Writing Out a Temporal Signal of Chunks: Patterns of Pauses Reflect the Induced Structure of Written Number Sequences

Peter C-H. Cheng (p.c.h.cheng@sussex.ac.uk)
Hector Rojas-Anaya (h.rojas@sussex.ac.uk)
Representation and Cognition Research Group
Department of Informatics, University of Sussex
Brighton, BN1 9QH, UK

Abstract

Writing may be an effective approach to the study of cognitive phenomena that involve the processing of chunks. This paper provides evidence for the existence of a substantial and robust temporal signal in the process of writing that reveals information about the structure of chunks in working memory. Specifically, it is demonstrated that in the writing of simple number sequences the duration of pauses between written elements (digits) that are within a chunk are shorter than the pauses between elements across the boundary of chunks. This temporal signal is apparent in un-aggregated data for individual participants in single trials.

Keywords: Writing, methodology, chunks, temporal signal, working memory, number sequences.

Introduction

The idea of chunking of information is one of the cornerstones of cognitive science. It underpins accounts of memory, skilled performance, knowledge representation, learning and so forth (e.g., Chi, Glaser, & Farr, 1988; Goodnow & Levine, 1973; Miyake & Shah, 1999; Cowan, 2001). Substantial work has been done to understand the role of chunks in perception and the central processes of the cognitive architecture (e.g., Chase & Simon, 1973; Gobet & Simon, 1998; Egan & Schwartz, 1979; Reitman, 1976; Vincente, 1988, Cheng, McFadzean & Copeland, 2001). Further, there is also evidence that chunks are important in the programs that govern motor behaviour and that such programs appear to have a hierarchical structure (e.g., Rosenbaum, Hindorff & Munro, 1987).

Many approaches have been used to probe the nature of chunks and related cognitive phenomena but writing has not been great among of them. Where ordinary writing by hand has been used for this purpose, it has typically been in the context of simple response latency tasks (e.g., Lochy, Pillon, Zesiger, & Seron, 2002). The relative neglect of writing is surprising given the range and extent of the potential benefits. Writing is an integral part of many tasks, so the problems associated with requiring participants to perform addition activities in order to generate a behavioural data stream is avoided (cf., concurrent verbalizations). Unlike response latency tasks that are rather artificial activities, writing can be more naturalistic, even in an experimental context. The density of data that can be obtained may also be substantial, both in terms of the range of measurable parameters and the number of data points per trial. Computer tools can automatically do much of the initial extraction, analysis and coding of digitally recorded writing actions, without substantial manual effort (although current tools are research prototypes).

To fully exploit the potential of writing some methodological and theoretical advances are required. The aims of this paper are (a) to show the potential of writing by introducing a particular approach to the studying of a chunking phenomenon and (b) to demonstrate the strength and robustness of the data that can be obtained. In particular, the focus is on the extent to which the patterns of pauses between written elements, in this case sets of digits, reveals the structure of chunks in working memory. As will be seen the duration of pauses between written elements strongly and clearly reflects the structure of the chunks and so provides a distinct *temporal signal of chunks*.

Given the fundamental role of chunks in the cognitive architecture, it would be a revelation if the process of writing did not depend on the chunks in memory and if behavioural measures of writing did not reflect the structure of the chunks. Nevertheless, it is imaginable that process of writing could be dissociated from the structure of chunks in working memory in substantive ways. Writing is a complex skill that takes years to fully master and could involve the development of complex processes specific to this activity. Such a process might, for instance, specially recode chunks into a fixed uniform size in order to optimize the planning and execution of actual motor behaviours, which would mask the structure of the original chunks. Existing evidence does suggest that this is not the case. One aim of the experiment reported here is to provide more direct evidence of the existence of a temporal signal of chunks.

