

Digit memory in Chinese and English: Evidence for a temporally limited store*

JAMES W. STIGLER

University of Chicago

SHIN-YING LEE

HAROLD W. STEVENSON

University of Michigan

Abstract

This paper describes 3 studies comparing short-term memory for digits between native speakers of Chinese and of English. The first study documents, with large samples of kindergarten, first-, and second-grade children, a Chinese advantage in memory for digits. The Chinese subjects, at all grade levels, remembered at least 2 more digits, on average, than did American or Japanese subjects. The second study compared digit memory of 6- and 7-year-old children, Chinese and American, under forward, backward and grouped conditions. The provision of a grouping strategy helped both Chinese and American subjects equally, which fails to support strategy use as the primary explanation of digit memory differences. Further, Chinese children performed more poorly than American children on the backward span. The final study, carried out on Chinese and American university students, investigated differences in pronunciation duration of Chinese and English number words as a possible explanation of span differences. Chinese number words were found to be of significantly shorter pronunciation duration than English number words; and total pronunciation duration for a subject's maximum span did not differ between Chinese and Americans. These findings provide evidence for a temporally limited store.

* We want to express our thanks to those who have collaborated with us in this project: Professors Seiro Kitamura, Susumu Kimura, and Tadahisa Kato of Tohoku Fukushi Daigaku in Sendai, Japan, and Chen-chin Hsu of National Taiwan University, Taipei, Taiwan, who worked with us on Study 1; Chang Hsin-wu of National Taiwan University who worked with us on Study 2; and Janellen Huttenlocher of the University of Chicago who worked with us on Study 3. We also thank Alfred Hu and Alexis Wong for their assistance in conducting Study 3, and Doug Bell for his assistance with Study 2. Study 1 was supported by grant MH 30567 from the National Institute of Mental Health, and Study 3 was supported by a grant from the Spencer Foundation.

Reprint request should be sent to James Stigler, Department of Behavioral Sciences, University of Chicago, 5730 South Woodlawn Avenue, Chicago, IL 60637, U.S.A.

It is common knowledge that Asian children far outperform their American counterparts in tests of mathematics achievement (Husen, 1967; Stigler, Lee, Lucker, & Stevenson, 1982). Numerous factors have been hypothesized to account for this difference, including differences in cultural emphases on the importance of mathematics, the amount of time devoted to mathematics instruction in school, and intelligence. While innatist formulations have been disputed on both theoretical and empirical grounds (e.g., Stevenson, Stigler, Lee, Lucker, Kitamura, & Hsu, 1985), the thought that Asians may be born with extraordinary mathematical abilities fascinates laymen and psychologists alike.

For this reason we were intrigued by a finding from an earlier study of cognitive performance and achievement among Japanese, Chinese, and American elementary school children (Stevenson et al., 1985). Children were administered a standard digit span test as part of a battery of cognitive tests. The digit span, which is a test of short-term memory for numbers, is one of the most widely studied tasks in psychology. Binet and Henri (1895) included digit span in their test of intelligence, and the task has been employed in literally thousands of studies since then. The average first-grader remembers about 5 digits, with a range from 3 to 7. This turned out to be the case in our samples of Japanese and American children. Children in Taiwan, however, performed remarkably better than children in the other two countries; for example, nearly 70% of the first-graders in Taiwan remembered 7 digits correctly, whereas less than 5% of the Japanese and American children did so.

The purpose of this paper is to present results from three studies that describe and clarify the nature of the Chinese digit-memory advantage. The first study presents digit-span performance of kindergarten, first- and fifth-grade children in Taiwan, Japan, and the United States, and compares short-term memory for digits with performance on a test of short-term memory for words. The second study assesses the possible role of information processing strategies in explaining the difference by comparing digit memory of Chinese and American children under different testing conditions. The third study addresses the hypothesis that cross-linguistic differences in the time required to pronounce number words may account for differences in memory for digits.

Study 1: Digit and word span in Chinese, Japanese, and American children

Data for the first- and fifth-graders were collected in 1980 as part of a large study of children's cognitive performance and academic achievement in Tai-

wan, Japan, and the United States. Data for the kindergarten children were collected in 1984, and included kindergarten children in the same three locations. Data on digit span were part of an extensive amount of information collected on large, representative samples of children. Together, these data document the existence of a profound difference in the short-term memory for numbers of Chinese children as compared to Japanese and American children. The design of the first study also permitted us to (a) answer the question of whether there is a general superiority in memory of Chinese children by comparing memory for digits with memory for words, and (b) rule out methodological problems as the cause of the cross-national differences.

Method: Elementary school children

Subjects

The study was conducted in Taipei, Taiwan; Sendai, Japan; and the Minneapolis metropolitan area. The rationale for selecting these cities has been discussed by Stevenson et al. (1985). Because few residents of Minneapolis are from minority groups or from non-English speaking families, our sample is not representative of all American children, but of children from white, English-speaking, native-born families. Taipei, Sendai, and Minneapolis are all economically successful cities with relatively traditional cultures. In all three cities, over 99% of all children of elementary school age attend school.

A stratified random sample of 10 elementary schools in each city was selected for the study. The selection was based on a demographic analysis of the characteristics of both public and private schools. Educational authorities in each locale were consulted to compile a list of schools broken down by region of the city, socioeconomic status of the families, and public vs. private funding. Schools were then selected at random with the restriction that they be representative of the city in terms of the demographic characteristics listed above.

Within each school two first-grade and two fifth-grade classrooms were selected at random, yielding a total sample of 20 classrooms at each grade in each country. All children in each classroom were tested with a reading achievement test approximately 4 months after the beginning of the school year. The sample of children included in the present study was constituted by randomly selecting two boys and two girls from the upper, middle, and lower thirds of each classroom according to their performance on the reading test. The final sample thus consisted of approximately 120 boys and 120 girls at each grade level in each city. (The actual numbers of children in each sample ranged from 237 to 246.)

