Using both an auditory and a visual short-term store to increase digit span

ROBERT W. FRICK

University of Washington, Seattle, Washington

Can digit span be increased by storing digits nonredundantly in both an auditory short-term store and a visual short-term store (dual storage)? In Experiment 1, when four digits were presented visually and then the remaining digits were presented auditorily, digit span increased three digits over a baseline measurement, but only when the auditorily presented digits were reported first (inverted response). Normal order of response for this presentation was not as effective. Possible reasons for the advantage of inverted response are suggested, based upon dual storage. For all-auditory and all-visual presentations that controlled for parsing, digit span increased only 1 to 1.5 digits over baseline, and inverted and normal order of response did not differ. Experiment 2 demonstrated that an improvement for dual storage could occur when the type of presentation was held constant and the order of response was normal rather than inverted. The results are consistent with the idea of separate auditory and visual short-term stores.

Auditory short-term memory and visual short-term memory have been shown to be separate, in the sense that two tasks performed simultaneously in differing sensory modalities do not necessarily interfere with each other. Specifically, performance of two tasks is more impaired when the tasks are performed in the same modality than when they are performed in different modalities (Brooks, 1968, 1970; Proctor, 1978; Rollins & Hendricks, 1980; Segal & Fusella, 1970). Rollins and Hendricks concluded:

The results . . . support the hypothesis that the visual and auditory modalities can process verbal material independently and without interference with each other This occurs with little practice and thus appears to be a basic property of these two information-processing systems. (p. 106)

The capacity of short-term memory is normally measured by the number of stimuli that can be correctly recalled in order immediately after presentation (memory span). If information can be processed and retained in two short-term stores without interference, an individual's total capacity should equal the sum of the capacities of those two short-term stores. Tests of memory span usually measure the capacity of an auditory store [Conrad, 1964; this store may be either acoustic or articulatory (Wickelgren, 1969)]. The capacity of this auditory store appears to be seven plus or minus two digits (Miller, 1956). Tests of memory span that use visual presentation and eliminate auditory recoding by having the subject

speak aloud during presentation reveal a store that is not auditory (Baddeley, Thompson, & Buchanan, 1975; Conrad, 1972; Murray, 1967). When the stimuli are presented simultaneously, a visual store is used (Frick, 1984). If the auditory store and this other nonauditory, presumably visual store could be used nonredundantly, perhaps digit span could be increased.

Previous attempts to increase immediate recall by using both auditory and visual stores have not been too successful. Dornbush (1968) presented auditory and visual digits concurrently and reported no improvement in recall. Murdock and his associates (Murdock & Carey, 1972; Murdock & Walker, 1969) found an improvement for presenting some digits visually and some auditorily. However, this improvement was small (.6 to .8 digits), and an important control was lacking: Having two types of presentation parses the digits (i.e., separates them into groups), and parsing can improve recall as much as 1 digit (Ryan, 1969). Therefore, the increase Murdock found could have been a consequence of parsing the digits.

The following experiments attempted to increase immediate recall by using both an auditory and a visual short-term store.

In Experiment 1, the subjects were first given a test of digit span using auditory presentation. This allowed the subjects to become familiar with the testing situation, and it provided an estimate of ability that was used for determining the starting level for subsequent testing and for data analysis. The subjects were then tested with an all-auditory, an all-visual, or a mixed presentation. In the mixed presentation, four digits were presented visually, and then the remaining digits were presented auditorily.

Digit span was measured with the method suggested by Woodworth and Schlosberg (1954). Testing started at a

Christopher Carter, James Corkern, and Elizabeth Johnson helped test subjects; Lawrence Dunkel assisted with computer programming; and Sandra Frankmann, Susan MacDonald, Thomas Nelson, and Mary Smith criticized earlier drafts of this paper. Requests for reprints should be sent to Robert Frick, Department of Psychology, NI-25, University of Washington, Seattle, WA 98195.

given level, and the number of digits presented was slowly incremented until the subject could not recall any digit strings on a level. Digit span was defined as a function of the total number of strings correctly recalled.

Several reasonable, but to varying degrees nontraditional, manipulations were made in order to facilitate dual storage of the digits. First, the subjects receiving a mixed presentation were asked to store the visually presented digits visually and the auditorily presented digits auditorily.

Second, except for the baseline measure, testing began at a level estimated to be slightly beyond the subject's auditory capacity. Within the range of the subject's auditory capacity, all-auditory storage is probably more effective than dual storage. If so, testing in this range would have encouraged all-auditory storage. By beginning testing outside the subject's auditory capacity, the subjects were encouraged to use dual storage.

