Span Theory: Laboratory and Classroom Applications[1][2]

Bruce L. Bachelder

Western Carolina Center[3]
Morganton, North Carolina 28655

Abstract

This is a three-part paper which illustrates the broad range of basic and applied research stimulated by Span Theory.

Part 1: Performance and learning. Two of Murdock's (1960) equations describing acquisition curves in multitrial free recall in college students contain an empirical quantity, m, which Murdock thought might be the immediate memory span. This experiment tested the implications of the equations that individual differences in memory span are directly related to performance and to rate of learning in this task. Twenty high- and low-span retarded subjects took 16 trials in a multitrial free recall task. As expected, the high-span subjects out performed the lower-span subjects from the earliest trials and the slopes of their acquisition curves were steeper.

Part 2: Span and multidimensional absolute judgment. Miller (1956) noted the similarity in scores for memory span and unidimensional absolute judgment tasks (7 $\pm$ 2), but concluded there is no common underlying ability because information transmission in multidimensional tasks greatly exceeds 7 $\pm$ 2. Span theory, however, can explain high information transmission in the multidimensional tasks while retaining the notion of a limited capacity in the 7 $\pm$ 2 range. Multidimensional absolute judgment data by Pollack & Ficks (1954) were selected for analysis. Their experimental situation was analyzed as a combination of an immediate memory span task and a unidimensional absolute judgment task. A model was developed and used to predict the data using

---

[1] A paper presented at the Eleventh Annual Gatlinburg Conference on Research in Mental Retardation, March 8-10, 1978, in Gatlinburg, TN. Because of time constraints part 2 was not actually presented at the meeting. [E:\NBDATA\LABCLASS\LABCLASS.NB]

[2] Note; June 1997: This paper was originally prepared in 1978 and has been distributed as a mimeographed manuscript. The present manuscript is the same paper scanned for reproduction via computer. The concepts and terminology of span theory have evolved considerably since 1978, but those developments are not included here. More recently prepared manuscripts are available from the author.

[3] The author has been in full-time independent practice of psychology since July 1980: Bachelder & Bachelder, 306 West Union Street, Morganton, NC 28655-3729, 828-437-5090, 828-437-1068, brucebachelder@hci.net.

memory span values which are typical for the subjects in question. The model nicely predicts the data in four variations of the basic task.

Part 3: Classroom applications. Several memory span tests have been developed to test the span abilities of non-verbal severely physically handicapped young adults who have generally been considered to be severely retarded. They respond by pointing or otherwise indicating response sequences on a symbol communication board. Their span scores ranged from about 1.7 to 5.4; 5.4 is about the mean for normal adults. The symbol board span scores of these subjects corresponded closely with objective assessment of their language comprehension, language production, and general classroom performance and learning.

Span Theory: Laboratory and Classroom Applications

Span Theory (Bachelder & Denny, 1977a, 1977b) is a theory of intelligence couched in stimulus and response terms. The central capacity construct is span ability which is measured by a memory span test using familiar words as stimuli and responses. Span ability can be defined as the ability to cope with task complexity and task complexity is defined as the number of jointly (conjunctively) relevant cues in a task. If we know the task complexity and a subject's span ability, we can make fairly precise quantitative predictions of learning and performance. So long as task complexity is equal to or less than span ability the subject is said to have the ability to perform, or learn to perform, the task, provided: (a) prerequisite skills and responses are established, (b) appropriate teaching methods are used, and (c) the complexity of the teaching method does not exceed span ability.

This paper presents three research projects which illustrate applications of span theory and which provide empirical support for the theory. First, evidence is presented in support of the idea that span ability is directly related both to performance and rate of learning in a verbal learning task. Second, the process of task complexity analysis is explained and used to make predictions of performance in a laboratory judgment task. Finally, a preliminary report is made of a project the intent of which is to measure the span abilities of nonverbal, seriously physically handicapped, young adults. Because of their physical disabilities and poor IQ test performance, all these people have been considered to be severely or profoundly retarded, yet in several cases span abilities and classroom performance suggest otherwise.

Span Ability, Learning, and Performance

Murdock (1960) has argued that the multitrial free recall task provides a clear operational distinction between retention and learning. In the multitrial free recall task the subject must learn a list of words. On the first trial he hears or sees all the words and tries to recall all of them. On the second trial he hears or sees the words in another random order and again attempts to recall them all. The subject is scored simply for the number of correct words regardless of their order of production. When the number of correct words is plotted against trials the usual finding is that the acquisition curve is an increasing but negatively accelerated function of the number of trials. According to Murdock, retention is measured by the number of correct responses on Trial 1 and rate of learning is assessed by the slope of the Trial X Trial acquisition curve. This paper departs from Murdock's language slightly by defining performance as the number correct on any trial; but Murdock's definition of learning is used essentially as he presented it. Equation 1 is Murdock's equation describing multitrial free recall in college students.

$$\underline{R} = \underline{c}\,(1-e^{-bn}) \tag{1}$$

R is the number correct on a trial, c is an upper limit or asymptote on R, b is a rate parameter controlling rate of change in R across trials, n is the number of presentations, and e is the mathematical constant (approximately 2.72). R is performance and b is rate of learning. This curve is negatively accelerated.

