

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Modelling Pattern Reproduction in The ACT-R Cognitive Architecture

Permalink

<https://escholarship.org/uc/item/6v5336pc>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

Authors

Liu, Yanze
Cheng, Peter

Publication Date

2023

Peer reviewed

Modelling Pattern Reproduction in the ACT-R Cognitive Architecture

Yanze Liu (yl680@sussex.ac.uk)
Peter C-H. Cheng (p.c.h.cheng@sussex.ac.uk)
Department of Informatics, University of Sussex,
Brighton, BN1 9QJ, UK

Abstract

We present two models of visual location memory developed within the ACT-R cognitive architecture and compare the model's performance to that of human participants in a pattern reproduction task. The *snapshot* model has a fovea-peripheral based activation mechanism, which simulates how more attention and processing resources are given to the centre of the visual field for short stimulus exposure trials (50ms and 200ms). For long exposure trials ($\geq 1s$), a *chunking* model was developed based on the snapshot model by adding chunking processes which can encode geometric patterns. Both models can match the task response accuracy and pause data of human participants. The results of the modelling reveal that for the short stimulus exposure trials the accuracy of recall is affected by the distance between the object location and the fovea vision location. For trials with long stimulus exposure times, participants were likely to use salient geometric patterns to encode the configuration of discs.

Keywords: Chunking, visual location memory, ACT-R

Introduction

The ability to remember the location of objects within our visual environment is an essential cognitive function for daily activities. Visual memory is studied to identify the information processing limitations of visual perception. Some studies have suggested that there is a limit on the total amount of information that can be processed, with the need to remember more objects reducing the amount of detail that can be encoded about each object (e.g., Ma et al., 2014; Nassar et al., 2018). Another theory proposed is the "four slots" limit found in working memory (WM) studies (e.g., Luck & Vogel, 1997; Cowan, 2001). These studies argue that there is a limit on the total number of objects that can be remembered regardless of the amount of information encoded per object. One concept related to this theory is subitizing, which is the ability to count up to 4 visual items quickly and accurately (Pylyshyn, 1989; Trick & Pylyshyn, 1994). Therefore, the number of items that can be processed precisely and quickly are often referred to as the subitizing range, while numbers larger than the subitizing range are typically referred to as the counting range (or estimate range). However, in some situations, the subitizing range can be exceeded in visual location memory experiments. For example, participants might encode configurations of dots in the stimuli as polygons (Yantis, 1992). Even with randomly distributed discs as the stimulus, we can process more than 4 objects at the same time (Haladjian & Pylyshyn, 2011; Haladjian & Mathy, 2015).

An interesting question about visual location memory tasks in which the subitizing range is exceeded is what strategies

participants use to encode the patterns present in the stimulus. One potential strategy is chunking, which is used in complex tasks, such as chess board memorization (Chase & Simon, 1973). It plays a crucial role in overcome limitations in WM.

Chunking, a fundamental concept in cognitive science (Miller, 1994), is a mechanism that groups associated information together, thereby reducing cognitive effort required to hold and recall relatively long streams of information. According to Ullman's visual routine theory (Ullman, 1987), visual perception processing can be divided into two distinct stages: a bottom-up stage of creating base representations, and a top-down stage of applying visual routines to these base representations. In the top-down recognition stage, the shape of several nearby discs can be recognized as a single salient pattern. Studies in cognitive psychology and brain imaging provide support for the chunking hypothesis. For example, research by Amalric et al. (2017) showed that human naturally use Euclidean geometric patterns to encode visuospatial sequences, which can reduce the complexity of stimuli and tasks. These studies provide evidence that perceptual chunks play an important role in the visual enumeration process.