Third, in the visual portion of the mixed presentation, the digits were presented simultaneously. Penney (1975) suggested that a simultaneous presentation might be more conducive to visual memory than a sequential presentation. (A pilot experiment had also suggested that a simultaneous presentation was more effective.) To control for other possible advantages of the simultaneous presentation, the all-visual control condition had a simultaneous presentation.

An auxiliary hypothesis of Experiments 1 and 2 was that speaking the digits in the visual store would interfere the digits in the auditory store, but speaking the digits in the auditory store would not interfere with the digits in the visual store. This was expected as an example of modality-specific interference, in which speaking is an auditory act that would interfere selectively with auditory short-term memory. Margrain (1967) found that the recall of auditorily presented digits was more impaired by spoken than by written recall of visually presented digits, and the recall of visually presented digits was more impaired by written than by spoken recall of auditorily presented digits. In Experiment 1, responses were spoken, so an inverted response beginning with the auditory portion was expected to avoid interference. The subjects were tested for both normal and inverted order of response.

Finally, in the mixed presentation, the visually presented digits were presented slightly longer than the auditory digits (600 vs. 450 msec per digit). A rapid presentation of the auditory digits reduced the length of time the visual digits would need to be remembered, yet was not expected to impair the passive storage of digits in the auditory short-term store (Hockey, 1973). Presentation times were still within the normal range for testing digit span, and the all-auditory and all-visual presentations had the same timing as the mixed presentation.

To control for effects of parsing produced by the mixed presentation, the auditory presentation was parsed rhythmically by inserting a pause after the fourth digit, and the visual presentation was parsed spatially by having the first four digits on one line and the remaining digits on a lower line.

EXPERIMENT 1

Method

Subjects. Sixty students from introductory psychology classes at the University of Washington volunteered in order to earn extra credit. English was their first language.

Design. There were three groups of 20 subjects each. All groups were first given a baseline test of digit span. Then the groups were tested twice, once with normal order of response and once with inverted order, beginning with the fifth digit. The normal response and the inverted response were counterbalanced for order. For one group, the experimental presentations were mixed, for one group the experimental presentations were all auditory, and for one group the experimental presentations were all visual. The three types of presentation and the two types of response created six conditions: mixed-inverted, mixed-normal, visual-inverted, visual-normal, auditory-inverted, and auditory-normal.

Stimuli. The digits 0 through 9 were presented by computer, either visually on a cathode ray tube or auditorily through a speaker. The auditory digits had been previously recorded and stored in the computer by an analog-to-digital conversion, and were presented by a digital-to-analog conversion. The actual spoken duration of the individual auditory digits ranged from 210 to 440 msec. A pause was inserted between digits during presentation so that the time from the middle of one digit to the middle of the next was held constant.

In the baseline test, digits were presented auditorily at 450 msec per digit. In the mixed presentation, four digits were presented visually and simultaneously for 2,250 msec, there was a 150-msec pause, and then the digits were presented auditorily at 450 msec per digit. In the all-auditory presentation, digits were presented at 450 msec per digit, with a 600-msec pause between the fourth and fifth digit. In the all-visual presentation, the digits were presented for 450 msec times the total number of digits plus 600 msec. The first four digits were on one line, and the remaining digits were one line down and began four spaces over.

The digits in a string were generated randomly, with several constraints: (1) If the first digit was less than 5, the second digit was greater than or equal to 5, and vice versa; (2) although digits could recur within a string, the recurrences were separated by at least two digits; (3) a digit could not differ from the preceding digit by one; and (4) two adjacent differences between digits could not be equal (e.g., the sequences 579 or 852 did not occur).

Measuring digit span. Digit span was measured with the method suggested by Woodworth and Schlosberg (1954, p. 697). In the baseline condition, testing began at a string length of five digits. For the two experimental conditions, testing began at a length based upon the baseline performance—two more than the highest level on which the subject made no errors and had not previously made two or more errors.

The subjects were tested three times at each length, except that, when the first two trials were correct, the third trial was presumed to be correct and was not tested. The length of the string was incremented by one digit, until a subject had missed all three trials at a length. If the subject did not get any correct on the initial string length, the string length was decremented by one, and testing proceeded backward until at least one trial at a string length was correct.

A response was defined as correct when all the digits were recalled in the correct order. Digit span was defined as the number of trials correctly recalled plus the number of trials presumed to be correct (string lengths less than the starting level and the third trial whenever the first two were correct), divided by three. For example, a subject who recalled all strings shorter than or equal to seven digits would receive a score of 7. A subject who did not recall one of the strings at length seven but recalled one string at length eight would also receive a score of 7, having recalled the same total number of strings. This score is approximately .5 digits less than the number of digits the subject would recall 50% of the time.

Procedure. The subjects were tested individually. The general procedure was explained to the subjects at the start of the experiment, and each condition was described before that condition was begun.

