Western Carolina Center
Papers and Reports

Span ability, word frequency, and free learning among retarded adolescents

Bruce L. Bachelder

Western Carolina Center
Volume IV, No. 16
1974

Abstract

Span ability, word frequency, and free learning among retarded adolescents

Bruce L. Bachelder
Western Carolina Center

Ten high-span children (mean word span = 4.78; mean IQ = 70.1; mean CA = 13.2; mean institutionalization = 16.1 months) and 10 low-span children (mean word span = 2.50; mean IQ = 49.3; mean CA = 15.1; mean institutionalization = 61.0 months) attempted 16 trials in a free-learning paradigm. Half the subjects attempted low-frequency words and half attempted high-frequency words. High-frequency words produced higher scores than low-frequency words ($F = 5.21$, 1/16 *df*, $p < .05$). High-span subjects produced higher scores than low-span subjects ($F = 7.14$, 1/16 *df*, $p < .025$). Size of Span and Word Frequency did not interact ($F = .001$, 1/16 *df*). There was a significant linear increase in scores over trials ($F = 28.46$, 1/11 *df*, $p << .0005$). Neither Size of Span nor Word Frequency interacted with Trials ($F = 1.12$, 15/240 *df;* and $F = .71$, 15/240 *df*, respectively). There was no Word Frequency X Size of Span X Trials interaction ($F = .59$, 15/240 *df*).

Span ability, word frequency, and free learning[1] among retarded adolescents

Bruce L. Bachelder[2]
Western Carolina Center[3]

In the memory span test the experimenter presents sequences of stimuli and the subject attempts to repeat, report, or label them in order. The memory span is a measure of the largest absolute number of stimuli a subject can reproduce reliably. Memory span subtests have long been used to measure intelligence, mental development, and to diagnose abnormality. The Stanford-Binet test and Wechsler tests all have span subtests. There are span subtests on tests of brain damage (Hunt, 1943), aphasia (Eisenson, 1954), and psycholinguistic abilities (Kirk, McCarthy, and Kirk, 1968). In addition, Wechsler (1958) noted that low span scores are related to psychosis and emotional disturbance. Despite the considerable clinical use of span tests, experimentalists have devoted relatively little systematic effort to the study of individual differences in span ability (Jensen, 1964). The writer knows of only two experiments which have explicitly examined the relation between traditional span measures and the performance of other tasks.

Whimbey, Fischhof, & Silikowitz (1969) tested 24 college students on a mental arithmetic test and found that digit span scores correlated .77 ($p < .0005$) with the mental arithmetic scores. They replicated the study with 40 more college students and found a correlation of .67 ($p < .005$). (The p values were determined by the present investigator.) Neither the digit span test nor the mental arithmetic test correlated with a vocabulary test (rs = .12 and .15, respectively). Baumeister, Bartlett, & Hawkins (1963) correlated WISC Digit Span Test scores with trials to criterion in a double alternation lever pressing learning task. Among two groups of retarded subjects (mean IQ = 56) the correlations were -.61 and -.35 (significant at the .01 and .05 levels, respectively). The correlations were negative because high span scores were associated with low trials to criterion. Among normal subjects (mean IQ = 99) the correlation was nonsignificant (r = -.12).

Several other experiments have ostensibly investigated the relation between "span ability" and performance in other tasks but in these cases span ability was not measured by traditional span measures but rather by performance in the free learning task.

According to these experiments, free learning ability is directly related to serial recall (Furukawa, Suydam, & Miller, 1969), acquisition and recall of foreign words under programmed instruction (Furukawa, 1970), the number of word sorting categories chosen in a free choice card sorting task (Miller, Fleishman, & Simpson, 1971), and paired-associates learning (Miller, 1971; Miller & Weinstock, 1971). The present investigator knows of no published experiment which has demonstrated a relation between free recall performance and span performance.

---

[1] This electronic manuscript was prepared in 2008. It was scanned from a copy of the original dated 1974. Minor changes were made to correct errors and update the style.