The children in Taiwan were approximately one month younger on the average than the children in Japan and the United States. The mean ages of the first-graders in Taipei, Sendai, and Minneapolis were, respectively, 6.7, 6.8 and 6.8 years; and the ages of the fifth-graders, 10.8, 10.9, and 10.9 years.

Procedure

All children were tested on two serial short-term memory tasks, one for digits and one for words. The order of presentation of these two tasks was intermixed among 8 other tasks according to four random orders of the 10 tasks.

Digits were selected randomly for the lists with three constraints: no digit was repeated in a list; no successive numbers (e.g., 2-3) followed each other; and sequences of numbers that contained mnemonic cues in any language were eliminated (e.g., 1-7 was not allowed because the words, *ichi-shichi*, rhyme in Japanese). Lists varied in length from 4 to 7 numbers. Two lists of each length were presented. After practice with a list containing 3 numbers, testing continued until the child made mistakes on both lists of a given length. The score was the length of the longest list recalled correctly. Examiners read the lists at a rate of one word per second, and children were told to listen carefully because they would be asked to repeat each list.

The lists of words for the serial memory task varied from 3 to 6 words for first-graders. The length of each of four successive lists of words increased by one word. Words were chosen that were judged to be equally familiar to children in each of the three locations, that would not contain differential mnemonic cues among the three languages, and that could be pronounced relatively easily. All of the words were concrete nouns (e.g., pencil, rabbit, airplane).

The lists of words for fifth-graders varied from 4 to 7 words. They included the 4-, 5-, and 6-word lists used with the first graders and two additional lists of abstract nouns (e.g., experiment, peace, curiosity, satisfaction) that contained either 4 or 5 words. Successive lists were presented until the child was unable to repeat two consecutive lists correctly.

Testing was conducted 6 to 8 months after the beginning of the school year. Each child was tested individually by an examiner who was a resident of the city in which the child lived. Examiners, who had professional experience with children or were preparing for careers in psychology, education, or social work, were carefully trained in the administration of the tasks.

Method: Kindergarten children

Subjects

The vast majority of children in Taipei, Sendai, and Minneapolis attend kindergarten during the year preceding their entrance into first grade. While the ratio of public to private schools differs in the three locations, our strategy was to sample kindergartens that were representative of each city in terms of public vs. private funding, geographic location, and socioeconomic status of the families whose children attend the schools.

In each city, 24 kindergarten classes were selected. Within each class, 6 boys and 6 girls were randomly chosen to participate in the study, yielding a total sample of 288 children in each city. The average age of the children was 6.0 years in Taipei, 6.1 years in Sendai, and 6.1 years in Minneapolis. In all three cities, children were tested approximately 8 months before they were to enter first grade.

Procedure

Children were tested on a digit span test that ranged from 2 to 7 digits in length. The construction of the test and the procedure for administering the test were exactly the same as in the study of first- and fifth-graders.

Results

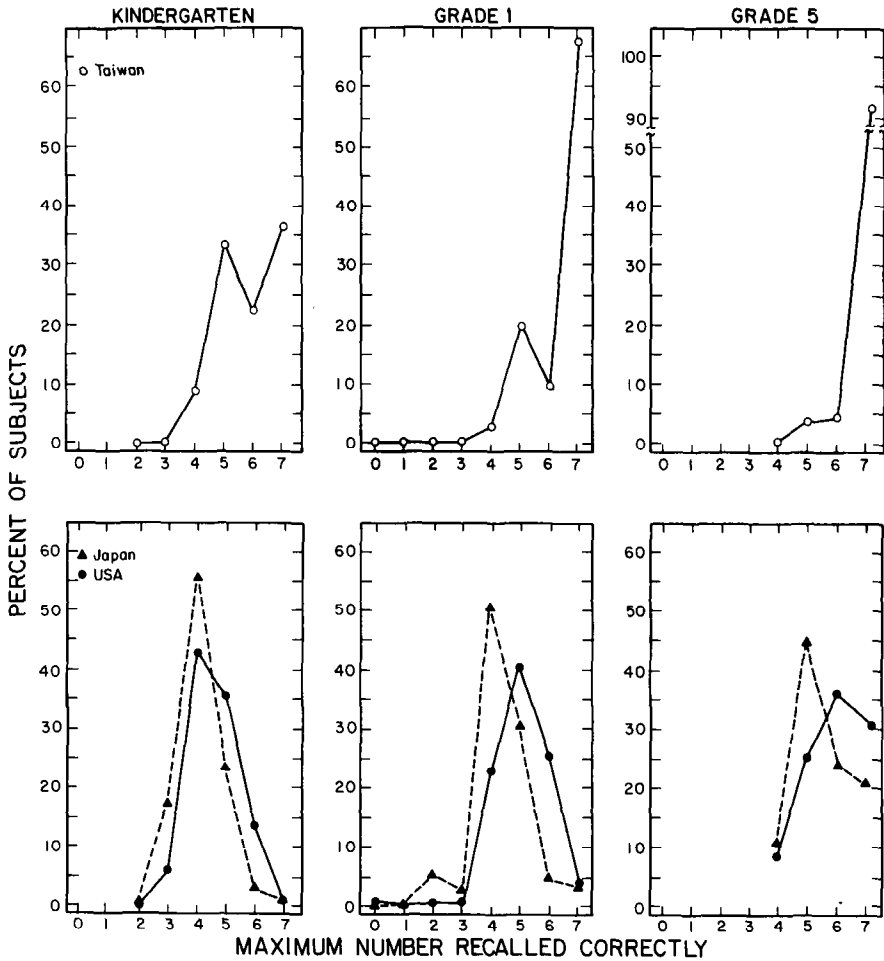
The performance of children in kindergarten and grades 1 and 5 on the digit span test is presented in Figure 1. The lower panel presents the distribution of scores for the Japanese and American children while that for the Chinese children is presented in the upper panel.

