Running Head: CHUNKING

Chunking

Curtis M. Kularski

March 15, 2010

Fayetteville State University

ABSTRACT

This paper is an exploration of the cognitive process of chunking. Three research studies on the topic of chunking will be explored to expound on the topic and provide perspective on the implications of chunking on the overall storage and retrieval system of the brain. Each of the studies is from a different area of memory. One is from language, one from number processing and the final study is from the topic of ideomotor principle. Each study provides a unique perspective on how chunking is used to facilitate human learning.

CHUNKING

Chunking is the process of recoding information for storage, as it applies to mental processes, usually referring to the grouping or splitting of information for short term memory. The term itself was defined in George A. Miller's "The Magical Number Seven, Plus or Minus Two : Some Limits on our Capacity for Processing Information." While the paper itself contains no original research, it does formalize some of the concepts that are important to the study of chunking in the modern era of psychology. Chunking is a component of information theory; in fact, it is at the very root of it. Chunking allows complex information to be handled in a simple form. This technique is so essential to information handling that it has inspired some data structures in computer science, including the TCP/IP packet structure and some language processing structures in artificial intelligence (Argamon-Engelson, et. al, 1999).

Chunking is a simple concept, but yet the very reason for why it works is unknown, and research is continuing into how the number of "chunks" can be influenced (Chen, 2005). Various experiments have been conducted involving the use and limits of chunking.

One such experiment was performed by Zhijian Chen and Nelson Cowan and published in their paper "Chunk Limits and Length Limits in Immediate Recall". The purpose of the study was to determine the correlation between word pairing and the number of words that could be remembered. The first experiment utilized 33 undergraduate students who received course credit as compensation for their involvement in the study. There is no information specified for how the participants were selected, however, it is likely that due to the composition and arbitrary number of participants that they volunteered for the study for the credit at their institution and no random

sampling was used. The second experiment utilized 32 undergraduate students. In this experiment all participants were native English speakers with normal vision and no known hearing impairments. None of the participants participated in both experiments. Both experiments utilized three key phases, training, immediate recall and final free recall. Both experiments were administered using words both on a computer screen and as verbal stimuli. Free recall testing was performed orally.

Experiment 1 was a “free recall” in which participants were asked to memorize the list of words and then they were to recall whatever words from the list that they could remember, both immediately and after a time delay (for the final recall). Experiment 2 utilized a somewhat different technique. The training and final recall were carried out the same, but the participants were asked to recall the words in order for the immediate recall. Both experiments utilized lists containing twelve, eight, six and four chunks. The various lists were designated with their number, and then a letter to specify if the list contained single word chunks (s), paired word chunks (p), or words that were not trained (n). The words were selected at random from MRC Psycholinguistic Database. The study found that single word chunks that had been trained in the short list (four chunks) were the most reliably recalled. Non-trained lists were the most reliably recalled of the list types, perhaps due to the novel nature of the list. Outside of the non-trained lists, the lists of four single monosyllable words were the most easily recalled. The difference between the single word lists and the paired word lists was not statistically significant for the final free recall trials, therefore it was confirmed that the pairs function as single chunks.

Another experiment used numbers instead of words to address the subject of chunking and to determine what the capacity for retaining them in memory might be. Using numbers instead of words removes the advantage to memory that is gained by relying on language

processing and the intrinsic weight of importance that some words may be given over others, as most humans treat all numbers equally. David W. Fendrich and Raina Arengo created and executed the experiment utilizing numbers instead of words for chunking. The experiment was published in their paper “The influence of string length and repetition on chunking of digit strings”. The goal of the study was to examine the effects of string length and repetition on performing a data entry task. The study utilized multiple experiments to reach a conclusion. The participants in all three of the experiments were undergraduate students from Widener University. The participants were involved either as a component of an introductory psychology course or for extra credit in a cognition course. The participants had no knowledge of the purpose of the experiments before completing them. For Experiment 1 there were thirty participants, Experiment 2 had forty-eight participants and Experiment 3 had forty. Experiment 1 required each participant to go through 108 trials involving the keying of numbers displayed on a monitor. The numbers were in strings ranging from three to eight digits in length. No digits were repeated in any string and zero was never used. For practice the strings were presented on the monitor while the participant typed the string on the keyboard. After the last digit was typed, the screen was cleared for a period of one second before the next string was displayed. After the practice period the participants were presented with the same blocks of numbers and required to type the strings after they had been presented for one second. The purpose of this portion of the study was to determine how boundaries are placed on numbers and the size of chunks that a participant elects to use. In this experiment the chunk sizes are not determined by the researchers, but instead by the participants themselves, allowing a more natural storage and retrieval of chunks.