Murdock was not centrally interested in individual differences but his equations bear directly on the present issue. In the process of curve fitting he derived two equations (equations 2 and 3) involving an empirical quantity m which he interpreted to be the immediate memory span. The first of these equations is:

$$\underline{R}_1 = \underline{kt} + \underline{m} \tag{2}$$

where $\underline{R}_1$ is the number correct on trial 1, k is an empirical constant, and t is a variable which was held constant in the present experiment. According to Equation 2, if kt is constant, then increases in m predict increases in R_1, in other words, Equation 2 predicts a direct relation between m and trial 1 performance so long as t (rate of presentation) is held constant.

The second equation involving m (Equation 3) relates learning (b) to immediate memory span:

$$\underline{b} \approx -\ln\,(.88 - \underline{m}/\underline{L}) \tag{3}$$

L is the list length and ln is the natural logarithm. Equation 3 predicts a direct relation between immediate memory span and b or learning rate: Assume that L is held constant and is always larger than m. Since m is always positive and, in practice, varies from about 1.5 to 8.5, the ratio of m to L will assume values from .15 to .85 or smaller (for L greater than 10) so the term in parentheses will be a positive number less than 1.0. Since the value in the parentheses is less than 1.0 the ln of this number will be negative (a property of logarithms) but b will be a positive number because it is the negative of a negative number. As m increases, the value of the number in parentheses decreases (becomes more negative) so b increases. For example, when L is assumed to be 10 and m is assumed to be 1.5, b is approximately equal to 0.31; the corresponding value of b for m equal to 8.5 is 3.5. These calculations illustrate that b varies directly with size of

m. In other words, Equation 3 predicts that the slope of the acquisition curve in multitrial free recall is directly related to size of m.

Method

Subjects. Ten high-span and ten low-span subjects were selected from the population of Western Carolina Center. The ten high-span subjects had a mean word span of 4.78 (IQ = 70.1; CA = 13.2) and the low-span subjects had a mean word span of 2.5 (IQ = 49.3; CA = 15.1).

Materials and equipment. Spans were tested with common single-syllable words judged to be highly familiar to the subjects; namely, tree, hair, cup, milk, shirt, phone, car, fish, soap, and grass. The test sequences were played on a Bell & Howell Language Master Machine.

The words for free recall were either high frequency (from the 500 most frequent words) or low frequency (once per 4 million words) according to the Thorndike-Lorge (1944) word list. They were presented by tape recorder. Four randomizations of each word pool were recorded with 2 sec between words and 60 sec between repetitions. The word 'ready' was recorded five sec before each list.

Procedure. Half of each ability group learned high-frequency words and half learned low-frequency words. Each subject took 16 trials.

Results and Discussion

The results are presented in Figure 1. These curves show the negative acceleration as expected from Equation 1 and which is typical for the multitrial free recall task. Because of the marked curvilinearity, comparison of the slopes of these curves is difficult. Figure 2 presents the same data plotted against log Trials. Plotted this way the curves are much more linear. The straight lines are the best-fit lines derived through linear regression analysis of the mean number correct against log Trial. The high-span subjects performed higher than the low-span subjects within each word-frequency condition. This finding supports the hypothesis that individual differences in span ability are directly related to performance levels. Within span groups the subjects performed uniformly higher on high-frequency words.

Insert Figure 1 about here

Insert Figure 2 about here

The high-span subjects produced greater slopes than the low-span subjects which supports the hypothesis that rate of learning is directly related to span ability. In contrast, Word Frequency did not affect slope. Within each span group the acquisition curves are parallel. Thus, these data indicate that span ability affects both performance and learning; but Word Frequency affects only performance.

Clearly, span ability predicts free recall in retarded subjects; but span ability and IQ correlated .82 in these subjects. Is it IQ or span ability which produces the observed effects? Span correlated .64 with total number correct across 16 trials. The corresponding correlation for IQ is .56. The partial correlation of span and total correct with IQ held constant is .44 ($p < .03$); but the correlation between IQ and total number correct with span held constant is .02. These statistics indicate that, compared with IQ, span is the more fundamental capacity variable in multitrial free recall. This conclusion must be qualified as follows. While it is true the span variable was a better predictor than IQ of the data of this experiment (the correlations were slightly larger and held even when IQ was controlled statistically), no attempt was made to control the range over which IQ and span varied. In terms of standard deviation units, the present range of span abilities may have exceeded that of the IQ scores. Under such conditions, the partial correlation technique is not really appropriate for comparison of the two tests. Nevertheless, the experiment as a whole does warrant the conclusion that the span test is better than the IQ for predicting individual differences in this task and these subjects. The correlations between span and performance were somewhat larger than the correlations of IQ and performance and the span test is a quicker and easier test to administer. Furthermore, the span scores enter into Murdock's empirical equations while IQ scores did not. In other words, span ability appears to be a more fundamental measure of individual differences in multitrial free recall because numerical values for immediate memory span emerge directly in multitrial free recall data. Undoubtedly, an equation in terms of IQ could be developed but it would be a more complex equation because a correction term of some sort would have to be incorporated to convert IQ to m values. The resulting equation would also lose its generality because normal IQ does not change appreciably as a function of age while span scores increase markedly during the developmental period. If IQ were to be substituted for m into Murdock's equations the value would have to be MA to apply generally to subjects during the developmental period.