It is immediately clear that the Japanese and American distributions are similar to each other, but differ greatly from the distributions for the Chinese children. While the Japanese and American distributions are roughly normal, the Chinese distributions are skewed highly toward the upper end of the scale. Among Chinese children, a span of 7 digits was found for 37% of the kindergarteners, 68% of the first-graders, and 92% of the fifth-graders.

Ceiling effects prevent our estimating the true length of the digit span for the Chinese children. Nevertheless, the mean digit spans of the Chinese, Japanese and American kindergarteners were 5.9, 4.1 and 4.6, respectively; 6.4, 4.4. and 5.1 for the first-graders; and 6.9, 5.5 and 5.9 for the fifth-graders.

Analyses of variance revealed highly significant effects of country at all

Figure 1. Distribution of digit span scores for kindergarten, first-, and fifth-grade children in Japan, Taiwan, and the United States.



grade levels, $F_s(2,717-869) > 158$, $p_s < .0001$. Sheffé contrasts revealed that at all grade levels Chinese children outperformed both American and Japanese children, $p_s < .0001$; and that American children outperformed the Japanese children, $p_s < .01$.

One explanation for the markedly different distributions for Chinese children is that the Chinese testers administered the task in a manner not comparable to that used by examiners in the other two locations. This explanation

is rendered implausible by the fact that the kindergarten children were tested 3 years after the first- and fifth-graders, and by a different set of examiners. It is highly unlikely that some deviation in the procedure would have been continued through such vast changes in circumstance.

An even better control for variations in method was provided by the inclusion of the word span test at grades 1 and 5. Since the same examiners gave the word span and digit span tests, and since the instructions for the two tests were identical, one would expect any methodological deviation to be reflected in both the digit span and word span results. This was not the case. Chinese first- and fifth-graders did not receive significantly higher scores than the American children, though both Chinese and American children outperformed the Japanese children, $p < .01$. The mean numbers of words recalled correctly across all trials for the Chinese, American, and Japanese children were, respectively, 3.7, 3.6, and 3.5 at grade 1, and 4.2, 4.3, and 4.1 at grade 5.

Another hypothesis is that the Chinese counting system was responsible for the differences in digit span. Counting in Chinese is as follows: one, two, three..., ten-one, ten-two, ten-three..., two-ten-one, two-ten-two... etc. Perhaps this system gives an emphasis to the sequence and organization of numbers that is lacking in counting words used in English, where different words are used to describe the increments of ten. This hypothesis can be rejected immediately, for both Chinese and Japanese use the same counting system and the Japanese children did not display exceptional performance on the digit span task.

In summary, a large advantage in digit span was demonstrated for Chinese children as compared to their Japanese and American peers. The difference appeared at kindergarten, first and fifth grade, and did not appear to be the result of methodological artifacts, a more general superiority in serial memory, or the counting system used in Chinese. It is important to note, too, that the Chinese and American children received significantly higher scores than the Japanese children at all three grade levels.

Study 2: Digit span in forward, backward, and grouped conditions

The second study was intended to make a start in specifying the cognitive basis for the observed Chinese advantage. Because the existence of the difference was no longer an issue, we believed that it was no longer crucial to use large, representative samples of children. Rather, in this study smaller samples of children were studied under several different conditions.

The question that motivated this study was that of whether or not the

Chinese digit-memory advantage was specific to a forward digit span task, or whether it also would be found when children were asked to repeat the digits backwards, or when the digits were presented in a grouped fashion. The grouped condition was intended to assess whether the cross-national differences in digit span could be tied to differential use of a grouping strategy between Chinese and American subjects. If Chinese children spontaneously employ some kind of grouping strategy, a strategy frequently used on the digit span test (see, for example, Dempster, 1981), grouping of digits should not be especially helpful in recall. If, on the other hand, Chinese children do not spontaneously produce strategies for remembering in any greater degree than do the American children, the performance of American and Chinese children should be improved to an equal degree with grouped presentation of digits.

Method

Twenty 6-year-olds (range from 5.9 to 6.2 years) and 20 7-year-olds (range from 6.9 to 7.2 years) were chosen from kindergarten and first-grade classrooms both in Taipei and in Chicago. Most of the 80 children attended schools associated with a university community, and most came from upper-middle socioeconomic backgrounds. Testing occurred near the end of the school year.

Procedure

All children were presented with three tasks: forward, backward, and grouped digit span, in that order. Children were randomly assigned one of the five sequences of randomly generated digits that were used for each task.

All three tasks were presented visually instead of aurally as in the earlier studies. For forward and backward conditions, digits were presented one at a time written on 5 cm × 5 cm square cards. Cards were placed on a table in front of the subject at a rate of one per second. Each successive card was placed on top of the previous card, thereby obscuring it. After all cards had been presented, the final card was covered with a blank card and the child was asked to repeat the digits in either forward or backward order.

Each test began with a string of 2 digits, and continued to a maximum length of 10 digits. Two different sequences of digits of each length were presented until both strings of each of two successive lengths were repeated incorrectly. Each child's score was the number of digits repeated correctly in the string of maximum length.

The grouped digit span task was exactly the same as the forward and backward tasks, except that the digits were presented in groups of three. For example, a string of 5 digits would be presented on two cards, the first containing 3 digits and the second 2. A string of 6 would be presented as 3-3, 7 as 3-3-1, etc. The digits were presented at a rate of 1 per second, so that, for example, a card with 3 digits would be visible for 3 seconds.