Chunking analyses for a visual pattern reproduction task

The study conducted by Haladjian and Mathy (2015), sought to examine the way in which individuals encode spatial information and whether increasing the duration of stimulus exposure could enhance the accuracy of memory recall. The participants were presented with a stimulus containing a number of randomly placed discs, which they were asked to view for a specified period of time. The discs were presented simultaneously for three different sets of durations: *very short*, 50ms; *short*, 200ms; *long*, $\geq 1s$ (1s per disc). Once the stimulus was removed, participants were instructed to place markers on a blank screen corresponding to the location of each disc they had seen, using a mouse. Haladjian and Mathy hypothesized that if participants encoded individual items or groups of items separately, then an increase in stimulus exposure time would allow for improved rehearsal and thus enhance the accuracy of spatial information recall. The accuracy of participants' responses was evaluated using measures such as the Euclidean distance between the coordinates of each stimulus-response pair and the spatial compression distortion (as previously defined by Sheth & Shimojo (2001) and Haladjian et al. (2010)). The results of the study revealed that the participants' localization accuracy did not improve significantly with longer viewing durations,

but a compression effect was observed across all types of stimuli. Based on these findings, the authors suggest that spatial information for a set of objects is extracted globally and quickly in a snapshot-like manner, without the involvement of any grouping effect.

However, chunking is such a common phenomenon that occurs in diverse tasks. We expected even in this simple visual pattern reproduction task chunk composition could exist. Given sufficient time it seems likely that participants could study the stimulus and encode the discrete patterns with salient features, such as squares or triangles.

We re-analysed the data of Haladjian and Mathy, focusing on both temporal (pauses between two successive responses) and spatial (geometric feature and spatial relation) measures to examine the disc production procedure of participants, to look for the presence of chunking. Temporal signals, the pauses, can serve as indicators of boundaries between different chunks in memory (Chase & Simon, 1973; Cheng & van Genuchten, 2018). This is because transitioning between chunks requires extra planning and retrieval processes, which are not present in within chunk actions. Longer pauses, therefore, are indicative of a switch to a new chunk. One interesting feature of the human response accuracy data in Figure 1 is the decrease of accuracy from 7 to 9 discs for the short and very short trials, with only a small drop for long exposure trials. We claim that the better performance on the long exposure trials for ≥ 7 discs suggests chunking. Our analyses indicate that in very short and short exposure trials, participants are likely to retain the stimulus as a mental image, as suggested by Haladjian and Mathy (2015). However, in long exposure trials, in addition to holding a mental image, participants were also likely to build chunks based on salient geometric patterns which contain a larger number of discs. Details of our re-analysis will be published elsewhere.

Here, we aim to better understand the factors that influence visual location memory by building models in the ACT-R (Adaptive Control of Thought–Rational) cognitive architecture (Anderson et al., 2004). The aim of this study is to develop a model for short (≤ 200 ms) exposure trials with no chunking mechanism and another model for long (≥ 1 s) exposure trials with a chunking mechanism. If the performance of the ACT-R models we built based on the conclusions of our analyses is comparable to the human data obtained from the experiments conducted by Haladjian and Mathy, it would serve as some evidence for our claim that chunking does occur in the pattern reproduction task. Thus, in this study we attempt to answer the following questions:

- 1) Are ACT-R's standard mechanisms and parameters sufficient to model performance in the short and long exposure conditions?
- 2) If not, what are the different representations and mechanisms needed to model the performance?
- 3) Are different mechanisms needed to explain the differences between the two conditions?

¹ Both models are available from GitHub: https://github.com/yanzeliu/ACT-R_pattern_reproduction_model.

ACT-R and visual location memory

ACT-R is a cognitive architecture that aims to explain how humans perform complex cognitive tasks by simulating the functioning of the mind at a high level of abstraction. The ACT-R architecture comprises a collection of modules for the mind's various cognitive functions, including: declarative memory, which is represented by a network of interconnected knowledge chunks; procedural memory, as sets of production rules; vision and motor control. Cognition in ACT-R occurs through a process of pattern matching, in which the system attempts to identify production rules whose conditions align with the current state of the system. Tasks are executed by the sequential activation of production rules. ACT-R has been successfully used to model a substantial range of cognitive phenomena, for example, memory, decision making, motor control, and perceptual processes (e.g., Anderson, 2007).

The capability of ACT-R to attend to and process visual objects displayed on a computer screen is important for the present study. The visual module of ACT-R has two buffers, one for holding chunks representing objects in the visual scene, and the other for holding chunks representing the locations of these objects. When ACT-R perceives an object on the screen, the features of that object ("what"), such as height, colour, and shape are encoded as a chunk in memory. Furthermore, the location ("where") of the object is also encoded as a separate chunk in memory. When chunks are created, they have an initial level of activation that decays over time, determining the probability that the chunks can be subsequently retrieved for future processing.