To concentrate on the effects of chunks on the writing process it is necessary to attempt to isolate the basic process from others that reflect chunk structure in behaviour measures, in particular, the effect of the recall of chunks from long term memory. Thus, simple sequences of numbers with regular structure were designed that could be quickly learned and that would not require deep coding by participants before they could begin writing. Further, the sequences were designed so that alternative patterns could be imposed on them, so that participants would process alternative chunk structures although the underlying set of numbers was the same. The three groups of three sequences used in the experiment are shown in Fig. 1. The alternative chunk patterns were induced in working memory by telling the participants the nature of the target pattern and having them read the target pattern written in words. For example, the interpretation and number sequence 1A is '500 repeated and separated by numbers counting from 1' and the wording was 'Hash, five hundred, one, five hundred, two, five hundred, three, etc'. For 1B they were, respectively, 'Counting up in ones from five thousand and one' and 'Hash, five thousand and one, five thousand and two etc'. The participants recited the target wording until they were sufficiently familiar with it that they could write it in a continuous unhesitating manner. The hash (#) at the beginning of each sequence is initially written so that the writing process is well underway before the first digit is generated.

To determine whether the structure of chunks has a role in the processes of writing the duration of pauses between drawn elements was recorded in the experiment. Using a graphics tablet different elements could be identified by whether the pen was in contact with the tablet. The pause before a particular element is operationally defined as the difference in the time between (a) when the pen was last lifted from the tablet on completion of the preceding element and (b) when it again touched the tablet at the beginning of the element under consideration. The participants wrote the digits within a horizontal row of equally space squares separate by small gaps, as shown in Fig. 2. This allowed the transition between successive digits to be automatically distinguished. The pauses can be coded at three levels: L0 – within a digit; L1 – a digit within a chunk; L2 – a digit beginning a new chunk (as defined in Fig. 1). The L2 coding depends on the given chunk pattern. L0 pauses were relatively rare and so are not considered in this paper.

Experiment: Writing numbers sequences

There were 10 participants, who were postgraduate students and members of research staff at the University of Sussex.

After familiarization with writing on the tablet and some training on a set of dummy sequences the participant wrote each of the nine sequences. One sequence from each set of three, in Fig. 1, was done in turn before returning to another sequence from the same set, otherwise the order of stimuli presentation was random.

Each number sequence was written on a card with the description of the pattern and the precise wording to be recited, shown most prominently. After familiarization with the number sequence, the experimenter checked the accuracy of verbalization of the participants, and they then wrote the number sequence in the row of squares provided whilst simultaneous reciting the sequence again. This was to ensure there was no recoding of the sequence by the participants.

1A. # 500 1 500 2 500 3 500 4 500 5 …
1B. # 5001 5002 5003 5004 5005 …
1C. # 5 001 5 002 5 003 5 004 5 005 …

2A. # 712 713 714 715 …
2B. # 71 2 71 3 71 4 71 5 …
2C. # 7 12 7 13 7 14 7 15 …

3A. # 303 5 404 5 505 5 606 5 …
3B. # 30 35 40 45 50 55 60 65 …
3C. # 3035 4045 5055 6065 …

Fig. 1. Experiment number sequence chunking patterns

A standard graphics tablet was used (Wacon, Intuos2®) connected to a personal computer. Points were sampled at a rate of more than 36 Hz and at an accuracy of better than an order of magnitude smaller than the shortest pauses. A specially written program, TRACE, was used to record the writing actions and to extract the pen positions, times of points and pauses (Cheng & Rojas-Anaya, 2004).

The data for each of the written sequences for each participant was initially treated individually.

Results

The presentation of the results will start with data for individuals doing a particular number sequence and proceed towards a global overview. Different level of aggregation of data will thus be considered and so it is worth introducing a little terminology to differentiate them. Data for an *individual-sequence* covers a trial of one written number sequence by one participant. Each participant produced nine such individual-sequences and there were a total of 90 in the experiment. A *participant* set of data covers all of the number sequences written by one participant. There were ten in the experiment. A *sequence* set of data covers all 10 participants writing the same number sequence and there are nine such in the experiment.