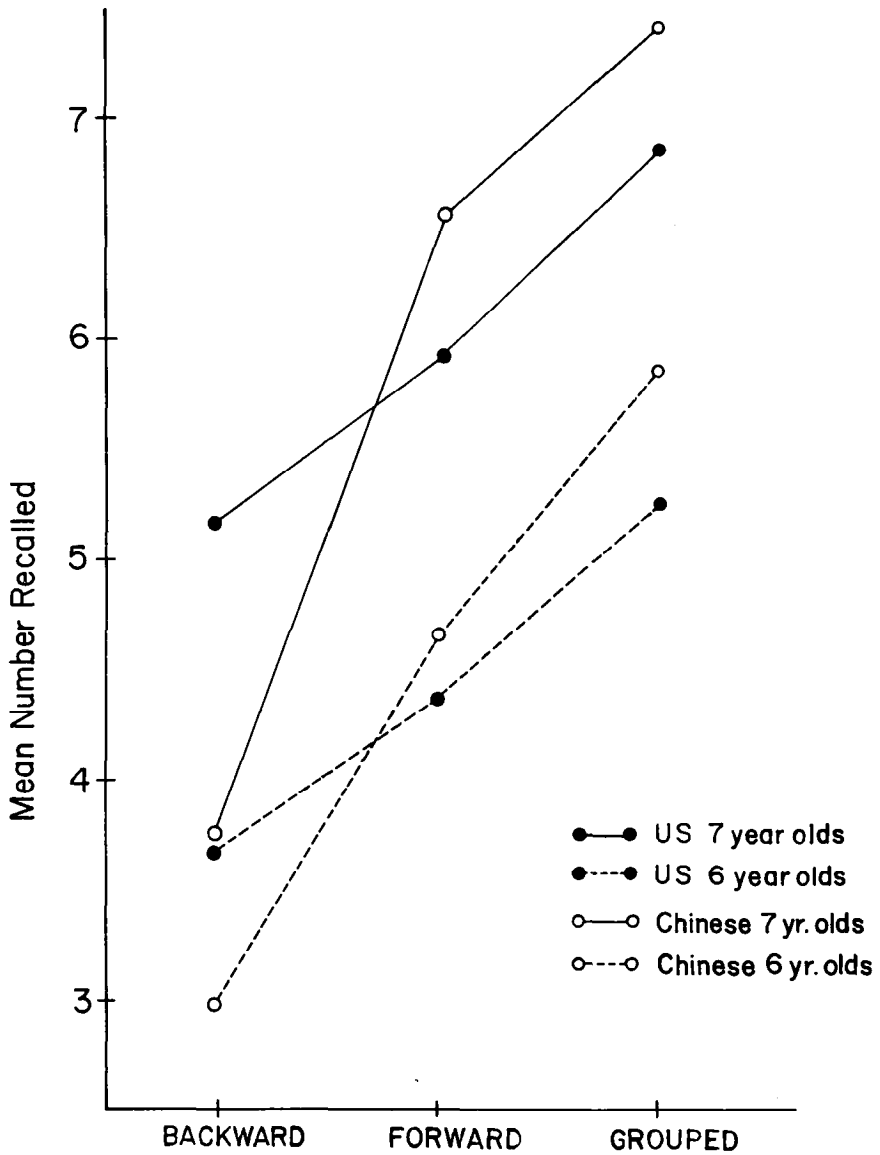
Results

Data for Study 2 were analyzed by a 2 (Country) \times 2 (Sex) \times 2 (Age) repeated measures analysis of variance, with Condition (3 levels) as a repeated measure. The overall analysis was highly significant, $F(95,144) = 6.04$, $p < .0001$. There were two significant main effects. The first main effect was for condition, $F(2,144) = 246.10$, $p < .0001$; the second was for age, $F(1,72) = 40.44$, $p < .0001$. We were primarily interested in the country \times condition interaction. This was the only significant interaction, $F(2,144) = 33.20$, $p < .0001$. Directions of the effects can be seen in Figure 2, where the data are graphed by age group, nationality, and condition. Seven-year-olds received higher scores than 6-year-olds, and the pattern across countries and conditions was comparable at the two age groups.

The Chinese tended to score higher than the Americans in both the forward and grouped conditions, but post-hoc comparisons showed that neither difference was statistically significant. In the backward condition we obtained a very different result: Chinese children performed significantly more poorly than American children at both age levels, $p < .0001$. Post-hoc analyses also revealed that the drop in performance between the forward and backward conditions was significantly greater for the Chinese than for the Americans, $p < .0001$, and that the improvement in performance of the Chinese and American children between the forward and grouped conditions did not differ significantly.

In summary, the results of Study 2 did not support the proposal that Chinese children made more effective use of strategies in the digit span task than the American children. When digits were presented in groups, thereby providing a grouping strategy, performance improved equally for Chinese and American children. On the other hand, in the more demanding task of repeating the digits backwards, the performance of both groups fell, and the decline was significantly greater for the Chinese than for the American children. Contrary to prior expectations, the Chinese backward digit span was significantly shorter than the span for American children.

Figure 2. Mean number of digits recalled in forward, backward and grouped conditions by 6- and 7-year-old children in Taiwan and the United States.



Study 3: Number word duration and digit span in Chinese and American students

The final study was designed to investigate the intriguing hypothesis that cross-linguistic differences in the time required to pronounce number words in Chinese and English may account for the observed differences in digit memory. The rationale for this hypothesis is derived from work by Baddeley, Thomson and Buchanan (1975). Baddeley et al. presented evidence that the number of words an individual can hold in short-term memory is limited, not by the number of “chunks” of information represented, but rather by the amount of time it takes to pronounce the words. Thus, more words can be remembered when their pronunciation requires a smaller amount of time. Moreover, given any particular list of words, the individual who pronounces the words more quickly will be able to remember more of the list than will the individual who pronounces them more slowly.

Implications of this theory for cross-linguistic comparisons of digit memory have been pursued in a study by Ellis and Hennesly (1980). Welsh digit words appear to have longer pronunciation durations than do English digit words. Ellis and Hennesly explored the effect this might have on short-term memory for numbers with bilingual speakers of English and Welsh, some of whom were dominant in Welsh, and some in English. There were two tasks. The subjects were asked to read a random list of 200 single-digit numbers as fast as they could, once in English and once in Welsh and then they were given a standard digit span test in both languages.