Task Complexity Analysis of a Multidimensional Absolute Judgment Task

Task complexity analysis is the process by which Span Theory is used to make specific quantitative predictions of the performance of individuals or groups. The multidimensional absolute judgment task was chosen to illustrate the process because it is amenable to a fairly straightforward, but not obvious, task complexity analysis; but, more importantly, because the analysis bears directly upon an apparent problem with Span Theory.

George A. Miller (1956), in his Magical Number 7 $\pm$ 2 paper, concluded that the various Magical Numbers 7 in immediate memory span tasks, undimensional absolute judgment tasks, and the span of apprehension tasks, are coincidental; that it is unlikely that a common cognitive ability underlies individual differences in these three tasks. One of the main empirical reasons for Miller's conclusion was the fact that performance in multidimensional absolute judgment tasks is unconstrained by the magical Number. Miller cited Pollack & Ficks (1954) who found that performance in a multidimensional absolute judgment task greatly exceeded the magical number. It will be shown, however, that Pollack & Ficks' data can be fairly accurately predicted from Span Theory assuming span abilities which are typical for the subjects they used. It will be clear from this analysis that performance in this multidimensional task is constrained by span ability.

Unlike task analysis, task complexity analysis does not, in general, attempt to resolve a target task into a configuration of S-R bonds. Rather, a task complexity analysis attempts to analyze a target task into one or more elementary paradigms. Prediction of performance in the target task is then based upon the empirically determined characteristics of performance in the elementary paradigms. The elementary paradigms that have been used most in task complexity analysis include the various verbal learning paradigms, the short term memory paradigms, the discrimination learning paradigms, and the absolute judgment paradigms.

The present task complexity analysis makes use of just the immediate memory span paradigm and the unidimensional absolute judgment paradigm. The span paradigm makes use of a pool of stimuli for which the subject has well established discriminative responses. A number of the stimuli are presented and the subject responds to each stimulus using the responses he brings to the experimental situation. There are two critical characteristics of the span task important for the present analysis. First, performance is a sharply decreasing ogival function of the number of stimulus items presented on a test trial (Figure 3). Second, there is a broad range of individual differences in the span paradigm measured as the number of stimuli for which performance is 50% correct.

The unidimensional absolute judgment paradigm makes use of a pool of stimuli which vary along a unitary dimension such as size, tonal frequency, or loudness. The subject is instructed to judge the value of each stimulus by means of responses specified by the experimenter, such as the digits 1 through 10. Testing procedes one stimulus at a time in random order. The critical characteristics of this paradigm are quite similar to those of the span paradigm. When tested for each stimulus in a pool, overall percent correct is a decreasing ogival function of the number of stimuli in the judgment pool. There are individual differences in absolute judgment which can be measured as the number of stimuli which produce 50% correct performance. When measured in terms of numbers of stimuli, span abilities and absolute judgment abilities are quite similar (Miller, 1956; Bachelder & Chestnut, Note 1).

The complexity of a span task is defined as the number of stimuli in a stimulus string. The complexity of an absolute judgment task is defined as the number of stimuli in a stimulus pool. According to span theory, span ability (the ability to cope with task complexity) underlies individual differences in both tasks. Span ability is a limited capacity so that if a subject uses some of his span ability for a memory span function, he has less available span ability to apply to a concurrent absolute judgment task.

Tables 1 and 2 will help explain Pollack & Ficks' multidimensional task. They presented complex sounds (Table 1) for judgment. Each sound varied along the six dimensions shown. A subject listened to a complex sound then responded on a written answer sheet (Table 2) to indicate which of the several possible stimuli within each dimension had occurred. There were three versions of the basic task. In the binary version each dimension assumed either of two stimulus values which were always step 1 and step 5 of the chart. In the trinary version each dimension assumed just one of three values, step 1, step 3, or step 5 of the chart. In the quinary version each dimension assumed one of five stimulus values.

Table 1
The Dimension and Stimulus Values of
Pollack & Ficks' 6-Dimensional Tasks

Dimensions	Stimulus Steps				
	1	2	3	4	5
Frequency of tone (cps)	100	300	1000	3000	6000
Loudness (db)	40	75	90	100	105
Rate of interruption (ips)	0.4	1	2	4	10
Percentage time "on" (%)	10	30	50	70	90
Total duration (sec)	5	8	11	14	17
Direction (degrees)	-90	-45	0	+45	+90

Table 2
A Partial and Approximate Reconstruction
Of The Answer Sheet Used by Pollack & Ficks

		Frequency		
Very low___	Low___	Middle___	High___	Very high___
		Direction		
Left side___	Left middle___	Middle___	Right middle ___	Right side___

This task is analyzed as follows. It has the structure of a span task in that the subjects heard a series of six elementary stimuli then produced a series of six corresponding responses on their answer sheets. The task has the structure of a unidimensional absolute judgment task in that the specification of which stimulus element occurred within each dimension is an absolute judgment paradigm. In other words, the task is a unidimensional absolute judgment task imbedded within an immediate memory span task and will require the subject to split his span ability between the two tasks.