In ACT-R, memory retrieval happens when a production rule includes a retrieval request to the declarative memory module with one or more cues. For instance, if a production rule includes the features of a visual object, ACT-R will query the declarative memory using these features to retrieve a previously stored chunk that contains those features and that has sufficient activation to be retrieved.

Two models of object location memory

Two versions of the ACT-R model¹ were developed for the pattern location reproduction tasks, a *snapshot* model used to replicate the performance of participants in trials with a short exposure durations, and a *chunking* model used to simulate the responses of participants in trials with a long duration.

Snapshot model

For very short and short trials, our analyses showed that there is no evidence that participants use a chunk composition strategy. To simulate the no chunking response strategy, the snapshot model does not build sub-chunks although it does have an elementary chunk for each disc and an overall snapshot chunk encompassing all discs.

The procedure of the snapshot model can be summarized as follows: (1) The model fixates on the cross mark before

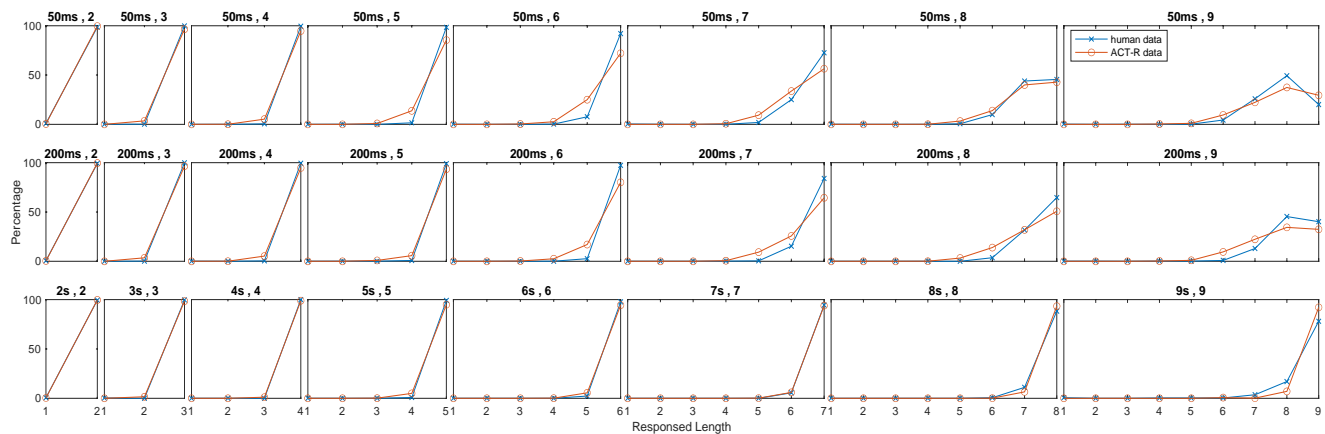


Figure 1: Human and ACT-R performance on the pattern reproduction task. Each row is one exposure time (very short (50ms), short (200ms), and long (≥ 1 s)). Each plot shows the percentage of trials for particular numbers of responses. The blue curves – human data, red – ACT-R models.

the stimulus appears. (2) As soon as the stimulus appears, the model will “take a snapshot” by storing all the disc location information in declarative memory simultaneously. (3) The model shifts its attention to the location of the next nearest disc and stores all the disc location information into declarative memory. This will increase the activation of each elementary disc location chunk, because the declarative memory will merge the chunks with identical contents and give it a higher base level activation. (4) The model checks whether the stimulus is still visible or not; if it is still visible, steps (2) and (3) are repeated until the stimulus disappears. (5) After the mask appears the model begins to rehearse the location chunks by retrieving them from declarative memory and simultaneously checks whether the mask has disappeared. (6) If the mask has disappeared, the model then starts to retrieve disc location chunks beginning with the most active non-retrieved one and making clicks to place the discs.