The results were in line with what would be predicted by Baddeley et al. Regardless of the dominant language, the average duration of pronunciation for Welsh digits was significantly greater than that for English digits; and digit span was significantly greater in English than it was in Welsh.

On the basis of these findings, we decided to test our own hunch that number words in Chinese require shorter times for pronunciation than the number words in English. If this was found to be the case, then it would seem to be a possible explanation for the cross-national differences in digit span.

Method

Subjects

There were 31 subjects: Thirteen were graduate students at the University of Chicago who were native speakers of Mandarin Chinese from Taiwan. In order not to bias the results by sampling heavily from among students in the

physical sciences and mathematics whose work often requires daily involvement with numbers, we sampled only students working toward advanced degrees in the humanities or social sciences.

A comparison sample of 18 native born English-speaking Americans was constituted from an already existing data set. The English-speaking subjects were all undergraduates at the University of Chicago, and were selected without regard to major.

The materials and procedures differed somewhat across the comparison samples, for the data came from two separate studies conducted by different experimenters at different times. These differences will be described in detail in the following section. We should state at the outset, however, that we do not believe these differences influenced the validity of the results obtained.

Materials and procedures

All subjects were tested on two tasks, always presented in the same order. Both tasks were presented through headphones. The subjects were tested in a soundproof booth. Two tape recorders were used: one presented the stimulus materials; the other recorded both the stimuli and the subject's responses to the stimuli. All stimuli for the Chinese subjects were in Mandarin Chinese, and those for the English speakers were in English.

The first task was a digit naming task. The subject was asked to repeat a number word from one to nine spoken by a voice on the tape. For example, the voice on the tape would say "one," and the subject would say "one." Subjects were not given any instructions regarding the speed with which they should do so. Chinese subjects were presented with each of the nine digits twice; English speakers received each digit only once. In each case, the digits were presented in a single random order for all subjects, and no digit was repeated until all nine digits had been presented once.

The second task was a standard assessment of digit span. The subjects were asked to repeat back randomly ordered strings of digits after a voice on the tape. All subjects began the task by attempting to repeat a string of 5 digits. They were then asked to repeat a different string of 5 digits. Subjects who were able to recall correctly one of the strings of 5 digits were permitted to move on to a string of 6 digits, and so on until both strings of a given length were recalled incorrectly.

For both groups of subjects, stimuli were composed of randomly ordered strings of digits from 1 to 9, generated under the constraint that no digit could be repeated until all nine digits had been presented once.

A major difference across the two groups of studies was the rate of presentation: while Chinese subjects were presented the digits at the standard rate

of one per second, English-speaking subjects heard the digits at a rate of 2 per second. Although this difference appears significant, we do not believe that it has an important effect on the results obtained for two reasons. First, previous research has shown that variations in rate of presentation within the range included here does not affect digit span (Dempster, 1981). Second, we were more concerned with differences in pronunciation duration than with differences in span, since the span difference had already been adequately demonstrated in Study 1. While the slower presentation rate might have given the Chinese subjects more time to engage in rehearsal, it is unlikely that such rehearsal would alter the pronunciation duration of individual number words.

Coding and reliability

Tapes of subjects' responses were digitized and displayed on an oscilloscope. A computer program allowed for the placement of cursors at the beginning and the end of the subjects' speech, so that the duration of subjects' articulation could be measured in milliseconds. Variables that were measured are described below.

Digit naming task

Pronunciation duration for each of the digits 1 to 9 was measured for each subject. Durations were averaged across all of the nine digits to produce a summary variable. Pronunciation duration was defined as the time elapsed from when the subject first began to pronounce the word to the point at which pronunciation was completed. (Since Chinese subjects were given two trials for each of the nine digits, separate analyses were conducted for the average duration, duration on the first trial, and duration on the second trial. All three analyses yielded essentially the same results, thus results are based on the average duration across the two trials.)

Digit span task

Several variables were derived from the digit span task. A subject's digit span was defined as the longest string of digits recalled correctly. The duration of each subject's response was measured as well. This was defined as the time elapsed, on correct trials only, between the moment the subjects began responding to a given string of digits, and the moment when pronunciation of the final digit in the string was completed. Two variables were derived from these measurements: first, each subject was assigned an average duration per digit for all strings recalled correctly by the subject. This was computed by dividing the duration of the entire response string by the number

of digits in the string. Second, each subject was assigned a duration per digit value for the longest string of digits recalled correctly. (The use of the second measure will be clarified in the results section.)

Reliability

All coding was carried out by one individual. Reliability of coding was assessed by having the coder recode two subjects' data (one Chinese and one American) independently and blindly, several weeks after they were first coded. The Pearson r between values obtained for both the digit naming task and the digit span task on first and second coding were greater than .95 for both subjects. Thus, the coder made consistent judgments and coding for the two languages was reliable.