Table 3 schematizes the method used to predict the data. The top row of digits simply denote the six successive judgments for a single complex sound. The second row of digits indicate the complexities of each dimension. For the binary task each dimension has a complexity of 2; the trinary, 3; the quinary, 5. The third row of digits indicates the amount of span ability available for each successive judgment. The amount of span ability available was derived as follows. Each subject begins with his total span ability. Once the first stimulus value has been specified through absolute judgment the response is retained as in a memory span task. This retained response uses one unit of span so there is one less span ability unit for the next judgment. As the subject procedes through each complex sound there is less and less available span ability. Eventually a point is reached where the available span is less than that required for the next judgment. The model then assumes the subject guesses all remaining stimulus values with probabilities of being correct equal to .50, .33, and .20 for the binary, trinary, and quinary tasks.

Table 3
An Outline of the First Method of Making Predictions of Pollack & Ficks' Data

A Subject With Span Ability Equal to 5.0						
Steps of stimulus specification:	1	2	3	4	5	6
Complexity of each dimension:	3	3	3	3	3	3
Available span at each step:	5	4	3	2	1	0
Number of correct responses achieved through span ability:	3					
Number of correct responses achieved through guessing:	1 (.33 x 3)					
Total predicted number correct:	4.0					
Predicted information transmitted:	6.32 (4.0 x 1.58)					
A Subject With Span Ability Equal to 7.0						
Steps of stimulus specification:	1	2	3	4	5	6
Complexity of each dimension:	3	3	3	3	3	3
Available span ability:	7	6	5	4	3	2
Number of correct responses achieved through span ability:	5.0					
Number of correct responses achieved through guessing	.33					
Total predicted number correct:	5.33					
Predicted information transmitted:	8.42 (5.33 x 1.58)					

A span-5 subject has sufficient span available to correctly judge each of the first three stimuli. The remaining three are guessed with a probability of being correct equal to .33 so one correct response is expected through guessing. The total correct responses expected for a span-5 subject is 4. Pollack & Ficks did not report the number correct, rather they reported information transmitted which they calculated by multiplying the number of correct responses by $\log_2$ of the number of possible different stimulus values within each dimension. In the binary task each correct response earned 1.0 bit; in the trinary task, 1.58 bits, and in the quinary task, 2.32 bits. The present model predicts that a span-5 subject will achieve a mean of 4.0 correct responses in the trinary task, which when multiplied by 1.58 equals 6.32 bits of transmitted information.

Analogous calculations were carried out for the binary, trinary, and quinary tasks for assumed span abilities of 5, 6, and 7. These assumed span abilities were not chosen arbitrarily. Pollack & Ficks' subjects were 36 college students who produced a fairly wide range of individual differences in performance. The data reported were from three groups of subjects representing low, medium, and high performance levels. It was assumed that these differing performance levels reflected individual differences in span abilities. Accordingly, an attempt was made to estimate the typical low, medium, and high span abilities of the subjects who served in the Pollack & Ficks study. According to Span Theory, span ability is directly measured by the staircase word span test (Bachelder, 1970; Bachelder, 1977). Appropriate word span scores for 99 college students were available in the data files from a variety of experiments. From these 99 students, 36 were selected at random. These 36 were then ranked according to their word span scores and divided into six equal groups varying systematically in mean span ability. Pollack & Ficks reported the data of the top, third, and fifth sixths of their subjects so the corresponding groups were selected from the 36 college students. The mean spans of the final groups were 5.2, 5.8, and 6.97 which were rounded to 5, 6, and 7 for convenience.

Table 4 presents the observed data and the predicted data for the six-dimensional tasks. In general, the predictions are quite accurate with the overall error of prediction less than 3 percent. Some errors, however, were quite large (+17% and -26%) and there is no obvious explanation for them.

Table 4
The Observed and Predicted I_ts for
the 6-Dimensional Tasks

Performance Groups	Low	Medium	High
	Binary Task		
Observed I_ts	4.8	5.3	5.6
Predicted I_ts	5.0	5.5	6.0
Error	+4.2%	+3.8%	+7.1%
	Trinary Task		
Observed I_ts	5.4	6.6	7.6
Predicted I_ts	6.32	7.38	8.42
Error	+17.0%	+11.8%	+10.8%
	Quinary Task		
Observed I_ts	6.3	7.2	7.9
Predicted I_ts	4.64	6.50	8.35
Error	-26.3%	-9.7%	+5.7%
Overall mean error = 2.71%, SD = 13.13			

The method just described is quite adequate to make the point that multi-dimensional absolute judgment data can be closely related to span scores and thereby resolve the problem for Span Theory posed by Miller's paper. The method makes assumptions, however, which are not strictly true of unidimensional absolute judgment. The method assumes that if task complexity is at or below available span ability the subject performs perfectly and that if task complexity is beyond span, subjects guess the responses. In fact, performance is a decreasing ogival function of increasing task complexity (see Figure 3). Another method of analysis has been developed which reflects this fact and which makes better predictions of Pollack & Ficks' data. This second method is based upon another concept of Span Theory; namely, relative difficulty. Relative difficulty is defined as the ratio of task complexity to span ability. According to Span Theory, performance is a decreasing ogival function of relative difficulty, an expectation supported by unpublished data by Bachelder & Chestnut (Note 1).