[2] I wish to thank Mr. William F. Cobb, III, who conducted this experiment for undergraduate credit from Davidson College, Davidson, North Carolina. This research was supported by the Department of Psychology and Education at Western Carolina Center in Morganton, North Carolina 28655.

[3] Independent Practice of Psychology since 1980; retired in 2004. Contact: 306 West Union St, Morganton, NC 28655; Bruce@BruceBachelderPhD.com; wwwBruceBachelderPhD.com

The present experiment attempted to relate span ability to performance in a free learning experiment. If span ability is related to performance in verbal learning tasks, then the great volume of data on short tern memory and verbal learning might begin to be integrated via the concept of span ability. This would be an important theoretical advance because span ability is clearly related to both intelligence and developmental level. Furthermore, rather little work has been devoted to analysis of individual differences in verbal learning beyond gross comparisons of age groups or intelligence groups.

Span ability as an individual differences construct has certain desirable characteristics. Span ability is clearly related to intelligence and developmental level. Span ability is general and abstract in the sense that, over a fairly broad range, the particular stimuli and responses used to measure span have relatively small effects on span measures, compared to the range of individual differences in span ability (Brener, 1940). Span ability can be measured quickly and reliably by unsophisticated technicians (Bachelder, 1970; Jensen, 1970). Span ability is markedly resistant to improvement with practice. Where practice effects have been demonstrated, they have been at great effort, were small compared to the practice involved, and were transient (Gates & Taylor, 1925; Martin & Fernberger, 1929). No study to the writer's knowledge has demonstrated that practice effects, when they do occur, generalize to span tests for other materials. In other words, it is quite possible that those practice effects which do occur are stimulus-response specific and cannot be considered the result of change in a more fundamental span ability.

One night expect a relation between span ability and free learning merely because of the structural similarity of the two paradigms. In each, the subject attends to a sequence of stimuli and then attempts to reproduce them with either vocal or written responses. The chief differences between the two paradigms is that in the free learning task stimulus strings are supraspan, order is unimportant in scoring, and stimulus presentation rates are usually slow (about 2 seconds or more per stimulus).

In the present experiment, high-span subjects and low-span subjects attempted 16 trials in a standard free learning paradigm. By giving many trials rate of improvement could be studied as a function of span ability. Half the subjects learned high frequency words and half learned low frequency words. Since the performance of normal subjects is directly related to word frequency (Hall, 1971), the inclusion of this variable should help integrate the present results with the data of normal subjects.

Method

Subjects

Ten high-span (HS) subjects and 10 low-span (LS) subjects were selected from the population at a residential center for the mentally retarded. The 10 HS subjects had a mean word span of 4.78 (mean IQ = 70.1; mean CA = 13.2; mean institutionalization = 16.1 months) and the 10 LS subjects had a mean word span of 2.50 (mean IQ = 49.3; mean CA = 15.1; mean institutionalization = 61.0 months). The word spans were measured by the staircase technique (a threshold-like measurement; Bachelder, 1970), and the words were common words such as grass, phone, and door. In the staircase span test whenever the subject produces a string perfectly he attempts a string one digit or word larger on the subsequent trial. Similarly, whenever he errs he attempts a stimulus string one word or digit smaller on the subsequent trial. The subject attempts 9 trials and his score is the average of the string sizes presented to him.

Materials and Equipment

Two word pools were randomly selected from the Thorndike-Lorge (1944) word list. The 10 high-frequency (HF) words came from the 500 most frequent words; the 10 low-frequency (LF) words from the words occurring once in 4 million words. Because of the age of the children certain words were eliminated from the LF list, namely, douche, groin, pimp, and pus. The HF words were: mind, school, word, life, bank, face, dress, car, air, and sun. The LF words were; hoax, slob, jinx, norm, bash, bike, fuzz, grid, skit, and squid. Other word attributes such as imagery value and meaningfulness were not controlled. Not only were the relevant norms not available, but Pavio, Yuille, & Madigan. (1968) reported that the correlations between Thorndike-Lorge frequency and imagery and meaningfulness are low (.23 and .33, respectively).