The subjects receiving a mixed presentation were told that the theory of the experiment was that digits could be stored both auditorily and visually. They were asked to form a picture of the visually presented

digits and to remember the sound of the auditorily presented digits. To control for the effect of giving instructions, the all-auditory group was told that the theory of the experiment was that people can remember digits better if the digits are presented rhythmically, and the all-visual group was told that the theory of the experiment was that people can remember digits better if the digits are grouped.

A trial began with a message on the screen; the message lasted 5 sec and indicated how many digits would be presented. The word "ready" then appeared on the screen for 3 sec, and then a dot appeared on the screen for 1 sec where the visual digits would start. Then the digits were presented, and the subjects spoke their responses. The experimenter recorded the responses and signaled the computer whether or not these responses were correct. After an approximately 5-sec pause, the next trial began.

Results

The average digit span for each condition is presented in Figure 1. Despite the fact that the groups were treated identically until after the baseline test of digit span, their scores on the baseline test differed substantially [$F(2,57) = 4.25$, $p < .05$]. This difference in overall ability between groups made a direct comparison between experimental conditions inappropriate. Comparing the increase over baseline would also have been inappropriate because of a possible regression toward the mean. Therefore, an analysis of covariance was performed, with the baseline test used as the covariate. The assumption of homogeneity of variance was met [$F(2,108) = 0.76$, $p > .05$]. For a 3×2 analysis of covariance, there was a significant effect for type of presentation [$F(2,112) = 22.44$, $p < .001$] and no effect for type of response [$F(2,112) = 2.12$, $p > .05$]. However, there was a significant interaction between type of presentation and type of response [$F(2,112) = 5.71$, $p < .01$], which suggested a need to examine simple effects.

A within-subjects t test is appropriate for comparing inverted and normal response within the same type of presentation. There was no difference for visual presentation [$t(19) = 1.20$, $p > .10$] or auditory presentation

Figure 1. Average digit span for auditory, visual, and mixed presentations in Experiment 1 and the replication of Experiment 1.

Figure 2. Average maximum span for auditory, visual, and mixed presentations in Experiment 1 and the replication of Experiment 1.

[$t(19) = 0.11$, $p > .10$], but the mixed inverted condition was superior to the mixed-normal condition [$t(19) = 3.74$, $p < .002$].

A simple pairwise comparison using an analysis of covariance was used to compare individual conditions across differing types of presentations. The mixed-inverted condition was significantly better than all other conditions: For all four comparisons, $F(1,37) > 15.44$ ($p < .001$).

The other pairwise comparisons were not planned, but they may be of interest. The mixed-normal condition was better (at $p < .05$) than the visual-inverted condition and both auditory conditions, and the visual-normal condition was better (at $p < .05$) than the two auditory conditions. Another method of measuring digit span is the length of the longest string the subject recalled correctly ("maximum span"). This measure, presented in Figure 2, revealed the same pattern of results.

Discussion

The mixed presentation with inverted response was superior to all other conditions. The first question is why inverted response was better than normal response for the mixed presentation. If dual storage was achieved, there are three plausible reasons. The first is modality-specific interference. Speaking a digit, either retrieved from long-term memory or read from a screen, interferes with auditory short-term memory (Dallet, 1964; Jahnke & Nowaczyk, 1977). Therefore, assuming dual storage was achieved, reporting the nonauditorily stored digits first would probably interfere with the auditory store. The question is whether speaking digits in the auditory store creates less interference with the contents of visual short-term memory (or whatever nonauditory memory was being used to store the visually presented digits). Speaking might interfere selectively with the contents of auditory short-term memory, as suggested by Margrain's (1967) results.

Second, if dual storage was achieved, the two short-term stores were storing different quantities of digits. Because the probability of correct recall decreases more rapidly with time when more digits are to be remembered (Melton, 1963), there may be an advantage to reporting the store with the most digits as soon as possible. Assuming dual storage, the auditory store contained the most digits. A third possibility, assuming dual storage, is that interference from emitting digits in one store upon a second store might be less when the second store has fewer digits. Further experimentation could discover to what extent these different factors contributed to the difference between the inverted response and the normal response.

If dual storage was not achieved, an advantage to inverted response for only the mixed presentation is not easily explained. It is plausible, *a priori*, that beginning recall with the middle digits would improve recall: The middle digits might be the digits most vulnerable to being forgotten, so reporting them earlier might increase the probability of overall recall. When subjects hear digits in both ears and attend only to one ear, the unattended digits are less likely to be recalled; however, reporting them first does not improve overall recall (Bryden, 1971). Immediate recall can be improved by reporting digits in inverted order (Howe, 1966; Posner, 1964). However, Posner (1964) found that this increase occurred only for slow presentations (2 sec per digit), not fast presentations (625 msec per digit). The results of Experiment 1 confirm Posner's finding for fast presentation rates. In fact, Experiment 1 offers an explanation of why inverted response might improve immediate recall at slow presentation rates: At slow presentation rates, subjects might be able to recode the first few digits visually, thus achieving dual storage.