Representation. The location of each disc is represented as an elementary chunk in the snapshot model. Beside the x and y coordinates information, each disc location chunk also holds a reference to a symbol chunk labelled as “snapshot” for further retrieval processes. This representation is a flat version of list memory representation proposed by Anderson et al (1998). In our representation, the chunk which represent the digits group was removed from the Anderson et al.’s network representation. The elementary disc location chunk is then retrieved during the response stage one by one.

Initial activation of disc location chunk. We tested only changing ACT-R’s decay parameter to match human data. However, this approach did not work; see below. Thus, we implemented a fovea-peripheral based activation mechanism. The activation of each location chunk is allocated based on its distance from the fovea. A higher activation level is assigned to the locations closest to the fovea, thereby simulating peripheral processing, which is the process that

devotes more attention and processing resources to the centre of the visual field. The activation of each disc location chunk is a power function of its distance to the current fovea location in viewing angle. This is based on the function of receptive field diameter and eccentricity proposed by Kandel et al. (2000). Another aspect of activation allocation is that the total amount of activation for all the chunks is fixed. This fixed total activation simulates the limited attention and processing resources of the mind. The total activation parameter was adjusted based on human data. The parameter was obtained by minimizing the squared error between human response data and ACT-R model results.

One modification which violated the constraints of ACT-R in our study was the simultaneous creation of multiple disc location chunks and storing them into declarative memory on the models first exposure to the stimulus. This is different to the typical approach of creating chunks using the imaginal module. The reason for this modification is we tested a model which first created a “snapshot” chunk that contains all disc location information, then used a set of production rules to divide this “snapshot” chunk into individual disc location chunks. However, we found that solely relying on production rules for this process required an unpalatable long time that prevented the snapshot model from having further looks at the stimulus, which led to response accuracy performance in short exposure trials not matching human data.

Response processes parallelism. To let the pause data of our model match with human data, we introduce two levels of parallelism in the response stage for the model. The first level is the parallelism between the motor module and the declarative memory module of ACT-R. After the mouse moving production is fired, the model then checks if the next location is prepared. If not, the production for retrieving the next disc location will be triggered during the mouse moving action. The second level of parallelism is between the mouse moving action and mouse clicking action. Since these two

actions are closely adjacent, participants knew that after a mouse moving was finished, they would click the mouse immediately. Thus, we utilized the preparation and initialization features provided by the motor module of ACT-R to advance the clicking action (Bothell, 2020). Both moving and clicking actions were split into two stages: preparation stage and execution stage. After the move action enters the execution stage, the motor module will then start to prepare for the following click action. The parallelism of the response processes gave plausible pause data using participants' mean coefficient parameters and ACT-R's motor module default parameters.

When adapting the snapshot model's performance to human data, the coefficient parameter in the Fitts's Law equation of ACT-R was set to the mean of the fastest mouse movement of each participant, which is 0.04. The total activation parameter was 23, which was fit only using the 50ms trial data. This parameter also enables the model to match the 200ms trial data. Additionally, we tried to fit the parameter only with the 200ms trial data. It yielded similar result. The total activation parameter is not sensitive to increasing, due to the logarithmic function in the approximation equation used for calculating base-level activation in ACT-R. For the base-level learning and most of the other parameters, we used the default setting of ACT-R.

With the fovea-peripheral based activation allocating mechanism and the parallelism of response processes, the performance of the snapshot model is consistent with both the accuracy and pause time data of human participants, as we will see below.

We also tested the snapshot model under long exposure conditions. The extended exposure time enabled the model to take multiple snapshots, resulting in perfect responses for all the stimuli. This was due to the fact that each disc location chunk was given an overly large activation, with all chunks being activated during the response stage. The direction of this seems to be pulling in two ways: (a) change the decay parameter to address the high activation, considering that this activation is independent of declarative memory decay (i.e., Kanwisher & Wojciulik, 2000; Gold et al., 2005; Bays et al., 2009); (b) having multiple snapshots is unrealistic, so we need chunking.

Chunking model

In contrast to the very short and short exposure trials, our analyses of spatial and temporal signals of participants' data provided evidence that participants used chunking to encode the stimulus in long exposure trials ($\geq 1s$). Thus, a chunking model was developed by extending the overall structure of the snapshot model.