Insert Figure 3 about here

Bachelder & Chestnut studied the relation between span ability and unidimensional absolute judgment in normal and retarded adults. The span abilities of the subjects ranged from 1.5 to 7.9 and the retarded subjects were mostly severely, moderately, and mildly retarded. Each subject performed in three judgment tasks with theoretical complexities of 4, 6, and 8. The subjects were grouped by span ability into 12 fairly homogeneous span groups of five subjects each. Figure 4 presents the mean percent correct plotted against $\log_2$ of relative difficulty. The reason the data are plotted against $\log_2$ relative difficulty is to transform the curve from the ogival form to a more linear form to simplify analysis. This transformation was quite successful but the underlying ogival form is still fairly evident. Mean percent correct correlated .94 with mean $\log_2$ relative difficulty. The linear regression equation expressing mean percent correct as a function of $\log_2$ relative difficulty is

$$\% \text{ Correct} = -37.45 \log_2 \text{relative difficulty} + 72.27 \quad (4)$$

Insert Figure 4 about here

Table 5
A Schematic Presentation of the Second
Method of Prediction; Span-5 Subject, Trinary Task

	3	4	5	6	7	8
Steps of analysis	1	2	3	4	5	6
Task complexity	3	3	3	3	3	3
Available span	5	4	3	2	1	0
Relative difficulties	.60	.75	1.0	1.5	3.0	-
Predicted % correct	.999	.878	.723	.504	.33	.33

Total = 3.77 (X 1.58 = 5.96)

Equation 4 was used to predict the percent correct responses in Pollack & Ficks' tasks. Table 5 illustrates the process for a span-5 subject in the trinary task. The top row of digits are the familiar six steps of the analysis of one complex sound. The second row of digits are the task complexities of each step of the task. The third row are the available spans assuming that retained items subtract from total span ability just as was assumed in the first method. The fourth row of numbers are the relative difficulties, that is, the ratio of the numbers in the second row to the numbers in the third row. The fifth row of numbers are the percents correct predicted by equation 4 for each of the relative difficulty values. Note, however, that where the predicted percent correct fell below the probability of guessing correctly, the number entered is the probability of guessing correctly. Total performance is predicted by summing the values in the fifth row and multiplying by 1.58, the number of bits earned by each correct response.

Table 6
The Observed and Predicted Values;
Second Method

Performance Groups.	Low	Medium	High
	<u>Binary Task</u>		
Observed values	4.8	5.3	5.6
Predicted values	4.66	5.16	5.66
Error	-2.9%	-2.6%	+1.1%
	<u>Trinary Task</u>		
Observed values	5.4	6.6	7.6
Predicted values	5.94	7.01	8.06
Error	+10.0%	+6.2%	+6.1%
	<u>Quinary Task</u>		
Observed values	6.3	7.2	7.9
Predicted values	5.56	7.01	8.64
Error	-11.7%	-2.6%	+9.4%
Mean error = +1.4%, SD = 7.12			

Table 6 presents the values reported by Pollack & Ficks with the values predicted by this second method. The errors of prediction range from -11.7% to 10.0% with a mean of 1.4% and a <u>SD</u> of 7.12.

<u>Span Testing of Nonverbal Physically Handicapped Young Adults</u>

One obvious implication of Span Theory and the work presented so far is that the span test should have considerable psychometric utility. In fact, it will be argued that for certain purposes and for certain populations the span test may be the most appropriate of the intelligence tests. Table 7 lists several advantages of the span test over most IQ tests. This table is not meant to imply the general superiority of span tests over IQ tests. On the contrary, the span test has distinct limitations and should be viewed as an adjunct to IQ testing. As conceived in Span Theory, span ability is a very narrowly defined ability although conceived to be fundamental to diverse tasks. According to the theory, the span test does not measure such important aspects as motivation for school learning or achievement levels in key intellectual tasks. The span test can be no match for an omnibus test of an individual's ability to achieve in our cultural setting. Assessment of this overall ability, intelligence if you will, is what is most important in service settings.

Table 7
Advantages of the Span Test in Comparison with IQ Tests

1. One simple span test, or minor variations of the test, may be used with quite diverse populations.
2. Through the process of span task analysis a span score may be used to make quantitative predictions of learning and performance in a wide range of tasks.
3. The span test assesses the general ability to cope with numbers of conjunctively relevant stimuli; it does not test the achievement of specific responses and skills.
4. The particular stimuli and responses used to test span are largely irrelevant. So long as we can specify a small number of associated stimuli and responses known to exist in a specific subject's repertoire, we can measure the subject's span.
5. The extent to which a subject is educationally deprived has no theoretical bearing upon the subject's span score, so long as stimuli and responses can be found with which to measure an individual's span.

With these points in mind several versions of the span test were developed for use with nonverbal seriously physically handicapped young adults. The subjects to be reported on had generally been considered to be severely retarded because of their physical handicaps and because of their low IQ scores on standard tests. Tables 8 and 9 summarize the characteristics of these people. Not only could they not communicate normally but they had had long histories of institutionalization. Until shortly before this study began their primary treatment had been fine nursing care. For people such as these the span test may be the only test of cognitive ability which is at all appropriate. Because of their physical handicaps and their long institutionalization, they simply have not been exposed to classroom experiences which might have developed basic knowledge, concepts and cognitive skills.