The second question is why a mixed presentation leads to better recall. One method of increasing recall is to parse the digits (e.g., Ryan, 1969), which the mixed presentation does. However, the all-visual and all-auditory presentations were also parsed. Rhythmic parsing was chosen for the all-auditory presentation because it was the most effective method Ryan examined. Changing the modality of presentation may be a more effective method of parsing, but it did not lead to much of an improvement in recall in the mixed-normal condition (especially in comparison with the visual-normal condition), suggesting that mixed presentation does not have a large advantage in parsing.

Immediate recall can also be improved by chunking the stimuli (e.g., Bower & Springston, 1970). A simultaneous presentation plausibly facilitates chunking, and the visual portion of the mixed presentation was presented simultaneously. However, to the extent that a simultaneous presentation facilitates chunking, it would equally have facilitated chunking in the mixed-normal condition, and even more so facilitated chunking in the all-visual conditions. Therefore, chunking cannot explain why the mixed presentation produced the largest increase in digit span.

It is possible, *a priori*, that beginning report with the fifth digit is easier for the mixed presentation than for the

all-visual or the all-auditory presentation. However, errors on the fifth digit were infrequent and balanced across conditions. One subject in the auditory presentation had difficulty recalling the fifth digit. (Ignoring this subject raises the average for the auditory-inverted condition .1 digit.) Otherwise, the fifth digit was missed three times in the mixed-inverted condition, twice in the visual-inverted condition, and three times in the auditory-inverted condition.

To explain the interaction found in this experiment without invoking dual storage, some combination of two or more factors must be posited, such as (1) starting with the fifth digit is troublesome in the all-auditory and all-visual conditions, and (2) there is an advantage to reporting the middle digits first. (However, these two factors have difficulties, as discussed above.) The plausibility of separate auditory and visual storage rests on previous demonstrations of visual memory (e.g., Cermak, 1971) and demonstrations that the visual and auditory systems can function independently in processing and retaining information (e.g., Rollins & Hendricks, 1980).

The improvement over baseline was substantially greater in the mixed-inverted condition than in the visual-normal condition (3.06 vs. 1.63). The absolute scores, however, were not very different (9.13 vs. 8.55), although this difference was statistically significant [$t(38) = 2.31$, $p < .05$]. To confirm the use of comparing scores with the baseline score, Experiment 1 was replicated with a comparison of the mixed-inverted condition with the visual-normal condition. Ten subjects were tested, first with a normal auditory presentation and then with a mixed presentation with inverted response and an all-visual simultaneous presentation with normal response. These latter two conditions were counterbalanced for order. As in Experiment 1, the subjects were asked to form a picture of the visual portion of the mixed presentation. When the all-visual presentation was tested after the mixed presentation, the subjects were told they could remember the all-visual presentation however they wanted. When the all-visual presentation was tested before the mixed presentation, no directions for performing the task were given.

The results of this replication are presented as a part of Figure 1. The mixed-inverted condition increased digit span 3.23 digits to 9.50, and the visual-normal condition increased digit span 1.48 digits to 7.75. This replicated the improvement over baseline found in Experiment 1. The difference between these two conditions was significant [$t(9) = 5.15$, $p < .001$].

In this replication, the subjects were asked after testing to recall the visual portions of the mixed presentations. The subjects were encouraged to guess. Only four visual portions were correctly recalled after testing, including two that were easily recognized as years (1962 and 1975). Therefore, there was little observable long-term memory for the visual portion. Furthermore, the auditory portion was correctly recalled on only one of the four trials on which these sequences appeared, so the ob-

servable long-term memory contributed little, if anything, to performance.

The subjects in this replication were also asked if they had subvocalized the visually presented digits during presentation. All subjects reported subvocalizing almost all of the visually presented digits, for both the mixed presentation and the visual presentation. Therefore, a comparison between subjects who vocalized and those who did not was not possible.

EXPERIMENT 2

In Experiment 2, an attempt was made to produce an advantage for normal order of response, by presenting the to-be-remembered-auditorily digits first.

The computer used in Experiment 1 and its replication had broken, so a computer-controlled auditory presentation was not possible. Therefore, all digits were presented visually, and the subjects read aloud the to-be-remembered-auditorily digits. The digits were presented simultaneously in two rows, with a variable number of digits in the top row and four digits in the bottom row. In the experimental conditions, the subjects read the top row aloud and were asked to form a picture of the bottom row. In the control conditions, the subjects read both rows aloud. Therefore, in the control condition, both an auditory and a visual storage of the digits was possible, but this storage was redundant; that is, the same digits were stored in both stores.