Representation. The representation of the chunking model has the same hierarchy as the list memory representation proposed by Anderson et al (1998). There are two chunk levels: chunking level and disc location level. In the chunking level, there are two types of chunks: pattern chunk, which represents a salient geometric pattern; and group chunk,

which represents a set of co-located discs. The pattern chunk contains the x and y coordinates of vertexes. Meanwhile, the group chunk saves the disc chunks that are part of this group. The structure of the disc location chunk is identical with that of the snapshot model.

Perception processes. During the first 50ms from the stimulus appearance, the chunking model also takes all the discrete location information and stores it into declarative memory in the same way as the snapshot model. The disc chunk creation and initial activation allocation mechanism are identical to the snapshot model. This shared process is intended to simulate the fact that people can get all the location information of the stimulus at first sight. Then, the chunking model begins to search for salient patterns in the stimulus. For example, if four discs form a square pattern, the model will create a chunk in the imaginal buffer with the square pattern tag and coordinates of vertexes, then stores it into declarative memory. The model then continues to search for other salient patterns until time runs out and the stimulus is masked. The procedure of searching patterns is based on the result of Delaunay triangulation on all the discs in the stimulus (Bradski & Kaehler, 2008). After all the patterns have been found, the model will start to group the remaining discs into chunks, depending on the configuration of the stimulus. Each chunk will have 2 to 4 nearby discs. If there is only 1 disc left in the area, the location information of that disc will be stored as a disc location chunk. This disc location chunk will then be merged with the disc location chunk created at the beginning in declarative memory to increase its base level activation, because these two chunks have identical content.

Response processes. The two levels of parallelism in the response stage of the chunking model are identical to those of the snapshot model. The key differences are in the processes of retrieval and finding the next disk location. In the long exposure trials, participants had enough time to construct a complete tree structure of hierarchical representations and plans for the response sequence. Therefore, during the response stage, the model's procedure is as follows: (1) The model retrieves a pattern or group chunk from declarative memory, based on which chunk has the highest activation, and places it in the imaginal buffer. (2) If the retrieved chunk is a pattern chunk, the corresponding series of production rules will be fired to reproduce the pattern on the screen. If the retrieved chunk is a group chunk, the model will use the names of disc chunks to find the corresponding disc location chunks from declarative memory. (3) After finishing all the pattern and group chunks, if there are still missing discs, the model will begin searching for non-responded disc location chunks in declarative memory. The use of pattern chunks and corresponding production rules for reproduction simulates the idea that the knowledge about salient patterns is stored in long-term memory. Each pattern has its own set of rules for reproduction, which are in long-term memory.

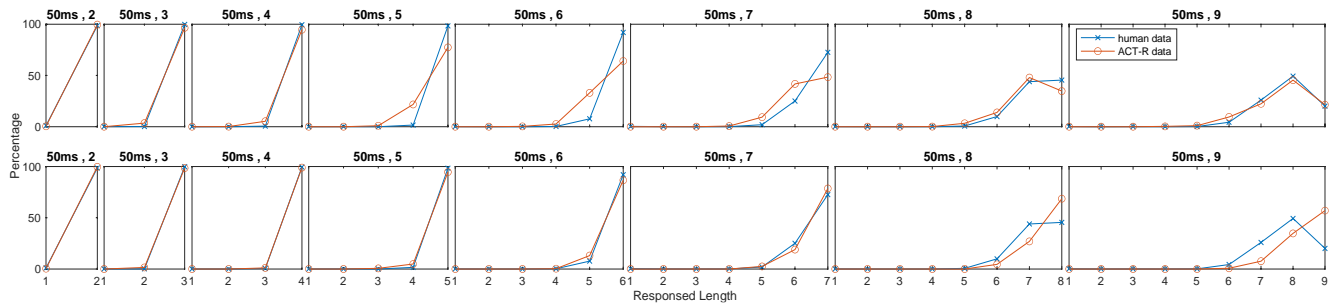


Figure 2: Plots of the percentage of trials for particular numbers of responses. The two rows of data are from the ACT-R model which did not utilize the fovea-peripheral mechanism. The first row presents the data from using a relatively large decay parameter, while the second row shows the data from using a smaller decay parameter. The format of titles is “*exposure time, number of discs in stimuli*”.