Results

As predicted, the Chinese subjects were able to recall significantly longer strings of digits than were the American subjects. The mean digit span for the Chinese subjects was 9.2, and for the Americans, 7.2. As can be seen in the distributions of digit spans for the two samples depicted in Figure 3, spans

Figure 3. *Distribution of digit span scores for Chinese and American university students.*

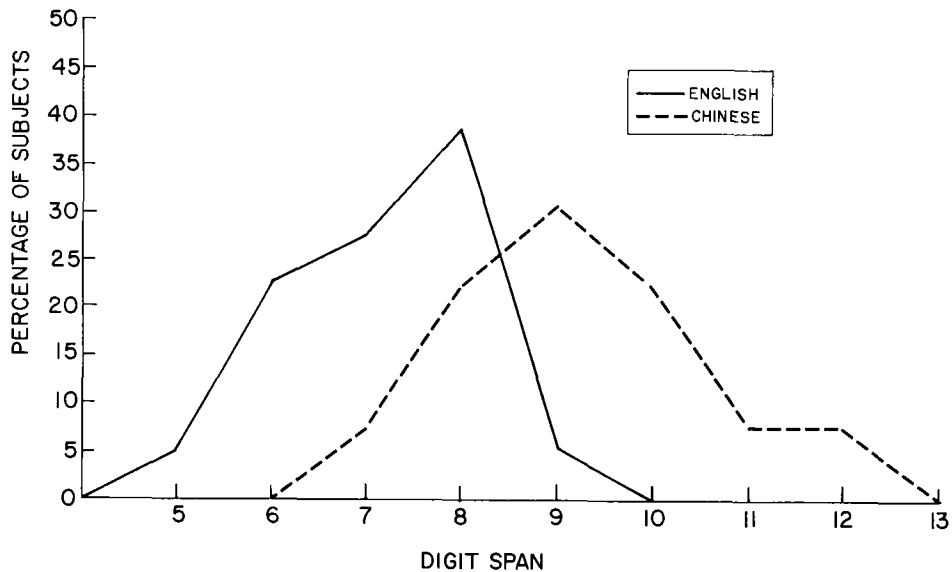
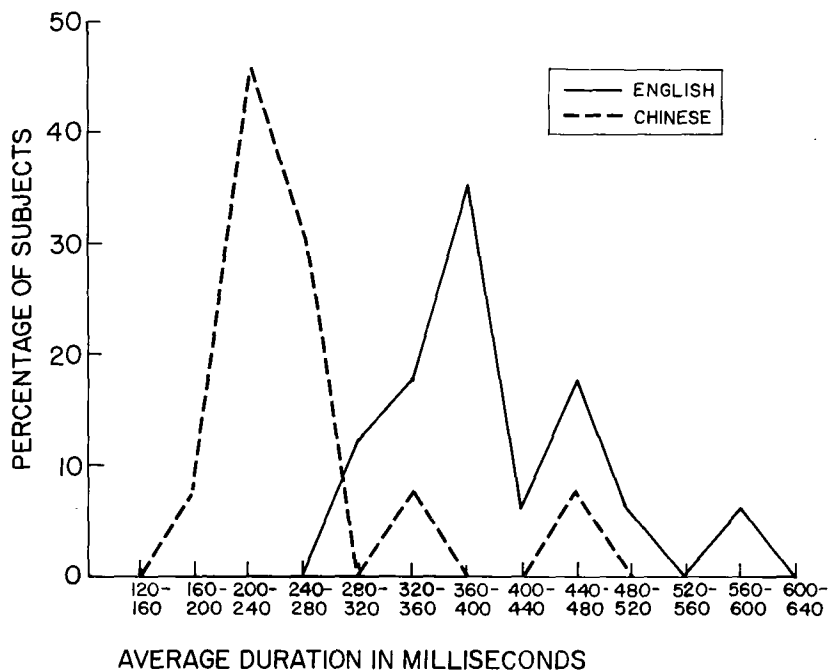


Figure 4. *Distribution of average pronunciation durations per digit recalled by Chinese and American students in the digit span task.*



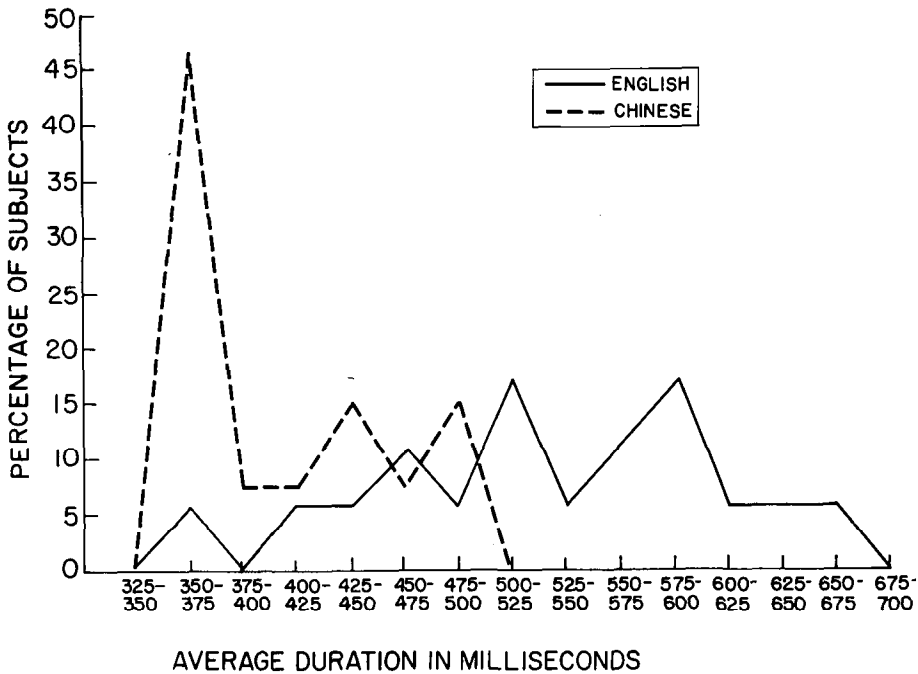
for the Chinese subjects ranged from 7 to 12 digits, while those for the Americans ranged from 5 to 9 digits.

In Figure 4, the corresponding distributions of response durations for the digit span task are presented in milliseconds per digit. The Chinese subjects had markedly faster response times. While the modal Chinese subject recalled the digits at a rate of one every 200 to 240 milliseconds, the modal English-speaking subject recalled the digits at about half that speed.

While this overall difference in speed of responding is an interesting one, it is not exactly what Baddeley and his associates would theorize to be the crucial measure. Overall speed of responding is not only a function of the pronunciation duration of the words to be recalled, but also of the space subjects leave between the digits they are recalling.