Order of response was again varied: The subjects sometimes reported the top row first (normal response) and sometimes reported the bottom row first (inverted response).

To establish generalization, a different method of testing and measuring digit span was used. The longest string that a subject could recall (maximum span) was measured. The subjects were given three chances to recall a given string length. When a trial was correctly recalled, the length of the string was increased.

Method

Subjects. Thirty-one subjects from introductory psychology classes at the University of Washington volunteered in order to earn extra credit. One subject could not read the digits aloud fast enough to participate in the experiment, so his data were discarded.

Stimuli. The digits were presented simultaneously on a green NEC monitor for 0.33 sec times the number of digits in the string. The digits were presented in the middle of the screen in two rows; four digits were in the second row, and the remaining digits were in the top row.

The digits were randomly generated, with the constraint that a digit could not recur without four other digits intervening.

Design. Two variables were manipulated: (1) The subjects either read the top row aloud or read both rows aloud, and (2) recall either was in the order the digits were presented on the screen (normal response) or began with the second row and ended with the first row (inverted response). This created four conditions. Each subject was tested for only three conditions, reading the top row aloud with both orders of response and reading both rows aloud with one order of response. The order and presentation of conditions were counterbalanced. The subjects were tested in three conditions and then retested on the same conditions in the same order. The scores were then averaged.

Measuring maximum digit span. Testing began with six digits in a string. When a subject correctly recalled a string, the number of digits

in a string was increased by one. When the subject could not correctly recall any of three trials at a given length of string, testing ended for that condition. The dependent measure was the longest string the subject correctly recalled.

Procedure. The subjects were tested individually. They were informed of the procedure before the start of the experiment, and each condition was explained again before it began. When the second row was not to be read aloud, the subjects were told not to say it to themselves, but rather to form a picture of it.

Before each presentation, the word "ready" appeared in the top left corner of the screen and then disappeared. Then the digits were presented, and the subject read the appropriate digits aloud and tried to recall aloud the digits immediately after the presentation had ended. The experimenter then signaled the computer whether or not the trial was correct, and a new trial began.

Results

The results are presented in Figure 3. The average maximum digit spans when subjects read all digits aloud during presentation were 9.2 for normal response and 9.3 for inverted response. Because this difference was not significant, these two conditions were combined. The average maximum digit span for reading the top row aloud with normal response was 9.8 digits, and that for reading the top row aloud with inverted response was 9.5 digits. An overall analysis of variance for these three conditions revealed a difference between the three [$F(2,58) = 5.11, p < .01$]. One pairwise comparison was statistically significant: Reading the top row aloud with normal response was better than reading all the digits aloud [$t(29) = 2.97, p < .01$].

Reading the top row aloud with normal response was marginally better than reading the top row aloud with inverted response [$t(29) = 1.88, p < .10$]. A more refined measure of performance was made post hoc: One-third was subtracted from the maximum digit span when the maximum recall occurred on the third trial of that length, and $\frac{1}{3}$ was added to this score for every trial at the level above maximum span on which the subject made only one mistake in recall. This measure tended to have higher test-retest reliability and between-subject correlations, suggest-

Figure 3. Average digit span for reading all the digits aloud or reading just the top row of digits aloud in Experiment 2.

ing that it was more reliable. With this adjusted measure, the difference between normal response and inverted response, .4 digits, was significant [$t(29) = 2.45$, $p < .05$].

Discussion

The improvement in immediate recall was not as large in Experiment 1. There are several possible reasons for this. First, reading aloud a parsed simultaneous presentation seemed to be a very effective method of remembering digits. Second, when only the top row was read aloud, many subjects reported that they read the second row to themselves (despite the instructions not to), which may have interfered with the auditory memory of the first row. A third problem was that the to-be-remembered-visually digits were not recalled as effectively as in Experiment 1. Among other differences, in Experiment 1 the to-be-remembered-visually digits were presented alone; in Experiment 2, the to-be-remembered-visually digits were accompanied by other visually-presented digits, which may have created interference.

If digits were stored only auditorily, it is difficult to explain why *not* reading the bottom row aloud could improve recall. Comparisons between subvocalization and reading aloud have found an advantage only for the last few digits read aloud (e.g., Conrad & Hull, 1968). Reading aloud is thought to create a short-term auditory store, often called echoic memory, with a capacity of one or two digits (e.g., Crowder, 1976). When only the top row is read, the advantage of having this short-term store is probably transferred from the last few digits of the bottom row to the last few digits of the top row. If the last few digits of the top row were the most vulnerable to forgetting, their "fortification" might be more advantageous. However, for 9- and 10-digit strings with all digits read aloud, the next to last digit of the bottom row was the digit least likely to be recalled, even though it presumably had the advantage of the echoic store. This suggests that the top row did not need fortifying more than the bottom row. Furthermore, fortification of the top row would be at least as valuable for the inverted response as for the normal response, yet an advantage to reading only the top row aloud occurred only for normal response. Therefore, a transfer of echoic store does not seem to explain why reading only the top row improved recall.