The pause durations for all written elements was calculated. Each of the marks made were coded as being an element within a digit (L0), as a digit within a chunk (L1), or as the first digit of a chunk (L2). As is typical with pause data for chunking production behaviours, the magnitudes of between chunk pauses is skewed, so medians and non-parametric statistics will typically be reported here.

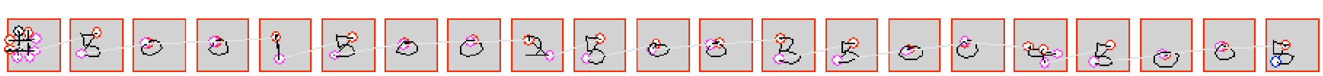

Figure 2: Number sequence 1A written by a participant (DR) with extracted elements and transitions between digits shown.

Patterns in basic un-aggregated data

Fig. 2 shows a screen snap shot from the TRACE graphical recording and analysis program for the writing of sequence 1A by one participant. The small circles superimposed on each written digit indicate the beginning and end of the production of those digits with the pen touching and leaving the paper. The lines between digits indicate transitions between squares where the pen is off the paper. Note the two pairs of dots on the '4' digit, which indicates that it was written in two parts. TRACE calculates the pauses between all of the elements.

Fig. 3 shows two graphs of the sequence of pause durations for the same participant writing sequence 1A and 1B. The number sequence is shown below the graph with each chunk aligned to their respective data points. The solid line gives the pause durations. The dashed line (arbitrary units) indicates expected chunk level, whether it is an L2 data point (100 units), an L1 point (50 units), or an L0 point (zero units). An indication of the match between the expected chunk structure and the durations of the pauses can be judged by comparing the shape of the solid and dashed lines (not their absolute magnitudes). The duration of the pauses reflects these chunk related levels reasonable well, but not perfectly. The pauses with the greatest magnitudes typically occur at the beginning of a new chunk.

The distribution of the longer pauses is different in the two graphs despite the same underlying set of numbers being written. The pattern of pauses clearly reflects the chunk structure imposed by the specific stimuli interpretation given with each number sequence.

The graphs in Fig. 3 are quite typical for this participant across the sequences and also quite representative of the other participants. It is noteworthy that the data correspond to a single individual on a single trial (i.e., not aggregated data). Inspection of all the graphs for all the participants' gives a distinct impression there is a temporal signal that reflects the individual chunk structure of each sequence.

Within and between chunk pause durations

For each individual-sequence the L1 (within chunk) and L2 (between chunk) pauses were considered. All 90 individual-sequence medians were computed and Mann-Whitney U computed to test the whether the difference between L1 and L2 pauses was likely to have been due to chance variations. Table 1 presents the participant medians for the within and between chunk pauses, and the difference between them (L2-L1). The outcomes of the Mann-Whitney test are summarized in terms of the number of sequences in which the difference between the L1 and L2 pauses was significant.

For every number sequence across all the participants the median of the within chunk pause duration was less than that of the between chunk pause duration. Of the 90

