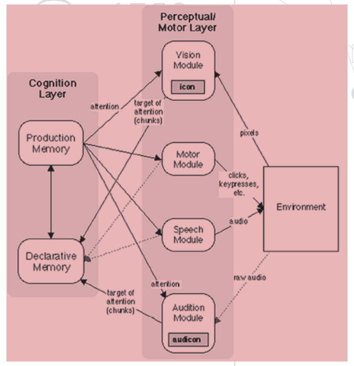
Modelling Eye Movements with the ACT-R Cognitive Architecture

David Peebles
Department of Behavioural Sciences, University of Huddersfield, Queensgate, Huddersfield, HD1 3DH, UK

University of Huddersfield



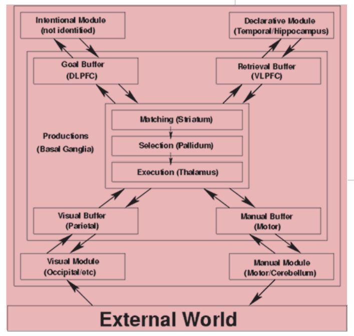
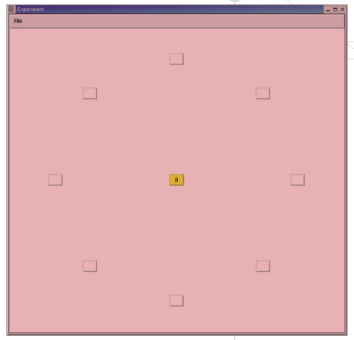
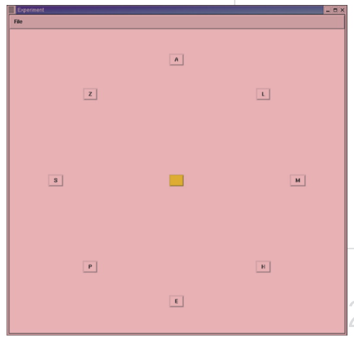


Figure 1. (top) the ACT-R architecture and (bottom) the organisation of information in ACT-R

Figure 2. Screen shots of an experiment trial





The ACT-R cognitive architecture

ACT-R (Anderson & Lebierre, 1997) is a well-established theory of cognition that is implemented as a software system in which computational models of cognitive processes can be developed and tested against empirical data.

ACT-R is one of a new generation of "embodied" cognitive architectures that contain cognitive, perceptual, and motor components, enabling models to interact with the same experimental software as human participants and allowing the modelling of perceptual-motor actions (e.g., shifts in visual attention over a computer's display, keystrokes on a computer keyboard, and mouse movements).

ACT-R is a hybrid architecture containing both symbolic and subsymbolic processes. ACT-R contains a procedural memory consisting of a set of production rules and a declarative memory in the form of a network of chunks. ACT-R also contains five buffers that store information about such things as the current goal, the item of declarative knowledge that is currently available to the system, and the current state of the perceptual and motor modules. In addition, ACT-R contains both symbolic and subsymbolic learning mechanisms that allow it to model various effects of practice on performance.

Modelling visual attention and eye movements in ACT-R

ACT-R's Vision Module determines what ACT-R "sees". The computer's display image is represented (in Common LISP) as a visual icon and production rules can direct visual attention to elements of this icon. Productions can initiate a search for a screen object with specific features (e.g., colour, shape). If the search is successful, a declarative chunk representing the object's location is created in memory and attention is then directed toward the object. When attention is focussed upon the object, a declarative chunk representing the object is created in memory. While they are the current contents of buffers, these chunks can be accessed by other productions for cognitive processing.

ACT-R assumes a default latency of 185 ms for shifts in visual attention. However ACT-R can compute more precise latencies for attention shifts and eye movements using an extension called EMMA (Salvucci, 2001). EMMA makes the time between the shift request and the generation of the chunk representing the visual object dependent on two factors: frequency (how often the object is encountered and encoded); and the eccentricity between the requested location and the current point of gaze. According to this model, encoding time increases with decreases in frequency and increases in eccentricity distance. EMMA divides the eye movement process into two stages: preparation and execution, and assumes latencies for these sub-processes based on Reichle et al.'s (1998) E-Z Reader model.

ACT-R is becoming increasingly successful in providing precise, testable theories of human performance in a number of human-computer interaction and diagrammatic reasoning tasks (e.g., Peebles & Cheng, 2003).

Modelling location learning in a visual search task

This experiment was designed to investigate two aspects of stimulus presentation on the incidental learning of location information in a visual search task: **location frequency** (the number of times a stimulus was presented at a particular location) and **presentation spacing** (the number of trials between presentations of the same stimulus). A second aim of the research was to discover whether the perceptual-motor and learning mechanisms of ACT-R enable it to model location learning and the spacing effect.

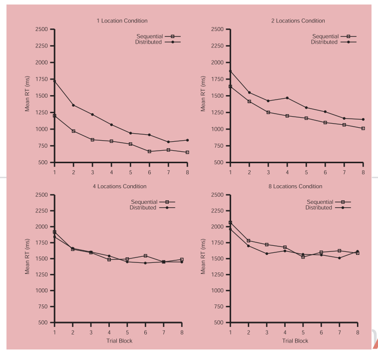
Thirty-six participants had to repeatedly locate each of eight target letters (G, K, O, Y, S, T, V, X) in a circle of distractor letters (A, H, P, E, L, Z, W, M) 64 times (512 trials = 8 blocks of 64 trials).