Comparing the models with human data

We attempted a modelling approach that only adjusted the decay parameter of ACT-R to match the dramatic decrease in the number of responses of human data for trials with more than 7 discs. However, as Figure 2 shows, this approach did not work. The model either failed to match 5 to 7 discs, while 8 and 9 discs were matched, or vice versa. The decrease in response accuracy for 8 and 9 might be faster than the logarithmic function used in ACT-R's base level calculation, as suggested by Bothell (2020).

In the light of that simply changing the decay parameter did not work, the fovea-peripheral based activation mechanism was introduced, see above. This mechanism works well to reproduce the pattern of responded number of discs from the human data. As seen in Figure 1, the snapshot model not only matches the slight decrease in response accuracy for relatively fewer numbers of discs, but also replicates the tail-off pattern for 8- and 9-disc stimuli as observed in human participants. The number in the title label of each graph in Figure 1 is the number of discs in the stimulus, and the points in the graph are the percentage of responses of each length made by participants responses and ACT-R models. We also conducted a comparison of the number of activated disc location chunks after the mask disappeared in short exposure trials (50ms) with data from 40ms trials from the study of Cheyette and Piantadosi (2020). We found that the number of activated disc chunks in our model and the estimated number of discs in Cheyette and Piantadosi's experiment were closely matched for various stimulus sizes (ranging from 1 to 9 objects). This fact that the default model did not work, but the snapshot model can match human data in multiple dimensions suggests that the allocation of attention and processing resources is crucial in memorizing the locations of multiple objects. The activation allocation mechanism also serves as a backbone to the chunking model.

We then tested the snapshot model under long exposure conditions. The extended exposure time resulted in perfect responses for all the stimuli, which failed to match human data. Thus, we implemented the chunking model by adding chunking mechanism to the snapshot model. The parameters

of the chunking model were the same as the snapshot model. The results of the response accuracy data of the chunking model and human data is shown in the last row of Figure 1.

The two models were both run 500 times for each disc number and exposure time combination. The results of response accuracy and pause data were compared to human data in Figure 1 and Figure 3 respectively. The first two rows of Figure 1 show the response accuracy data of the snapshot model can match the human data in both very short (50ms) and short (200ms) exposure trials well. The model matches the pattern for the smaller stimuli and, more impressively, matches the tail-off for 8 and 9 disc stimuli sizes. The third row of Figure 1 show the response accuracy data of chunking model. The mean squared errors of response accuracy for very short, short, and long between human data (blue curves in Figure 1) and ACT-R data (red curves in Figure 1) are 34.58, 35.72, and 10.12, respectively. These mean squared errors are all lower than those of the default ACT-R model, Figure 2, which are 66.82 and 56.47. It is important to note that the models generate a series of actions, not just one action. All of the models' actions match the performance of human participants.

For the pause data, Figure 3 shows the comparison between the pause data of humans and ACT-R. Figure 3a illustrates the pause data of responses for very short exposure trials for the human and snapshot model. We just show the very short exposure time data as that for the short exposure is similar. A clear pattern in Figure 3a is that the pause of first responses is slightly longer than the others for both human and model data. Figure 3b is the boxplot of pause data of human and model for 9-disc stimulus. In particular, the first pauses are longer, and the increase from the 2nd to the later pauses in the human data is simulated by the chunking model. Both figures demonstrate our models can match the human data. The median values of human and ACT-R data are close.

The human data in Figure 3 only includes trials in which participants reproduced 9 discs without missing any. There are 258 data points for each human data bar for very short exposure trials (Figure 3a), and 351 data points for long exposure trials (Figure 3b).