Table 8
Clinical Data for the Physically Handicapped Students

Student	CA	Years of Institutionalization	Physical Diagnosis
C. W.	14-3	9-0	Spastic quadriplegia
G. W.	20-3	1-3	Severe spastic quadriplegia
G. M.	22-1	9-0	Severe athetosis Spastic quadriplegia
T. C.	22-9	6-3	Severe athetosis
K. S.	23-11	17-9	Severe athetosis Spastic quadriplegia
S. O.	19-9	10-6	Severe athetosis Mild spastic quadriplegia
Mean	20-6	9-0	

Table 9
Clinical Data for the Physically Handicapped Students

Student	Span	Intellectual Diagnosis
C. W.	1.9	Severe MR Profound-severe MR
G. W.	2.5	Profound MR Severe MR Speculative IQ of 30 - 50 "Standard tests inaccurate"
G. M.	3.5	"No appropriate test" Severe MR Trainable MR Mild MR
T. C.	4.5	Moderate-severe MR
K. S.	4.3	Profound MR Severe MR High severe-low moderate MR
S. O.	5.5	"No test given, probably higher than can be tested." Severe MR "Probably moderate MR"

At this writing five of these people had been placed in a special class for seven months to learn to communicate via a symbol board although most have had as much as two years' previous experience with the symbol boards. The symbol board is a matrix of symbols and words to which they can point and thereby converse with their teachers and classmates. Figure 5 is a detail of one of these boards. The actual size of the cells is typically 25 mm by 25 mm but varies with the needs of the individual. Each student had learned to point to specific symbols upon request and it is this skill which made symbol board span testing possible.

Insert Figure 5 about here

The students had different vocabularies and different response modes so several versions of the symbol board span tests were developed. Test vocabularies for individuals were determined by asking their teacher to list the most familiar and accurately used symbols for each student. Nine of these words were selected (single-syllable nouns preferred) for each student and special span testing symbol boards were made. Figure 6 is the span testing symbol board using Rebus symbols. Figure 7 is the span testing symbol board using Bliss symbols.

Insert Figures 6 and 7 about here

Some of the students could point out responses with their hand. With these students span testing was easy. The tester said the word sequences and the student pointed out the responses. Some of the students could not point but could nod or otherwise indicate a target symbol when the board was "scanned" by the tester. By scanning is meant pointing successively at columns or cells. The tester scanned the board first for the correct column. When the column was indicated the tester then scanned down that column until the student indicated a specific symbol. The tester then called out this response and began to scan for the next symbol. One subject was deaf so the words were presented in American Sign Language and she responded by pointing.

Table 10 presents the highest score for each student and the mean of his or her two most recent tests. Obviously, there is little normative data for

Table 10
The Symbol Board Span Scores for the
Physically Handicapped Students

Student	Highest Score	Mean of Two Most Recent Tests
C. W.	1.9	1.7
G. W.	2.5	-
G. M.	3.5	-
T. C.	4.5	4.3
K. S.	4.3	3.9
S. O.	5.5	5.0

these tests, so percentiles corresponding to the individual scores cannot be presented. The scores can, however, be compared cautiously with those found using the standard staircase word span test (Bachelder, 1970; Bachelder, 1977). Figure 8 presents the distribution of span scores for a group of 20 retarded adults ranging in IQ from 18 to 80, most of whom ranged from severely to mildly retarded. The data were lifted intact from Bachelder & Chestnut (Note 1). Most of these scores range between 1.7 and 4.7. Figure 8 also presents the distribution of standard staircase word span scores for 99 college students who had participated in one or another study in the present research program. Most of these scores range from 4.7 to 7.2. The lower scores among the physically handicapped students fall in the range of scores common among retarded people; but the higher scores of the physically handicapped students are not uncommon among normal adults. It seems fair to conclude that these span test scores suggest that several of the students have cognitive abilities well beyond that normally attributed them by staff at Western Carolina Center.

Insert Figure 8 about here

These findings show that span tests can be given to this population and that the results are potentially more useful than the IQs and psychological impressions which were in their clinical folders. The span tests suggested that this group is not highly homogeneous in cognitive ability. The range of ability revealed was quite large and suggested that a few of the individuals might have near normal capabilities. This finding was particularly important at Western Carolina Center because the impressions conveyed by the psychological examinations (Table 9) had seriously impeded attempts to provide classroom and life experiences commensurate with the abilities some individuals felt the students to have. The very fact that a few of the students scored quite high on the span test considerably softened the resistance of key staff members and helped ensure continuity of education services and involvement in a wide variety of stimulating activities previously judged to be a relative waste of resources.

Even though the symbol board span tests appear to be theoretically valid as span tests and the initial findings appeared consistent with past experience with both span scores and the students involved, much more objective evidence is needed to validate the use of the symbol board span tests. How comparable are the symbol board span test scores with the standard staircase word span scores? How closely related are the span scores to other measures of cognitive functioning? Some data are available which bear on each of these questions.