A tape recorder was used to present the words. Four randomizations of each word pool were recorded, with 2 seconds between each word and 60 seconds between each randomization. The word “ready" was recorded five seconds before each list.

Procedure

The experimenter brought the subject to the experimental room and read these instructions:

> I am interested in how people learn words and I want to pay you tokens to learn some words. What you do will be very easy and probably fun and you can earn fifty tokens. I would like you to stay and help me but you may go if you want to. Will you stay and help?
>
> I am going to play some words on this tape recorder and I would like you to try your best to learn all of them. I will give you plenty of practice so don’t worry about not getting them all at first. Try to do your very best; when you finish I will give you 50 tokens. (Tokens were used to buy recreation, spending money, edibles, and privileges in an on-going therapeutic token economy.)

Each subject served in just one condition for 16 trials. After each 4 trials the experimenter rewound the tape and presented the four randomizations again. The subjects responded in the 60-second interval between successive presentations of the stimulus words.

Results

The mean scores are plotted in Figure 1. A 2 X 2 X 16 (Size of Span X Word Frequency X Trials) analysis of variance was performed on the number of words correctly recalled on each trial. A summary of the analysis is presented in Table 1.

The HF words produced higher scores than the LF words (mean words recalled across 16 trials = 6.1 and 4.6). The HS subjects had higher scores than the LS subjects (mean words recalled across 16 trials = 6.2 and 4.3). The mean words recalled on Trial 1 were 3.4 and 2.2

for the high- and low-span subjects, respectively. The linear, quadratic, and cubic trend components of variance for Trials were all significant.

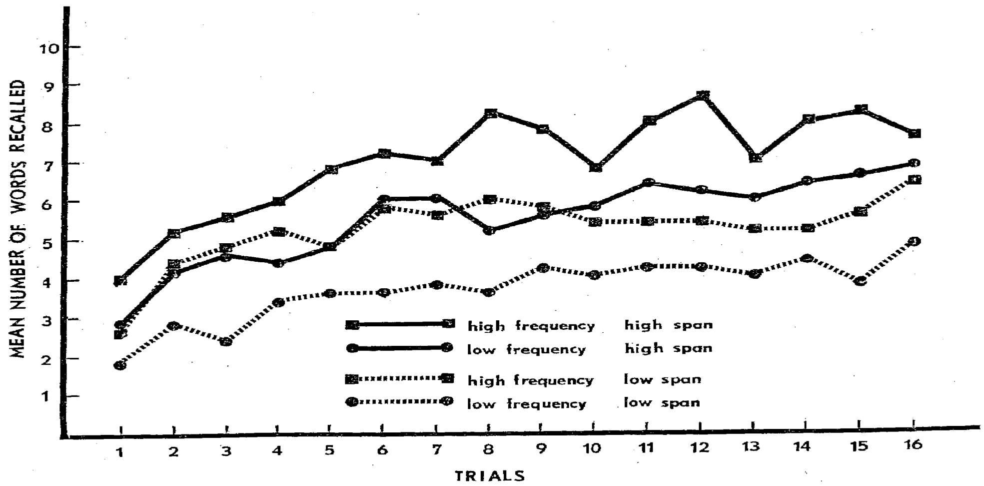

Figure 1. The mean number of correct responses as a function of Trials, Span Ability, and Word Frequency.

None of the interactions approached significance. The correlation between word span and the number of words recalled on Trial 1 was .48 ($p < .01$). The correlation between word span and the total correct words across all 16 trials was .64 ($p < .005$). These correlations underestimate the true correlations between span ability and free learning because of the inclusion of variance due to word frequency.