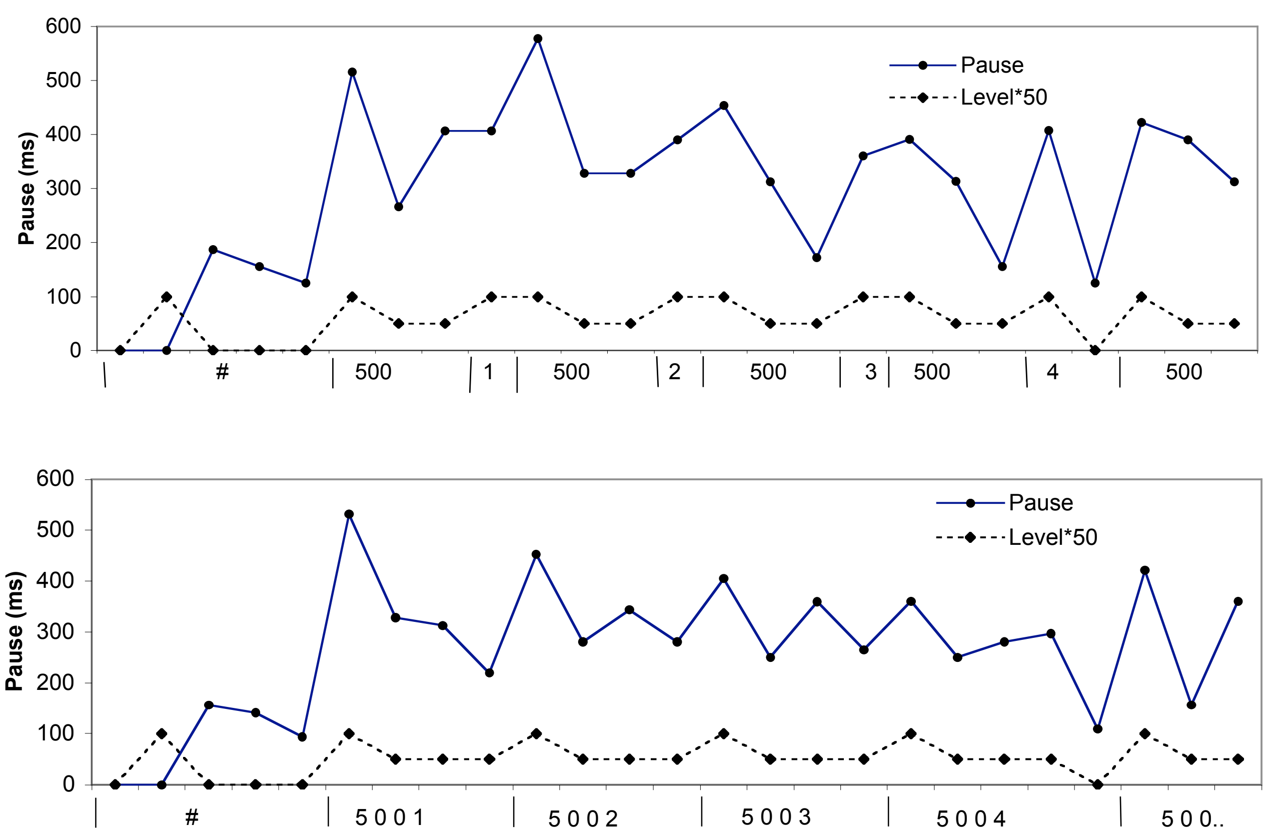

Figure 3: Graphs of successive pauses for one participant (DR) writing number sequences 1A (top) and 1B (bottom).

Table 1: Pause duration measures for each participant

Participant	BG	DL	DR	MR	MS	MT	VD	YB	RG	EV	Median
L1 participant median	250	297	313	203	242	219	313	375	266	282	276
L2 participant median	532	430	407	282	344	320	469	578	578	493	443
L2-L1 participant medians	305	172	95	86	141	132	148	203	312	219	181
L1—L2 significant difference											Mean
Number of sequences of p<.05	8	9	9	7	8	7	7	8	9	9	8.1

individual-sequences, this difference was significant for 81 of the cases with $p<.05$. This is noteworthy given the data is not aggregated over participants or sequences.

Overall, the magnitude of L2 pauses was 60% greater than the L1 pauses, with a typical difference between the levels of 181 ms.

Relation of pause duration and chunk size

Fig. 4 presents pause data for each number sequence. The means of the individual-sequence medians are plotted, as the data across individuals for each sequence is not skewed (in contrast to the data across sequences for individuals). The variability of L1 pauses is smaller (range = 86 ms) than the variability of the L2 pauses (range = 168 ms). The data points in the graph have been ordered with respect to the magnitude of the means of the difference between L2 and L1 durations. The pattern for each sequence is given, under the name of the sequence. It indicates the length of the chunks and the underlying repetition of the chunk pattern. There is a general trend. For the larger chunks the L2-L1 difference is greater than for the sequence with smaller chunks, with the former being double the latter.