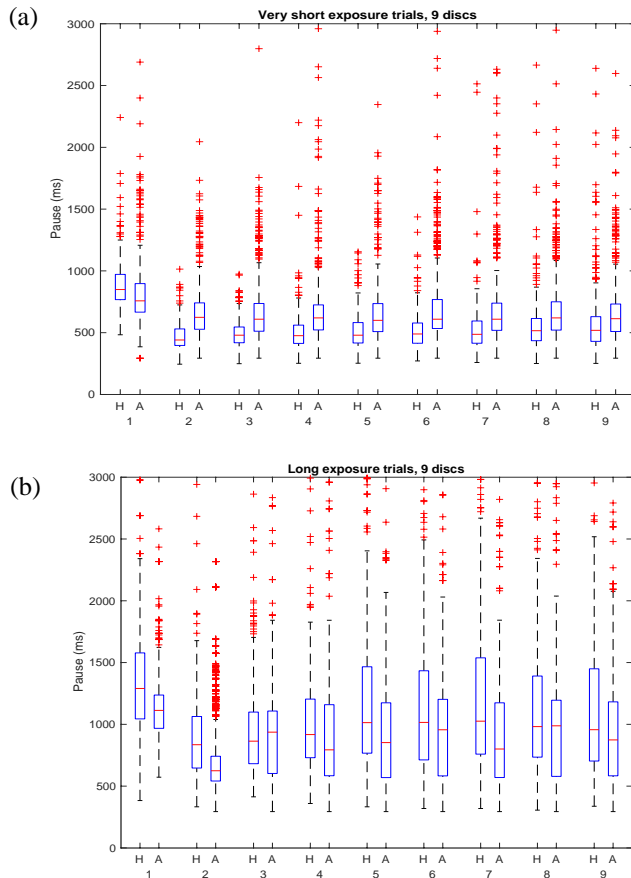


Figure 3: Comparison of pause between human (H) and ACT-R (A) for each response of 9-disc trials.

Discussion

We can answer the three questions posed in the Introduction. The standard mechanisms of ACT-R are not sufficient to model the performance in the short and long exposure conditions in the Haladjian & Mathy (2015), even with variations of the decay parameter. Thus, it necessary to add different representations and mechanisms to explain the differences between the two experimental conditions. With some modifications, ACT-R can model visual pattern reproduction tasks and that matches human data in multiple dimensions closely. The retrieval and decay processes of declarative memory module, the parallelism between modules, and the separate preparation and initiation mechanism of motor module of ACT-R are all crucial, and successfully serve as a common backbone to both the snapshot and the chunking models. The fovea-peripheral based activation mechanism works better to reproduce the pattern of responded number of discs from the human data than simply changing the decay parameter. And by adding the chunking mechanism to the snapshot model it is possible to produce a close match of the models' performance and the human data. These close matching of responded number of discs suggests that chunking will likely occur if there is more time to memorize the stimulus and more discs as stimuli. In

particular, this result differs from Haladjian and Mathy's (2015) claim that there is no grouping in both short and long exposure trials.

For the pause data of the snapshot model, the pause of first response is slightly longer than the others (Figure 3). This is because when the model begins to respond, the disc location chunk has not yet been retrieved from declarative memory. As a result, parallelism between the motor module and the declarative memory module is not applied at this point. Instead, the retrieval and motor action were executed in sequence. For 2 to 9 responses, the medium values of pauses generated by the snapshot model was found to be 200ms longer than that of the human data, as shown in Figure 3(a). This difference may be attributed to the fact that participants may have attempted to reproduce the stimulus as quickly as possible before the "snapshot" decayed, resulting in faster mouse movements than usual. The difference in pause data can potentially be addressed by adjusting the preparation and initialization parameters of the ACT-R motor module and selecting a smaller coefficient parameter of Fitts's Law. However, in the absence of compelling evidence, we decided not to change the default parameters of ACT-R. While Kieras (2009) proposed that modifying the preparation parameter of the motor module is a plausible method for modelling, but we did not find any literature that clearly supports the claim that preparation of motor action is faster in time-starved tasks. Additionally, using smaller coefficient parameter of Fitts's Law may suggest that there were less final iterative targeting movements in this pointing task. And the fewer final targeting movements could lead larger spatial accuracy for very short and short exposure trials. However, this assumption contradicts the conclusion of Haladjian and Mathy's study (2015) that there was no significant difference in spatial accuracy between short and long exposure trials.

The pause data for the chunking model shows a pattern that is similar to that of human data. As shown in Figure 3(b), the 2nd responses have shorter pauses in comparison to the other responses. This is because the chunking model starts its response from pattern or group chunks, which typically contain more than 2 or 3 discs. As a result, the 2nd responses are mostly inter-chunk responses. In contrast, the other responses are a mix of within-chunk and between-chunk responses, resulting a greater variance in the later responses.