If a single short-term store that can store auditory and visual representations is assumed, parsing may be able to account for the difference between reading both rows aloud and reading the top row aloud with normal response. A single store with different modalities of representation does not as easily account for the difference between normal and inverted response when only the top row was read aloud. Although the difference between reading the top row with normal response and reading the top row with inverted response was not large, it was marginally significant and in the predicted direction; furthermore, only the condition of reading the top row with normal response

was statistically different from the control condition. Normal and inverted response were equivalent when both rows were read. Assuming dual storage, normal response had the same advantages that inverted response had in Experiment 1.

GENERAL DISCUSSION

These experiments suggest that the dual storage of digits can improve immediate recall. Experiment 1 suggests that under the proper conditions, this improvement can be substantial: 1.5 digits compared with a parsed control, and 3 digits compared with an unparsed control.

Order of report was not relevant for any control conditions, but was relevant for all presentations designed to elicit dual storage. Three possible reasons for this all involve dual storage: (1) There is modality-specific interference in responding, such that speaking interferes with auditory short-term memory, but not visual short-term memory; (2) when digits are stored in two different stores, the store with the fewer number of digits should be reported first; and (3) a store with fewer number of digits might be less disturbed by output from the other store.

Identity of the Stores

The nonredundant use of two stores to increase capacity would be of interest no matter which stores were used; nonetheless, knowing the identities of these stores would be useful. Use of an auditory store is well accepted. A visual short-term store seems to be the most plausible candidate for the second store. A visual store exists (e.g., Cermak, 1971), it can store digits (Frick, 1984; Laughery, Welte, & Spector, 1973), and it can process digits independently and without interference from the auditory modality (e.g., Rollins & Hendricks, 1980). Furthermore, the conditions that apparently elicited dual storage had a visual presentation and instructions to form visual images, which would seem to facilitate a visual store.

Other possible stores do not seem to possess the appropriate characteristics. An attention-demanding central processor, such as Baddeley's (1981) central executive, is a possible store for digits. However, if the subjects used this store for the visual portion of the mixed presentation in Experiment 1, then interference should have occurred while they listened to the auditory digits, when they switched from listening to responding, and when they repeated the auditory digits aloud. It seems more plausible that attention and initiating action require central processing and that the storage of digits occurs in a more passive store.

If long-term memory is used for delayed recall, it can perhaps be used in immediate recall. However, in the replication of Experiment 1, the subjects were usually unable to remember any of the visual portions from the mixed presentations, even though in some cases they were being asked only a few minutes after testing. This suggests that the store used to remember the visual portion is useful for immediate recall but does not persist. A store

with these properties fits the definition of short-term memory, not long-term memory.

The existence of a modality-free semantic store for digits is possible (Shulman, 1971). Hinrichs and Novick (1982) reported that the serial position curve for magnitude encoding of digit strings favors the earlier digits over the later. In Experiment 1, the serial position curve for the visual portion of mixed-inverted condition was bowed, indicating what Hinrichs and Novick called nominal encoding. In the replication of Experiment 1, the eight subjects who demonstrated dual storage made few errors; the serial position curve could be described as either flat or bowed, but not as decreasing. In Experiment 2, the serial position curve for the unspoken row in the normal response condition was bowed. However, the largest number of errors was made at the third serial position, and there were more errors in the fourth position than in the second, suggesting perhaps some magnitude encoding. The serial position curve for subjects who did the best on this condition was the same as that for subjects who did not.

This discussion has assumed that two tasks can be performed independently in a visual and an auditory modality. A spatial/verbal classification has been proposed for explaining modality-specific interference (e.g., Brooks, 1970). That classification does not account for the results of Experiments 1 and 2, because the stimuli were all verbal. A spatial/auditory classification also has been proposed (Baddeley & Lieberman, 1980). If the visual modality is defined as processing patterns in space and the spatial modality is defined as the location of these patterns, then the second store had visual and spatial properties—visual storage of the digits and spatial storage of their order.

Total Capacity

Miller (1956) noted that, for a variety of stimuli, memory span fell into the range 7 ± 2 . In the mixed-inverted conditions of Experiment 1 and its replication, however, 73% of the subjects recalled at least one string of 10 or more digits; in the condition of reading only the top row aloud and normal response in Experiment 2, 80% recalled at least one string of 10 or more digits. When dual storage can be achieved, subjects' performance seems to fall outside the range Miller suggested. The capacity of 7 ± 2 apparently applies to individual short-term stores such as auditory or visual short-term memory, not to the combined capacity of these stores.