A more relevant test of Baddeley's hypothesis lies in the pronunciation durations derived from the digit naming task, where the duration of each individual number word was measured. Distributions of pronunciation durations, averaged for each subject across the nine number words, are presented in Figure 5.

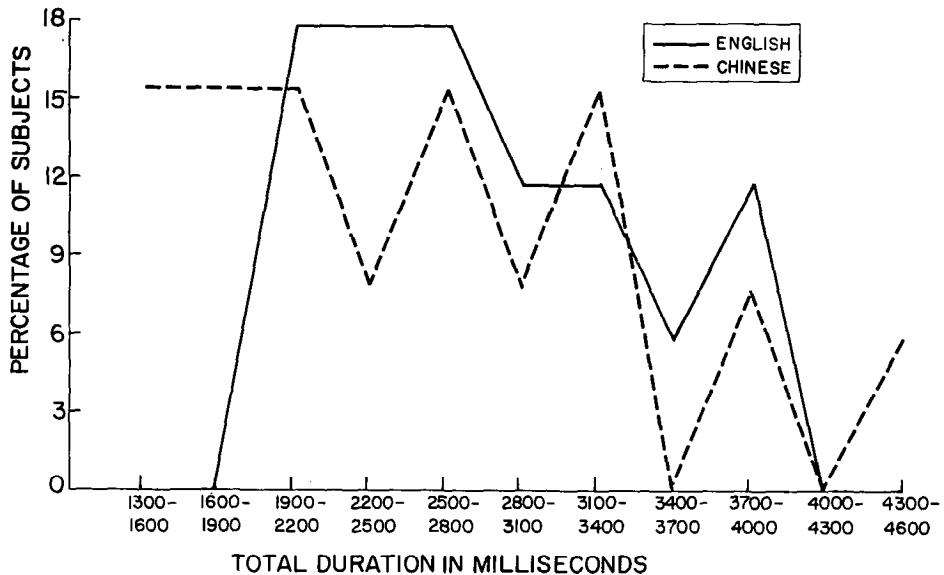
Figure 5. *Distribution of average durations for pronouncing digits in the digit naming task by Chinese and American students.*



Two striking differences between the Chinese and English distributions appear in Figure 5. First, there is a large difference in the average duration with which the number words are pronounced in the two languages. For Chinese subjects the average was 406 ms/digit, while in English the average was 527 ms/digit, $t(29) = 5.19$, $p < .001$. Second, there is an obvious difference in the shape of the distributions across the two samples. The distribution for English speakers is much more variable than that for Chinese speakers.

One further analysis is of interest. Baddeley proposed that while the pronunciation duration may vary across different subjects and different materials, the total duration of the maximum number of items a subject can recall should be constant across subjects; i.e., a constant of the human information processing system. In other words, while the number of items that can be retained in short-term memory will vary, mostly due to variation in pronunciation duration of the items, the total short-term memory span as measured by duration will remain constant. In order to assess this hypothesis within the present comparison, response durations for the longest string of digits cor-

Figure 6. *Distribution of total response durations for longest string of digits recalled correctly in the digit span task by Chinese and American students.*



rectly recalled by each subject were analyzed. The distributions of these times are presented in Figure 6. While the range of maximum durations was quite large for both groups, the most interesting feature of this graph is the fact that the average durations for maximum spans did not differ significantly between the two groups, $t = 1.85$, $p > .05$. (The average for Chinese subjects was 2423 ms, and for Americans, 2905 ms.)

Discussion

These data offer support for the hypothesis that differences in digit memory between Chinese and English speakers are derived, in part, from differences in the time required to pronounce number words in the two languages.

Such differences in pronunciation duration could originate from two distinct sources. They could be due either to inherent differences in the linguistic structure of the words themselves or to the amount of time spent practicing the pronunciation of number words. Such practice could, for example, underlie observed developmental decreases in the pronunciation duration of words (Huttenlocher, 1984).

It is premature to conclude that the cross-linguistic difference that we have documented in pronunciation duration of number words is due primarily to either linguistic or experiential factors. However, some of our observations may support the practice hypothesis. For example, the fact that our Chinese subjects assumed that their responses to the digit span task should be made as quickly as possible corroborates informal observations of mathematics classes in Chinese schools where speed and repetitive practice are emphasized far more than they are in American classrooms. Years of having speed of response emphasized could well culminate in faster pronunciation of number words. (Such an explanation, however, would not account for the data from the kindergarten children.) Further, if language were a more important factor than practice in producing longer pronunciation durations, the variability in pronunciation durations should be similar across the two cultures but the means should be different. As is evident in Figure 5, however, the times for the Chinese subjects were less variable than were those for the American subjects, perhaps indicating that many of the Chinese subjects but few of the American subjects had sufficient experience to reach an asymptote in their speed of pronouncing number words.

Although it was not the purpose of the present group of studies to investigate possible bases for the poor performance of Japanese children in the digit span test, the hypothesis of pronunciation duration appears to be relevant. Japanese digits are likely to require longer times than either Chinese or English. Many of the Japanese digits have two syllables (i.e., *ichi*, *ni*, *san*, *shi*, *go*, *roku*, *shichi*, *hachi*, *ku*, *ju*), while the Chinese digits have only one (*i*, *er*, *san*, *si*, *wu*, *liu*, *chi*, *ba*, *jiu*, *shi*).