Table 1
A Summary of the Analysis of Variance

Source	*df*	MS	*F*	*p*
Between Subjects	**19**			
Word Frequency (F)	1	189.112	5.21	<.05
Size of Span (S)	1	259.200	7.14	<.025
F X S	1	.050	.001	
Subjects within groups	16	36.302		
Within Subjects	**300**			
Trials (T) (overall)	15	17.399	14.30	<<.0005
T (linear)	1	188.777	28.46	<<.0005
T (quadratic)	1	42.444	33.74	<<.0005
T (cubic)	1	17.877	10.61	≈ .005
T (quartic)	1	.00041	.001	
T (residual, 5th through 15th degree)	11	1.081	1.44	≈ .25
F X T	15	.859	.71	
S X T	15	1.360	1.12	
F X S X T	15	.717	.59	
T X Subjects within groups	240	1.217		
(linear)	16	6.633		
(quadratic)	16	1.258		
(cubic)	16	1.685		
(quartic)	16	.397		
(5th through 15th degree)	176	.752		
Total	319			

Discussion

The present results indicate that word frequency is an important variable controlling the free learning performance of retarded subjects. Since similar results have been found for normal subjects (Hall, 1954), the present results help integrate the verbal learning data of retarded subjects with the verbal learning data of normal subjects. If we compare Hall's results with the present results, it would appear that word frequency is a more potent variable for normal subjects. Hall found that his two extreme frequency conditions produced mean words recalled of 12.04 and 15.04, a difference of 3 words. In the present experiment the average words recalled across 16 trials were 4.6 and 6.1, a difference of 1.5 words. In the present experiment, the average words recalled on the first trial were 2.2 and 3.4, a difference of 1.2 words. More importantly, the range of word frequencies used in the present experiment was much greater than that used by Hall, namely, words occurring 50 to 100 times per million words versus words occurring 1 time per million words. The conclusion that word frequency is a more potent variable for normal subjects must be cautious, however, because Hall's experiment differed from this experiment in both rate of presentation and the number of stimulus words.

It was concluded that for retarded subjects span ability is directly related to free recall performance, but span ability is not related to improvement across trials. The latter conclusion must be somewhat cautious because two subjects reached (but did not maintain) ceiling performance by trials five and eight. The subjects were in the HS-HF group and had they had the opportunity they might have gained at a higher rate than the LS subjects. Note, however, the fairly obvious flattening of the curve for the HS-HF group. It is doubtful that the two high performers would have fully offset this flattening effect to the extent of producing a significant interaction between Span Ability and Trials.

There is additional evidence to suggest that span ability is related to free learning performance. Since normal subjects have higher span abilities than retarded subjects, we would predict from the present data that normal subjects would perform at a higher level on free learning. Murdock (1962) presented 10 words to college students at a 2-per-second rate just as in the present experiment and found a mean recall of 6.39 words on the first trial. This level of performance is considerably higher than the present first trial performance of 3.4 and 2.2 for the high- and low-span groups, respectively.

The present data may help us understand the paradoxical relation between IQ and learning. Hovland (1951) concluded that IQ is generally related to gross measures of learning in a variety of tasks but that IQ correlates little, if at all, with gain scores. It seems obvious that intelligence is intimately related to the ability to learn, yet empirically IQ is not related to gain scores which have the highest face validity as measures of learning (changes in behavior as a function of practice).

This paradox may be resolved as follows. Span ability and intelligence are not related to learning in the sense of changes of behavior with practice. Rather, span ability and intelligence are related to the ability to perform a given task. Thus in the present experiment the span groups differed from the first trial in their ability to perform the free learning task, but each group improved at similar rates. Under certain circumstances learning measures confound gain scores with individual differences in ability to perform a task. For example, if in the present experiment a criterion measure had been used, the high-span subjects would have appeared to learn faster. This would not have been because they gained faster, but rather because they started higher and would have reached criterion sooner. Under these conditions span ability or IQ would be correlated with trials to criterion even though both ability groups would have gained at similar rates across trials.