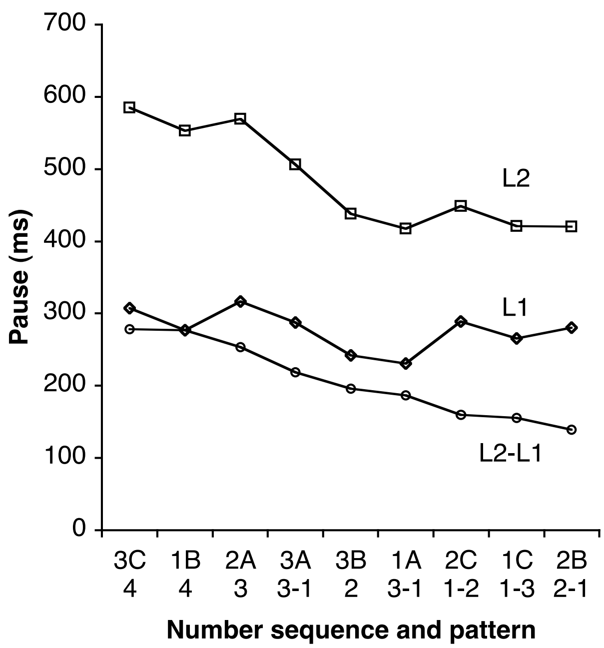

Figure 4: Means of individual-sequence medians

It appears that the particular structure of the chunk pattern underlying each sequence influences the duration of the pauses between chunks, with larger chunks having a greater duration. The effect on within chunk pauses does not show any particular trend with the size of the chunk.

Chunk structure in temporal patterns

To test more rigorously whether the patterns of pauses genuinely is a signal that reflects the imposed chunk pattern a further analysis was carried out. The pause values for each individual-sequence was coded using the other two alternative incorrect target chunk structures from the same set and new median values of the L2 and L1 pauses found. (In Fig. 3, this is equivalent to swapping the dashed lines for expected chunks patterns between the two graphs.)

Table 2 presents the sequence means for each number sequence for codings using all three target chunk patterns, including the correct pattern (in bold): the top number in each cell. The stimuli chunk patterns occupy the main rows and the coding patterns are in the main columns. The underlying chunk pattern for each sequence is provided in the brackets after the label for the sequence. The second number in each cell is the difference between the L2-L1 differences for the alternative and correct coding. Small values or negative values indicate that the alternative coding is as good as, or better than, the coding for the actual stimulus. The third number in the cells gives the number of individual-sequences for which the difference between the L1 and L2 pauses are significant, using a Mann-Whitney test with $p<.05$. Locations of cells in the tables will be referred using pairs of sequences labels, with the first referring to the stimulus pattern (rows) and the second to the coding pattern (columns) (e.g., 1A-1C = top row and right column).

Some interesting regularities are apparent. The magnitude of the L2-L1 difference for the alternative coding is less than the respective value for the correct coding, with two exceptions (1C-1B and 3A-3C). For the alternative codings the numbers of individual-sequences with a significant difference between the L2 and L1 levels is typically less than the correct coding and only in one case equal to it (3A-3C). This suggests that, in general, the match of the stimulus chunk pattern to the recorded structure of the pauses is due to the actual chunk pattern.

Other interesting regularities can be seen in Table 2. *Spurious positive* codings occur when (a) the alternative coding approaches the success of the correct stimuli coding, with the difference between its L2-L1 value approaching (or better) than that of the correct coding (middle number in a cell is less than, say, 70 ms) and (b) the number of individual-sequences that are significant approaches that of the correct coding. These positive codings are highlighted by the underlining of their values in Table 2. The *expected negative* alternative codings are those that do not satisfy the criteria. Now, examining the difference between the two types of coding it can be seen that the sequences that give the expected negative coding are quite distinct from the stimuli patterns, whereas the patterns that give the spurious positive codings can be interpreted as larger chunks that group together the stimuli chunks. For instance the 1A stimulus has repeated sets of 3 and 1 letters and the 1B alternative coding is a simple repeat of sets of 4 letters. Case 3C-3B is an exception, because the stimulus pattern (sets of 4) is not "subsumed" by the alternative coding (a 2-2 pattern), but the simple relation of two equal parts in one is likely to be responsible for the positive match of the 3B pattern.