A team of Chinese researchers (Zhang & Simon, 1985; Yu, Zhang, Jing, Peng, Zhang, & Simon, 1985) recently have proposed a model of short-term serial memory that successfully combines the effects of chunking and of pronunciation duration into a single model. They demonstrate that the number of chunks recalled (C) across a broad range of studies using both Chinese and English language materials can be predicted quite well by a combination of three parameters: T , which signifies Baddeley et al.'s fixed time interval that constrains short-term memory capacity; a , the amount of time it takes to bring each new chunk into the articulatory mechanism and articulate the first syllable; and b , the time it takes to articulate each syllable (S) in the chunk beyond the first one. Thus,

$$C = T/[a + b(S - 1)].$$

It is possible, through a few simple calculations, to see roughly how well this equation fits our data on the Chinese and English digit spans. Since, on average, the number words in Chinese and English contain only one syllable,

if we assume that each digit is a chunk then $b(S - 1)$ becomes equal to zero, leaving us with

$$C = T/a.$$

If we let T , which is generally assumed to be 2–3 s, be equal to 3000 ms, and if we substitute the mean number of digits recalled for Chinese and for English speakers from our Study 3 (9.2 and 7.2, respectively), then we can solve for a for each group. The resulting values of a are 326 ms for the Chinese group, and 417 ms for the English-speaking group. The difference between these values, 91 ms, is roughly comparable to the 121 ms difference found between the Chinese- and English-speaking subjects' average pronunciation durations for the number words. Thus, this admittedly approximate analysis is consistent with the hypothesis that the difference in digit span across the two groups is primarily due to differences in articulation duration of the number words, and not to differences in the time it takes to bring each new chunk into the articulatory loop.

Conclusion

We conclude that there is a profound difference in the forward digit span of Chinese-speaking and English-speaking subjects, and that differences in the time required to pronounce digits in Chinese and in English is the most likely explanation for this difference.

As this paper was being written, we discovered a recent paper by Hoosain (1984) that provided findings that paralleled our own. Hoosain found that Chinese superiority for digit memory occurs among Cantonese speakers of Chinese residing in Hong Kong. When Cantonese-speaking undergraduates were asked to pronounce a list of 200 random numbers as quickly as possible (following the procedure used by Ellis and Hennesly, 1980), Hoosain found a mean pronunciation duration per digit of 265 ms, a value considerably faster than the 321 ms found by Ellis and Hennesly for English. Even more interesting is that Hoosain, in a study of Cantonese–English bilinguals, found an interaction between language and order of response (forward vs. backward). As in our results in Study 2, forward digit span in English was shorter than in Chinese, but the advantage of forward over backward digit span was greater in Chinese than in English. While the reason for this interaction is unclear, the fact that it has been found in two studies employing different research designs and age groups, makes the phenomenon worthy of further investigation.

References

- Baddeley, A.D., Thomson, H., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 14, 575–589.
- Binet, A., & Henri, V. (1895). La psychologie individuelle. *L'Année Psychologique*, 2, 411–465.
- Dempster, F.N. (1981). Memory span: Sources of individual and developmental differences. *Psychological Bulletin*, 89(1), 63–100.
- Ellis, N.C., & Hennelly, R.A. (1980). A bilingual word-length effect: Implications for intelligence testing and the relative ease of mental calculation in Welsh and English. *British Journal of Psychology*, 71, 43–51.
- Hoosain, R. (1984). Experiments on digit spans in the Chinese and English languages. In H.S.R. Kao & R. Hoosain (Eds.), *Psychological studies of the Chinese language*. Hong Kong: The Chinese Language Society of Hong Kong.
- Husen, T. (1967). *International study of achievement in mathematics: A comparison of twelve countries*. New York: Wiley.
- Huttenlocher, J. (1984). Word recognition and word production in children. In H. Bouma & D.G. Bouwhuis (Eds.), *Attention and performance Vol. 10: Control of language processes*. Hillsdale, NJ: Erlbaum.
- Stevenson, H.W., Stigler, J.W., Lee, S.Y., Lucker, G.W., Kitamura, S., & Hsu, C.C. (1985). Cognitive performance and academic achievement of Japanese, Chinese, and American children. *Child Development*, 56, 718–734.
- Stigler, J.W., Lee, S.Y., Lucker, G.W., & Stevenson, H.W. (1982). Curriculum and achievement in elementary school mathematics: A study in Japan, Taiwan, and the United States. *Journal of Educational Psychology*, 74, 315–322.
- Yu, B.L., Zhang, W.T., Jing, Q.C., Peng, R.X., Zhang, G.J., & Simon, H.A. (1985). STM capacity for Chinese and English language materials. *Memory and Cognition*, 13(3), 202–207.
- Zhang, G.J., & Simon, H.A. (1985). STM capacity for Chinese words and idioms: Chunking and acoustical loop hypotheses. *Memory and Cognition*, 13(3), 193–201.

Résumé

Cet article rapporte 3 études qui comparent la mémoire à court terme pour les nombres de locuteurs dont la langue maternelle est le chinois ou l'anglais. À partir d'un grand échantillon d'enfants d'école maternelle et des deux premières années d'école primaire, la première étude a établi que les locuteurs du chinois sont capables de mémoriser plus de nombres que les locuteurs de l'anglais. Tous les sujets chinois pouvaient mémoriser en moyenne au moins deux nombres de plus que les sujets américains. La seconde étude comparait la mémoire pour les nombres d'enfants chinois et américains âgés de 6 à 7 ans dans différentes conditions de présentation ou de rappel. Une présentation groupée des nombres facilitait la tâche de la même manière pour les enfants chinois et américains, ce qui semble indiquer que la différence de mémoire pour les nombres n'est pas due à l'utilisation d'une stratégie. De plus, la performance des enfants chinois était moins bonne que celle des enfants américains lorsqu'on leur demandait de répéter les nombres dans l'ordre inverse de l'ordre de présentation. La dernière étude, entreprise sur des étudiants d'université américains et chinois, a permis de se pencher sur les différences de longueur de prononciation entre les nombres anglais et chinois en tant qu'explication possible des différences de mémorisation. Les nombres en chinois sont plus courts que les nombres en anglais; et la longueur de prononciation totale de la liste mémorisable était la même pour les sujets chinois et américains. Ces résultats militent en faveur d'une mémoire temporellement limitée.