise some of the assumptions of TEC and turn them into empirical predictions. For instance, some of the harder areas of TEC to formalise are the form of internal representation, or the amount of exposure to a specific domain required to learn a particular association between perceptual and action events. CHREST itself, with detailed timing parameters applicable across many domains, can be used to investigate such questions, using the model to derive strong predictions for the temporal information separating the output actions.

References:

Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In Chase, W. G. (Ed.), Visual Information Processing, Academic Press, New York.

Cheng, P. C-H. (1996). Scientific discovery with law encoding diagrams. Creativity Research Journal, 9, 145-62.

Cheng, P. C-H. (1998). A framework for scientific reasoning with law encoding diagrams: Analysing protocols to assess its utility. In Proceedings of the Twentieth Annual Meeting of the Cognitive Science Society. (pp. 232-235). Mahwah, NJ: Erlbaum.

Cheng, P. C-H., McFadzean, J. & Copeland, L., (2001). Drawing on the temporal signature of induced perceptual chunks. In J. D. Moore and K. Stenning (Eds.), Proceedings of the Twenty-third Annual Conference of the Cognitive Science Society, (pp. 200-205), Lawrence Erlbaum Associates, Mahwah, NJ.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. Bahavioral and Brain Sciences, 24, 87-186.

De Groot, A. D., & Gobet, F. (1996). Perception and memory in chess: Heuristics of the professional eye. Assen: Van Gorcum.

Feigenbaum, E. A., & Simon, H. A. (1984). EPAM-like models of recognition and learning. Cognitive Science, 8, 305-336.

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

Gobet, F. & Jansen, P. (1994). Towards a chess program based on a model of human memory. In H. J. can den Herik, I. S. Herschberg, & J. W. Uiterwijk (Eds.), Advances in Computer Chess 7, Maastricht: University of Linburg Press.

Gobet, F. & Simon, H. A. (2000). Five seconds or sixty? Presentation time in expert memory. Cognitive Science, 24, 651-682.

Lane, P. C. R., Cheng, P. C-H., & Gobet, F. (2000). CHREST+: Investigating how humans learn to solve problems using diagrams. Artificial Intelligence and the Simulation of Behaviour Quarterly, 103, 24-30.

Lane, P. C. R., Cheng, P. C-H., & Gobet F. (2001). Learning perceptual chunks for problem decomposition. In J. D. Moore and K. Stenning (Eds.), Proceedings of the Twenty-third Annual Conference of the Cognitive Science Society, (pp. 528-533), Lawrence Erlbaum Associates, Mahwah, NJ.