Hence, there is a plausible explanation for the unexpected successful match of the alternative coding patterns and stimuli: the coding pattern has a substantially similar structure to the stimuli pattern. The positive codings are not actually as spurious as they initially seem, but are manifestation of a pattern shared by the correct and the alternative coding. The existence of these alternative positive codings may be interpreted as lending weight to, rather than detracting from, the claim that patterns of pauses constitutes a temporal signal of the structure of the chunks.

Discussion

The experiment provides evidence for the existence of a strong and robust temporal signal present in the pattern of the durations of pauses between written elements that directly reflects the structure of the chunks in memory. This illustrates the potential of using writing as a means to study cognitive phenomena that are underpinned by the processing of chunks.

The signal is manifest as longer pause durations between written elements that fall at the boundary between chunks and shorter pause durations for elements within a chunk. The typical within chunk pause (L1) was 280 ms and the typical between chunk pause (L2) was 440 ms. The typical difference between pauses levels was 180 ms (and greater than the arithmetic difference because of the positive skew of the data). However, there is considerable variability across individual and between sequences. The L2-L1 differences for individuals, with data aggregated across sequences, ranges from 64 to 254 ms. Aggregating over the participants, the range of L2-L1 differences is from 126 to 304 ms, with a suggestion that the values increase with chunk size.

Table 2: L2-L1 differences for sequences.

Number in each cell are: (a) means of individual-sequence medians (ms); (b) difference of alternative less correct coding (ms); (c) number of individual sequences with p<.05.

Stimulus	Coding		
	1A (3-1)	1B (4)	1C (1-3)
1A (3-1)	**187** **0** **10**	183 4 8	20 167 0
1B (4)	153 124 7	**277** **0** **10**	86 191 2
1C (1-3)	87 69 2	220 -64 6	**156** **0** **7**
	2A (3)	2B (2-1)	2C (1-2)
2A (3)	**253** **0** **8**	47 207 2	128 125 5
2B (2-1)	96 44 4	**140** **0** **9**	-39 179 4
2C (1-2)	154 6 8	-47 207 0	**160** **0** **9**
	3A (3-1)	3B (2)	3C (4)
3A (3-1)	**219** **0** **9**	87 132 2	327 -108 9
3B (2)	41 155 0	**196** **0** **9**	194 2 5
3C (4)	147 131 1	212 67 7	**278** **0** **10**

The existence of the different pause durations for different levels of components within a chunk is consistent with established findings about the nature of chunking and memory recall. However, the results here contrast with previous findings. In particular the pause durations are much shorter than the 2-5 s thresholds that are often used as means to discriminate between elements at the boundary of chunks from those within (Card, et al., 1983). Part of the difference is explained by the minimal need to recall chunks for long-term memory in the present number sequence task (cf. Chase & Simon, 1973; Reitman, 1976; Egan and Schwartz, 1979). Some of the difference may also be due to the to the actual process of writing as opposed to other means of behavioural output, such as physical item placement or drawing (e.g., Cheng, et al., 2001). The relative simplicity of the task may be another contributor. The question of the relative contribution of these factors is a question for further investigation. Short pause duration were found in verbal production of letter sequences (McLean & Gregg, 1956).