These experiments have shown that capacity can be increased by dual storage. If auditory and visual short-term memory were completely independent, the capacity of an individual would be equal to the sum of the capacity of these two memories. It seems worthwhile to analyze whether these experiments are consistent with such an additivity.

If capacities are in practice additive, recall would have increased 4 digits. In Experiment 1, the increase over baseline was only 1.5 digits more than the parsed control conditions. However, for the purpose of measuring an in-

crease produced by using two stores, this comparison is inappropriate. If dual storage was achieved with the mixed presentation, the auditory portion was not parsed. An increase of 4 digits would be expected only in comparison with an unparsed auditory presentation.

The baseline test in Experiment 1 had an unparsed auditory presentation, so it provides an appropriate comparison, except that it was always tested first. The effects of practice cannot be determined directly, but other experiments using this design suggest little or no practice effect. (The effects of order in Experiment 2 were not statistically significant. In two pilot studies that counterbalanced for order of presentation, the disadvantage of testing first was $+.3$ digits and $-.2$ digits, both nonsignificant, and in other unpublished experiments using this method of testing digit span, testing first had an average disadvantage of $.1$ digit.) Therefore, an average improvement of around 3 digits in the mixed-inverted condition, compared with an unparsed auditory presentation, is a reasonable (although tentative) conclusion.

An increase of three digits over baseline still creates a discrepancy of one digit. One source of this discrepancy is subjects who cannot use or choose not to use dual storage. Also, an improvement of four digits would be expected only if the visual portion was always recalled correctly. Therefore, a second source of discrepancy is decay of the visually presented digits between presentation and recall.

The replication of Experiment 1 was clear in showing both of these effects. Two subjects increased their recall only 1 digit over baseline in the mixed-inverted condition, suggesting that for them either dual storage was not achieved or it was not effective. The remaining eight subjects increased their recall 3.33 to 4.33 digits, with an average increase of 3.79 digits. Six times these eight subjects correctly recalled the auditory portion but did not recall the visual portion. If these six visual portions had been correctly recalled, digit span for the mixed presentation would have been 4.04 digits more than digit span for the baseline auditory test; furthermore, the correlation between subjects' scores for these two measures would have been .91.

Therefore, the results of these experiments are consistent with the hypothesis that the capacities of auditory short-term memory and visual short-term memory are, in theory, additive. Given problems with decay, translating this theoretical additivity into recall equal to the capacity of auditory short-term memory plus the capacity of visual short-term memory may be impossible. Partial report of a concurrent visual and auditory presentation might demonstrate a larger potential capacity.

Reisberg, Rappaport, and O'Shaughnessy (1984) attempted to increase digit span by use of an auditory store and a store based upon finger movements. The effects of parsing were sometimes not controlled for, subject demand was not controlled for, and the possible use of a visual store was not eliminated (the digits were presented simultaneously). However, the increase in digit span was

large compared with an unparsed auditory control (3.2 digits after 5 days of practice in Experiment 2), and subjects' use of the finger movements was established. By the same logic as described above for a visual store, their experiment showed the existence of a third system, a motor memory. (They interpreted their results differently.) Interestingly, subjects in their experiment spontaneously adopted the strategy of reporting the auditorily-stored digits first.

Conclusion

The number of digits that can be recalled immediately after presentation can be increased by nonredundant storage in two different stores. A visual store is the most plausible candidate for the second store used in these experiments. The results do not demonstrate that the capacities of these two stores are additive, but they are consistent with this hypothesis, and therefore are consistent with the hypothesis of independent auditory and visual short-term stores.