There were two manipulated variables: **location frequency** (within-subjects: 2 target letters given 1, 2, 4, or all 8 locations); **presentation spacing** (between-subjects: "sequential" = each target letter presented 8 times in a row, "distributed" = each target letter presented once in every 8 trials). In both conditions, each target letter was presented 8 times per block of 64 trials.

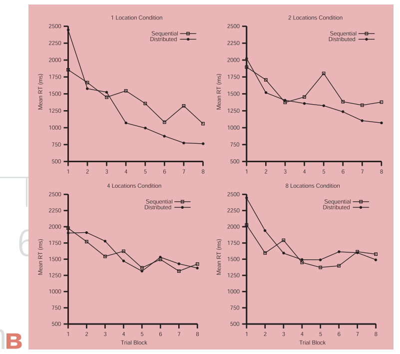
Results

RTs from the four frequency conditions were averaged over each consecutive set of eight blocks (Graph A). An ANOVA revealed a significant effect of presentation spacing, $F(1, 56) = 4.491$, $p < 0.05$, trial block $F(1, 56) = 54.067$, $p < 0.001$, and location frequency $F(1, 56) = 117.074$, $p < 0.001$, together with a significant interaction between presentation spacing and location frequency, $F(1, 56) = 8.086$, $p < 0.01$. However, the mean RT for the first stimulus presentation in a block (Graph B), shows that distributed condition participants are better at remembering the locations for some of the levels of the location frequency condition, indicated by a significant interaction between presentation spacing and location frequency $F(1, 56) = 5.325$, $p < 0.05$. This supports previous results demonstrating the spacing effect.

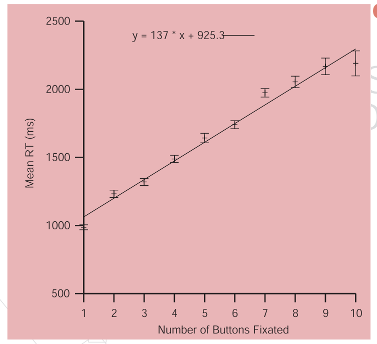
Eye-movements were recorded from 4 participants in each condition and were analysed by computing the frequency and duration of fixations on each of the buttons in the display. Fixations of 100 ms or more were recorded and a scan path consisting of the sequence of fixations for each trial was produced. Graphs C and D show the mean RT as a function of the number of buttons fixated and the reduction in the mean number of fixations during the experiment.



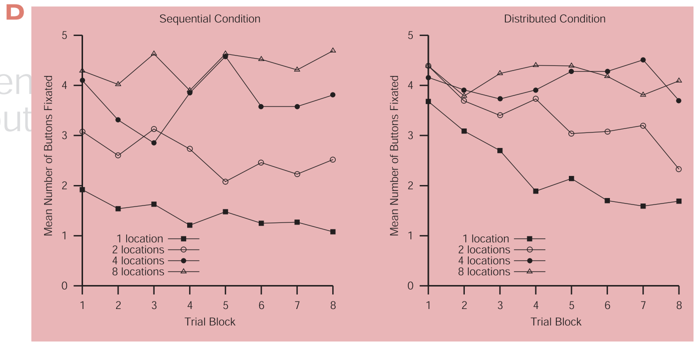
A



B



C

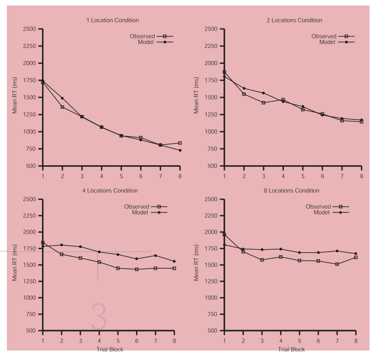


D

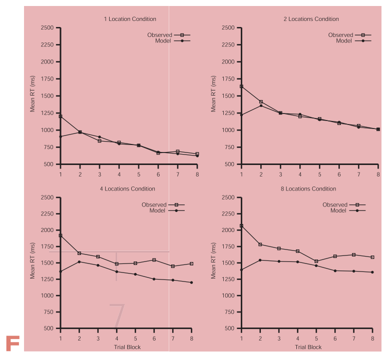
An ACT-R model of the experiment

An ACT-R model of the experiment is able to provide good fits to the mean RT data for the distributed (Graph E; $R^2 = 0.85$) and sequential (Graph F; $R^2 = 0.65$) conditions and also a close match to patterns of button fixations for distributed ($R^2 = 0.73$) and sequential ($R^2 = 0.73$) conditions.

The model (10 production rules) incorporates a simple strategy that initially tries to remember the location of a target letter and selects a location at random if the retrieval fails. However, the current version of ACT-R is unable to account for the spacing effect in the first stimulus presentation data. A modified version of ACT-R is currently being developed to account for this effect.



E



F