In order to determine if symbol board span test scores are comparable in absolute magnitude to scores from other span tests pilot data were collected from ten normal adults quickly assembled from involved staff or interested colleagues. Each subject took three tests in one session, usually in the following order: the auditory-vocal span test (the standard staircase word span test), the pointing symbol board span test, then the scanning symbol board span test. The results are presented in Table 11. The tests produced highly similar scores with these subjects much to the surprise of several of them. During testing these subjects often expressed the opinion that the scanning symbol board test was much more difficult than the auditory-vocal test. They expected their scanning scores to be lower than the others, which they were; but only to a slight, and not very important, extent. The means were compared by a one-way analysis of variance for repeated measures; $F(2, 18) = 0.95$.

Table 11
The Mean Span Scores of Ten Normal Adults for
Three Different Span Tests

Auditory-Vocal	Pointing Symbol	Scanning Symbol
5.30	5.32	5.11
SD = 1.22	SD = 1.09	SD = .964

Individual differences in the span scores of the handicapped students closely corresponded to individual differences in measures of their language comprehension. Each student had taken the Assessment of Children's Language Comprehension (ACLC) test (Foster, Giddan, & Stark, 1972). In this test the student must point out specific pictures upon command and the instructions vary

Table 12
The Span Scores and the Language Comprehension Scores (ACLC)

Student	Span	ACLC	% correct on each number of critical elements 1	2	3	4
C. W.	1.9	66/80	70	90	40	40
G. W.	2.5	61/80	86	90	50	40
G. M.	3.5	70/80	86	90	80	60
T. C.	4.5	73/80	100	100	70	60
K. S.	4.3	79/80	100	100	90	100
S. O.	5.5	78/80	100	100	100	80

in the number of critical elements necessary for correct choice (1-4 critical elements are tested). Figure 9 shows two choice arrays with examples of instructions. The top example tests for two critical elements and the bottom example tests for four. Table 12 presents ACLC scores and the span abilities of each of the students. The correlation between span and ACLC is .86. Of more interest is the fact that individual patterns of performance on the ACLC can be predicted from the span scores and Span Theory. According to Span Theory, correct performance on two critical elements requires a span of 2, correct performance on three critical elements requires a span of 3, and so on. In general, this is exactly what was observed in these students.

Insert Figure 9 about here

Language samples also seem to show a close relation to span scores. From time to time the teacher had collected a sample of the language of each student by asking several questions and writing down the responses. These samples are rather informal but they will serve to suggest the relation between span and language. There are two indices of the language samples. The upper bound index or UBI is the number of words in the longest utterance. The mean length of utterance or MLU is the mean number of words in each utterance. Table 13 presents the spans and UBIs and MLUs for the two most recent language samples. The correlations between span and mean UBI and mean MLU were 0.81 and 0.92, respectively. Tables 14 and 15 present transcriptions of fairly representative language samples for the lowest-span student and the highest-span student. It is fairly obvious that the language of the higher span student is much more complex.

Table 13
Symbol Board Spans, Upper Bound Indices, and Mean Lengths of Utterance for Two Most Recent Language Samples

						MEANS	
Student	Span	UBI	MLU	UBI	MLU	UBI	MLU
C. W.	1.9	2	1.0	2	1.2	2	1.1
G. W.	2.5	3	1.2	2	2.0	2.5	1.6
C. M.	3.5	3	2.0	-	-	3.0	2.0
T. C.	4.5	4	3.5	3	2.1	3.5	2.8
K. S.	4.3	7	4.1	10	4.2	8.5	4.15
S. O.	5.5	10	6.3	9	5.1	9.5	5.7

Table 14
A Sample of C. W.'s Expressive Language (Span = 1.9)

What happened last night?		
	(C. W.)	T.V.
What did you see?		
	(C. W.)	Wrestling (in sign language)
What are you looking at?		
	(C. W.)	Book
What book?		
	(C. W.)	Book
Yes, but what kind?		
	(C. W.)	Wrestling book
What do you do before school?		
	(C. W.)	Eat
Use all your symbols.		
	(C. W.)	Eat food
How do you get to school?		
	(C. W.)	Roll
Roll what?		
	(C. W.)	Chair

Table 15
A Sample of S. O.'s Expressive Language (Span = 5.5)

(S. O.) When new paper work

Jerry is working on it. What will you do if you get that job?

(S. O.) Go to all rooms and give out papers

Do you have anything to say to Gina?

(S. O.) Yes. Why you not want to work wheelchair
You can be able to use it
When she [Kay] have wheelchair?

I don't know.

(S. O.) I sick and tired part eye (see)* her
have to be pushed

*"part eye" is the local symbol board idiom for the verb, 'to see'.

Table 16
The Number of Vocabulary Items (Symbols and Written Words) Acquired Since August, 1977

Student	Span	Number of Items
G. W.	2.5	42
T. C.	4.5	70
K. S.	4.3	80
S. O.	5.5	85

Finally, gross assessments of the relation between span and classroom acquisition were made by comparing span ability with the size of vocabulary. In general, the class focussed on teaching new vocabulary items and added these to the boards of each student as the words were acquired. Table 16 presents the span abilities and the number of symbols on the symbol board of each student who was enrolled in the class. In general, the higher-span students had the larger vocabularies.

Reference Note

1. Bachelder, B. L. & Chestnut, W. Memory span and absolute judgment in normal and retarded adults. Paper in preparation, Western Carolina Center, Morganton, North Carolina 28655, 1976.