Experiment 2 was concerned with determining the impact of repetition in the chunking process. Repetition was carefully avoided during Experiment 1. For this experiment string lengths of between four and six digits were used. The experiment utilized a group of strings with repeated digits as well as a “control” group of strings that contained no repeated digits to compare the results for each participant. As with the first experiment the string was displayed in the center of the monitor, however, there was no training period for this experiment. Once the string was displayed the participant was to memorize the string and then press a key on the keyboard when finished, and then after a one second delay, the participant was to key the string that had previously been displayed. As in the first experiment, the second experiment used 108 trials.

Experiment 3 was designed to extend Experiment 2. Experiment 3 continued the use of repeated digits, but extended it to be multiple pairs of repeated digits as well as pairs of digits that had been split by other numbers. This experiment had the most restrictive strings yet, as they were all six digits in length. Similar procedures were used for Experiment 3 as were used on Experiment 2, with the exception that instead of being a predictable list for each trial, the strings were randomized for this experiment.

The experimenters felt at the conclusion of the study that they had began to form an understanding of the flexibility of chunking. Experiment 1 confirmed the hypothesis that a difference of string length would have an impact on the chunking strategies used. Evidence was also found that if told the length of the string that they will be typing, a participant will plan their keystrokes ahead of time. It was determined that when possible a participant will chunk a string into two or three digit chunks when possible, but will almost never use an inefficient method such as chunking into one digit chunks (even when a odd number of digits are presented). In

Experiment 2 and Experiment 3 it was confirmed that repeated digits improve the ability to recall the entire string, and overall the required study time is reduced. It is the opinion of the researchers that repeated numbers provide a basis for chunking, therefore reducing the overall demand on the short term memory system.

The first two studies addressed cover the usual area of consideration that is given to chunking and how it influences learning. Another study involving chunking is from a quite different area of learning, task sequencing. This study was conducted by Iring Koch, Andrea M. Philipp, and Miriam Gade and the results were published in their paper “Chunking in Task Sequences Modulates Task Inhibition”. The purpose of the study was to examine the impact of chunking on the learning of a sequence of behaviors. The participants in the study were eighty individuals (fifty-seven female, and twenty-three male). The participants were paid for their participation in the study. The stimuli used for the study were the letter ‘A’ and the digit 4. The color, size and orientation of these stimuli were varied to allow for eight possible items of stimuli. Additionally, a variety of cues were used to instruct a participant to report the form, color, or size of the stimulus. Half of the participants were informed ahead of time that there would be a pattern to the tasks. During the first trials the cues were presented in a rational pattern of form, size, form, color, size, color or color, form, color, size, form, size. In later trials there was a pseudorandom pattern used. During the debriefing session the participants were asked to identify any pattern they may have noticed.

The result of the experiment was that there was a learning improvement of 28 milliseconds of response time on each trial for both groups of participants during the predictable sequence phase. During the debriefing interview there was significant difference between the

amounts of sequences remembered between the two groups. The hypothesis that chunking occurs for task performance is confirmed by the experiment.

Chunking is an important part of learning as it is the basis for information processing and a key element in short term memory. It is flexible depending upon the size of the input, and can often be aided by internal relationships of the input itself.

References

Aragmon-Engelson, Shlomo; Dagan, Ido; Krymolowski, Yuval. (1999). A memory-based approach to learning shallow natural language patterns. Journal of Experimental Theories in Artificial Intelligence, (11), (369-390). Retrieved from Academic Search Premier. doi:095281399146463

Chen, Zhijian; Cowan, Nelson.. (2005). Chunk Limits and Length Limits in Immediate Recall. Journal of Experimental Psychology, (31, No. 6), (1235-1249). Retrieved from Academic Search Premier. doi:10.1037/0278-7393.31.6.1235

Fendrich, David; Arengo, Raina. (2004). The Influence of String Length and Repetition on Chunking of Digit Strings. Psychologial Research, (68), (216-223). Retrieved from Academic Search Premier. doi:10.1007/s00426-003-0142-9

Koch, Iring; Philipp, Andrea; Gade, Miriam. (2006). Chunking in Task Sequences Modulates Task Inhibition. Psychological Science, (17(4)), (345-350). Retrieved from Academic Search Premier. doi:10.1111/j.1467-9280.2006.01709.x

Miller, George A. (April 15, 1955). The Magical Number Seven, Plus or Minus Two : Some Limits on our Capacity for Processing Information. Psychological Review, (101), (343-352). Retrieved from http://spider.apa.org.