One limitation of this study is that our models were unable to reproduce the spatial location errors observed in the experiments conducted by Haladjian and Mathy. While ACT-R's motor module provides a randomized noise for pointing actions, the distribution of this motor action noise differs from the distribution of spatial location errors produced by human participants. Specifically, the spatial location errors observed in human participants appear to be dependent on the distance between the disc location and the centre of the screen (Haladjian & Mathy, 2015; Sheth & Shimojo, 2001). The underlying mechanism of these spatial errors remains an open question. Thus, we were unable to model this aspect of human performance in our models.

References

- Amalric, M., Wang, L., Pica, P., Figueira, S., Sigman, M., & Dehaene, S. (2017). The language of geometry: Fast comprehension of geometrical primitives and rules in human adults and preschoolers. *PLoS computational biology*, 13(1), e1005273.
- Anderson, J. R., Bothell, D., Lebiere, C., & Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language*, 38(4), 341-380.
- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological review*, 111(4), 1036.
- Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* New York, NY: Oxford University Press.
- Bays, P. M., Catalao, R. F., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of vision*, 9(10), 7-7.
- Bothell, D. (2020). ACT-R 7.21+ reference manual. Technical Report. Technical Report, Carnegie Mellon University.
- Bradski, G., & Kaehler, A. (2008). *Learning OpenCV: Computer vision with the OpenCV library.* " O'Reilly Media, Inc."
- Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive psychology*, 4(1), 55-81.
- Cheng, P. C. H., & van Genuchten, E. (2018). Combinations of Simple Mechanisms Explain Diverse Strategies in the Freehand Writing of Memorized Sentences. *Cognitive Science*, 42(4), 1070-1109.
- Cheyette, S. J., & Piantadosi, S. T. (2020). A unified account of numerosity perception. *Nature human behaviour*, 4(12), 1265-1272.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and brain sciences*, 24(1), 87-114.
- Gold, J. M., Murray, R. F., Sekuler, A. B., Bennett, P. J., & Sekuler, R. (2005). Visual memory decay is deterministic. *Psychological Science*, 16(10), 769-774.
- Haladjian, H. H., & Pylyshyn, Z. W. (2011). Enumerating by pointing to locations: A new method for measuring the numerosity of visual object representations. *Attention, Perception, & Psychophysics*, 73(2), 303-308.
- Haladjian, H. H., & Mathy, F. (2015). A snapshot is all it takes to encode object locations into spatial memory. *Vision research*, 107, 133-145.
- Kandel, E. R., Schwartz, J. H., Jessell, T. M., Siegelbaum, S., Hudspeth, A. J., & Mack, S. (Eds.). (2000). *Principles of neural science* (Vol. 4, pp. 1227-1246). New York: McGraw-hill.
- Kanwisher, N., & Wojciulik, E. (2000). Visual attention: insights from brain imaging. *Nature reviews neuroscience*, 1(2), 91-100.
- Kieras, D. E. (2009). Why EPIC was wrong about motor feature programming. In 9th International Conference on Cognitive Modeling—ICCM2009, Manchester, UK.
- Logie, R. H., & Marchetti, C. (1991). Visuo-spatial working memory: Visual, spatial or central executive?. In *Advances in psychology* (Vol. 80, pp. 105-115). North-Holland.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657), 279-281.
- Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature neuroscience*, 17(3), 347-356.
- Miller, G. A. (1994). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 101(2), 343-352.
- Nassar, M. R., Helmers, J. C., & Frank, M. J. (2018). Chunking as a rational strategy for lossy data compression in visual working memory. *Psychological review*, 125(4), 486.
- Pylyshyn, Z. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition*, 32(1), 65-97.
- Sheth, B. R., & Shimojo, S. (2001). Compression of space in visual memory. *Vision research*, 41(3), 329-341.
- Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological review*, 101(1), 80.
- Ullman, S. (1987). Visual routines. In *Readings in computer vision* (pp. 298-328). Morgan Kaufmann.
- Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive psychology*, 24(3), 295-340.