It is uncertain what the ability is that is involved in both the span test and the free recall test (and presumably the IQ test). Regardless, the present experiment allows us to make the tentative theoretical statement that free learning and span tests both involve a common ability. Having made that theoretical statement, we can draw upon span data and predict other results. For example, free recall performance should increase until 14 to 20 years of age, then continue to increase slightly to age 25, then begin a general decline (based on span norms by Wechsler, 1958). Similarly, free recall should be directly related to intelligence such that retarded subjects perform poorer than normal subjects and who should perform more poorly than superior subjects. Wachs (1969) has data which are partially supportive of these predictions. He found a significant age effect for free learning among 4th, 8th, and 12th graders. He also found that free learning was related to intelligence among groups of children with IQs greater than or equal to 115, between 90 and 110, and less than 90.

References

Bachelder, B. L. (1970/1971). The memory span paradigm: Its use for the analysis of mental retardation. (Doctoral dissertation, Michigan State University) Ann Arbor, Michigan: University Microfilms. No. 71-13, 158.

Baumeister, A. A., Bartlett, C. J., & Hawkins, W. F. (1963). Stimulus trace as a predictor of performance. *American Journal of Mental Deficiency, 67,* 726-729.

Brener, R. (1940). An experimental investigation of memory span. *Journal of Experimental Psychology, 26,* 467-482.

Eisenson, J. (1954). *Examining for aphasia.* New York: The Psychological Corp.

Furukawa, J. M. (1970). Chunking method of determining size of step in programmed instruction. *Journal of Educational Psychology, 61,* 247-254.

Furukawa, J. M., Suydam, M. M., & Miller, S. (1969). The effects of grouping instructions in short-term memory. *Psychonomic Science, 15.* 81-82.

Gates, A. I. & Taylor, G. A. (1925). An experimental study of the nature of improvement resulting from practice in a mental function. *Journal of Educational Psychology, 16,* 583-592.

Hall, J. F. (1954). Learning as a function of word frequency. *American Journal of Psychology, 67,* 138-140.

Hall, J. F. (1971). *Verbal learning and retention.* Philadelphia: Lippincott.

Hovland, C. I. (1951). Human learning and retention. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley.

Hunt, H. F. (1943). *The Hunt-Minnesota test for organic brain damage.* Los Angeles: Western Psychological.

Jensen, A. R. (1964). *Individual differences in learning: Interference factor.* (USDHEW Office of Education Cooperative Research Project No. 1867) Washington: Educational Resources Information Center.

Jensen, A. R. (1970). A theory of primary and secondary familial mental retardation. In N. R. Ellis (Ed.), *International review of research in mental retardation,* Vol. 4. New York: Academic Press.

Kirk, S. A., McCarthy, J. J., & Kirk, W. D. (1968). *Examiner's manual: Illinois test of psycholinguistic abilities.* Urbana: Board of Trustees of the University of Illinois.

Martin, P. R. & Fernberger, S. W. (1929). Improvement in memory span. *American Journal of Psychology*, *41,* 91-94.

Miller, S. (1971). Memory span effects in transfer. *The Proceedings of the 79th Annual Convention of the American Psychological Association, 39.*

Miller, S., Fleishman, P., & Simpson, R. (1971). Individual differences in recall. *Psychological Reports, 29,* 499-501.

Miller, S. M. & Weinstock, R. B. (1971). Memory span and intralist response similarity in paired-associate learning. *Psychonomic Science*, *23*, 139-140.

Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, *64*, 422-488.

Paivio, A., Yuille, J. C., & Madigan, S. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology, 76*(1, Pt. 2).

Thorndike, E. L. & Lorge, I. (1944). *The teachers wordbook of 30,000 words.* New York: Teachers College Bureau of Publications.

Wachs, T. D. (1969). Free-recall learning in children as a function of chronological age, intelligence, and motivational orientation. *Child Development, 40,* 577-589.

Wechsler, D. (1958). *The measurement and appraisal of adult intelligence* (4th ed.). Baltimore: Williams & Wilkins.

Whimbey, A., Fischhof, V., & Silikowitz, R. (1969). Memory span: A forgotten capacity. *Journal of Educational Psychology, 60,* 56-58.