The temporal signal appears to be strong and robust. The between chunk pauses are more than 60% greater than the within chunk pauses. In 90% of the individual-sequences significant differences were found; i.e., using un-aggregated data at the level of single trials done by one participant at a time. The recoding of the data for each individual-sequence using the other sequences in the same group provides further evidence of the reality of the temporal signal. When the pattern in the alternative coding was clearly distinct from the target stimulus structure the difference between the chunk levels disappeared. Further, when the pattern happened to be consistent by virtue of being a simple aggregation of pairs of chunks in the stimulus, then the difference in duration between the coded chunk levels remained.

It is an open question whether the strength of the signal will be diminished with more complex stimuli, involving larger chunks, less regular patterns or more hierarchical levels. Such stimuli could, of course, encompass written natural language.

The experiment presented here also stands as a demonstration of one methodology that uses writing in the study of chunking phenomena. Some features of the approach are worth emphasizing. The relative long sequences of items (20 digits here) gives a high density of data per trial and permits relatively complex chunk structures to be used. The regular grid (Fig. 2) facilities the automated coding of drawing strokes associated with the start of particular items and chunks. The "mismatch" analysis using the alternative coding patterns, Table 2, provides a simple means to test the reality of the patterns of pauses and identifies patterns of chunks that are related (i.e., the apparently spurious positive codings). Cheng et al. (2001) used a somewhat similar approach in the context of hierarchical geometrical drawings and also found a strong and robust temporal signal of chunks.

Acknowledgment

This work was supported in part by a grant from the UK Engineering and Physical Sciences Research Council, GR/T10053/01. Our thanks go to the members of the Representation and Cognition Research Group for their assistance with this study. We are also grateful to the three anonymous reviewers for their valuable suggestions.

References

Card, S. K, Moran, T. P., & Newell, A. (1983). The Psychology of Human-Computer Interaction. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cowan, N. (2001). *The Magical Number 4 in Short-term Memory: A Reconsideration of Mental Storage Capacity*. Behavioral and Brain Sciences 24(1): in press

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

Cheng, P. C.-H., McFadzean, J., & Copeland, L. (2001). Drawing out the temporal structure of induced perceptual chunks. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the Twenty Third Annual Conference of the Cognitive Science Society* (pp. 200-205). Mahwah, New Jersey: Lawrence Erbaum.

Cheng, P.C-H., & Rojas-Anaya, H. (2004). TRACE user guide (Unpublished Representational Systems Laboratory report).

Chi, M. T. H., Glaser, R., & Farr, M. J. (Eds.). (1988). *The Nature of Expertise*. Hillsdale, N.J.: Lawrence Erlbaum.

Egan, D. E., and B. J. Schwartz (1979). *Chunking in the recall of symbolic drawings*. Memory and Cognition, 7(2), 149-158.

Gobet, F. (1998). *Expert Chess Memory: Revisiting the Chunking Hypothesis*. Memory, 6(3), 225-255.

Goodnow, J., & Levine, R. (1973). *The Grammar in Action: Sequence and syntax in Children's Copying*. Cognitive Psychology, 4, 82-98.

Lochy, A., Pillon, A., Zesiger, P., & Seron, X. (2002). Verbal structure of numerals and digits handwriting: New evidence from kinematics. *Quarterly Journal of Experimental Psychology*, 55a(1), 263–288.

McLean, R., & Gregg, L. (1967). *Effects of Induced chunking on Temporal Aspects of Serial Recitation*. Journal of Experimental Psychology, 74(4), 455-459.

Miyake, A., & Shah, P. (Eds.). (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. New York: Cambridge University Press.

Reitman, J. (1976). *Skilled perception in Go: Deducing Memory Structures from Inter-Response Times*. Cognitive Psychology, 8, 357-381.

Rosenbaum, D. A., Hindorff, V., & Munro, E. M. (1987). Scheduling and programming of rapid finger sequences: tests and elaborations of the hierarchical editor model. *Journal of Experimental Psychology Human Perception and Performance, 13*(2), 193-203.

Vicente, K. (1988). *Adapting the memory recall paradigm to evaluate interfaces*. Acta Psychologica, 69, 249-278.