REFERENCES

- BADDELEY, A. D. (1981). The concept of working memory: A view of its current state and probable future development. *Cognition*, **10**, 17-23.
- BADDELEY, A. D., & LIEBERMAN, K. (1980). Spatial working memory. In R. S. Nickerson (Ed.), *Attention and performance VIII* (pp. 521-539). Hillsdale NJ: Erlbaum.
- BADDELEY, A. D., THOMPSON, N., & BUCHANAN, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, **14**, 575-589.
- BOWER, G. H., & SPRINGSTON, F. (1970). Pauses as recoding points in letter series. *Journal of Experimental Psychology*, **83**, 421-430.
- BROOKS, L. R. (1968). Spatial and verbal components of the act of recall. *Canadian Journal of Psychology*, **22**, 349-368.
- BROOKS, L. R. (1970). An extension of the conflict between visualization and reading. *Journal of Experimental Psychology*, **22**, 91-96.
- BRYDEN, M. P. (1971). Attentional strategies and short-term memory in dichotic listening. *Cognitive Psychology*, **2**, 99-116.
- CERMAK, G. W. (1971). Short-term recognition memory for complex free-form figures. *Psychonomic Science*, **25**, 209-211.
- CONRAD, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology*, **55**, 75-84.
- CONRAD, R. (1972). Speech and reading. In J. J. Kavanagh & I. G. Mattingly (Eds.), *Language by ear and by eye: The relationship between speech and reading* (pp. 205-240). Cambridge, MA: MIT Press.
- CONRAD, R., & HULL, A. J. (1968). Input modality and the serial position curve in short-term memory. *Psychonomic Science*, **10**, 135-136.
- CROWDER, R. G. (1976). *Principles of learning and memory*. Hillsdale, NJ: Erlbaum.
- DALLET, K. M. (1964). Effects of a redundant prefix on immediate recall. *Journal of Experimental Psychology*, **67**, 296-298.
- DORNBUSH, R. L. (1968). Shadowing in bisensory memory. *Quarterly Journal of Experimental Psychology*, **20**, 225-231.
- FRICK, R. W. (1984). *Testing visual short-term memory: Effects of sequential and simultaneous presentations*. Manuscript submitted for publication, University of Washington, Seattle, WA.
- HINRICH, J. V., & NOVICK, L. R. (1982). Memory for numbers: Nominal vs. magnitude information. *Memory & Cognition*, **10**, 479-486.
- HOCKEY, R. (1973). Rate of presentation in running memory and direct manipulation of input-processing strategies. *Quarterly Journal of Experimental Psychology*, **25**, 104-111.
- HOWE, M. J. A. (1966). A note on order of recall in short-term memory. *British Journal of Psychology*, **57**, 435-436.
- JAHNKE, J. C., & NOWACZYK, R. H. (1977). Processing demand and short-term memory: The response-prefix effect. *Journal of Experimental Psychology: Human Learning and Memory*, **3**, 325-336.
- LAUGHERY, K. R., WELTE, J. W., & SPECTOR, A. (1973). Acoustic and visual coding in primary and secondary memory. *Journal of Experimental Psychology*, **99**, 323-329.
- MARGRAIN, S. A. (1967). Short-term memory as a function of input modality. *Quarterly Journal of Experimental Psychology*, **19**, 109-114.
- MELTON, A. W. (1963). Implications of short-term memory for a general theory of memory. *Journal of Verbal Learning and Verbal Behavior*, **2**, 1-21.
- MILLER, G. M. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Bulletin*, **63**, 81-97.
- MURDOCK, B. B., & CAREY, S. T. (1972). Release from interference in single-trial free recall. *Journal of Verbal Learning and Verbal Behavior*, **11**, 398-402.
- MURDOCK, B. B., & WALKER, K. D. (1969). Modality effects in free recall. *Journal of Verbal Learning and Verbal Behavior*, **8**, 665-676.
- MURRAY, D. J. (1967). The role of speech responses in short-term memory. *Canadian Journal of Psychology*, **21**, 263-276.
- PENNEY, C. G. (1975). Modality effects in short-term verbal memory. *Psychological Bulletin*, **82**, 68-84.
- POSNER, M. I. (1964). Rate of presentation and order of recall in immediate memory. *British Journal of Psychology*, **55**, 303-306.
- PROCTOR, R. W. (1978). Attention and modality-specific interference in visual short-term memory. *Journal of Experimental Psychology: Human Learning and Memory*, **4**, 239-245.
- REISBERG, D., RAPPAPORT, I., & O'SHAUGHNESSY, M. (1984). Limits of working memory: The digit span. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **10**, 203-221.
- ROLLINS, H. A., JR., & HENDRICKS, R. (1980). Processing of words presented simultaneously to eye and ear. *Journal of Experimental Psychology: Human Perception and Performance*, **6**, 99-109.
- RYAN, J. (1969). Grouping and short-term memory: Different means and patterns of grouping. *Quarterly Journal of Experimental Psychology*, **21**, 137-147.
- SEGAL, S. J., & FUSELLA, V. (1970). Influence of imaged pictures and sounds on detection of visual and auditory signals. *Journal of Experimental Psychology*, **83**, 458-464.
- SHULMAN, H. G. (1971). Similarity effects in short-term memory. *Psychological Bulletin*, **75**, 399-415.
- WICKELGREN, W. A. (1969). Auditory or articulatory coding in verbal short-term memory. *Psychological Review*, **76**, 232-235.
- WOODWORTH, R. S., & SCHLOSBERG, H. (1954). *Experimental Psychology*. London: Methuen.

(Manuscript received September 21, 1983;
revision accepted for publication June 20, 1984.)