References

Bachelder, B. L. The memory span paradigm: Its use for the analysis of mental retardation. (Doctoral Dissertation, Michigan State University, 1970.) Dissertation Abstracts, 1971, 32(1), 576B-577B. (University Microfilms No. 71-18,158)

Bachelder, B. L. Memory span tests and the ability to imitate word sequences and sentences. The Ohio Journal of Speech and Hearing, 1977, 13(1), 5-15.

Bachelder, B. L., & Denny, M. R. A theory of intelligence: I. Span and the complexity of stimulus control. Intelligence, 1977, 1, 127-150.

Bachelder, B. L., & Denny, M. R. A theory of intelligence: II. The role of span in a variety of intellectual tasks. Intelligence, 1977, 1, 237-256.

Foster, R., Giddan, J. J., & Stark, J. Manual for the assessment of children's language comprehension. Palo Alto: Consulting Psychologists Press, 1972.

Miller, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 1956, 63, 81-97.

Murdock, B. B. Immediate retention of unrelated words. Journal of Experimental Psychology, 1960, 60, 222-234.

Pollack, I., & Ficks, L. Information of elementary multidimensional auditory displays. Journal of the Acoustical Society of America, 1954, 26, 155-158.

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

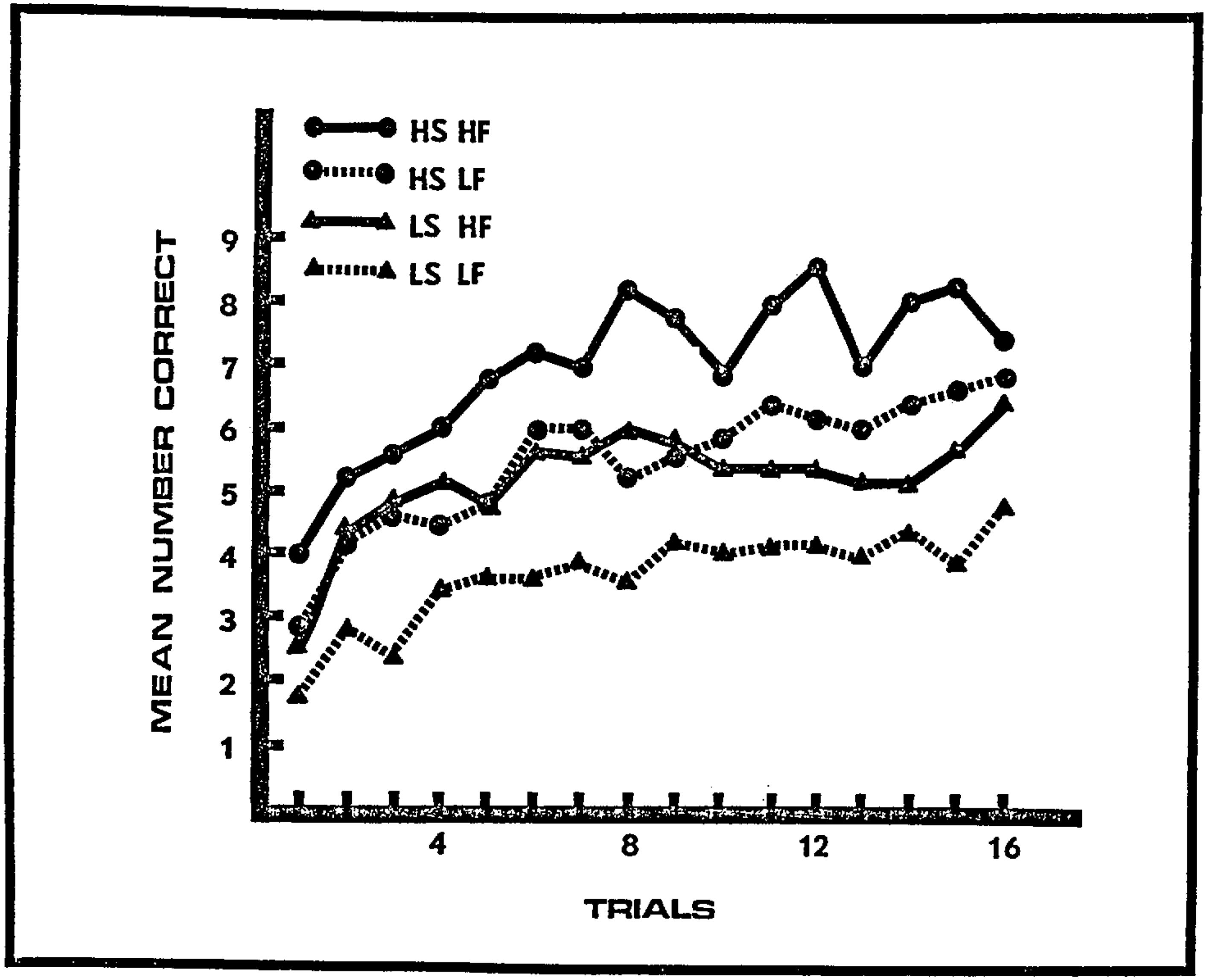

Fig. 1. The mean number correct in a free recall task as a function of Trials, Span (high span versus low span), and Word Frequency (high frequency versus low frequency).

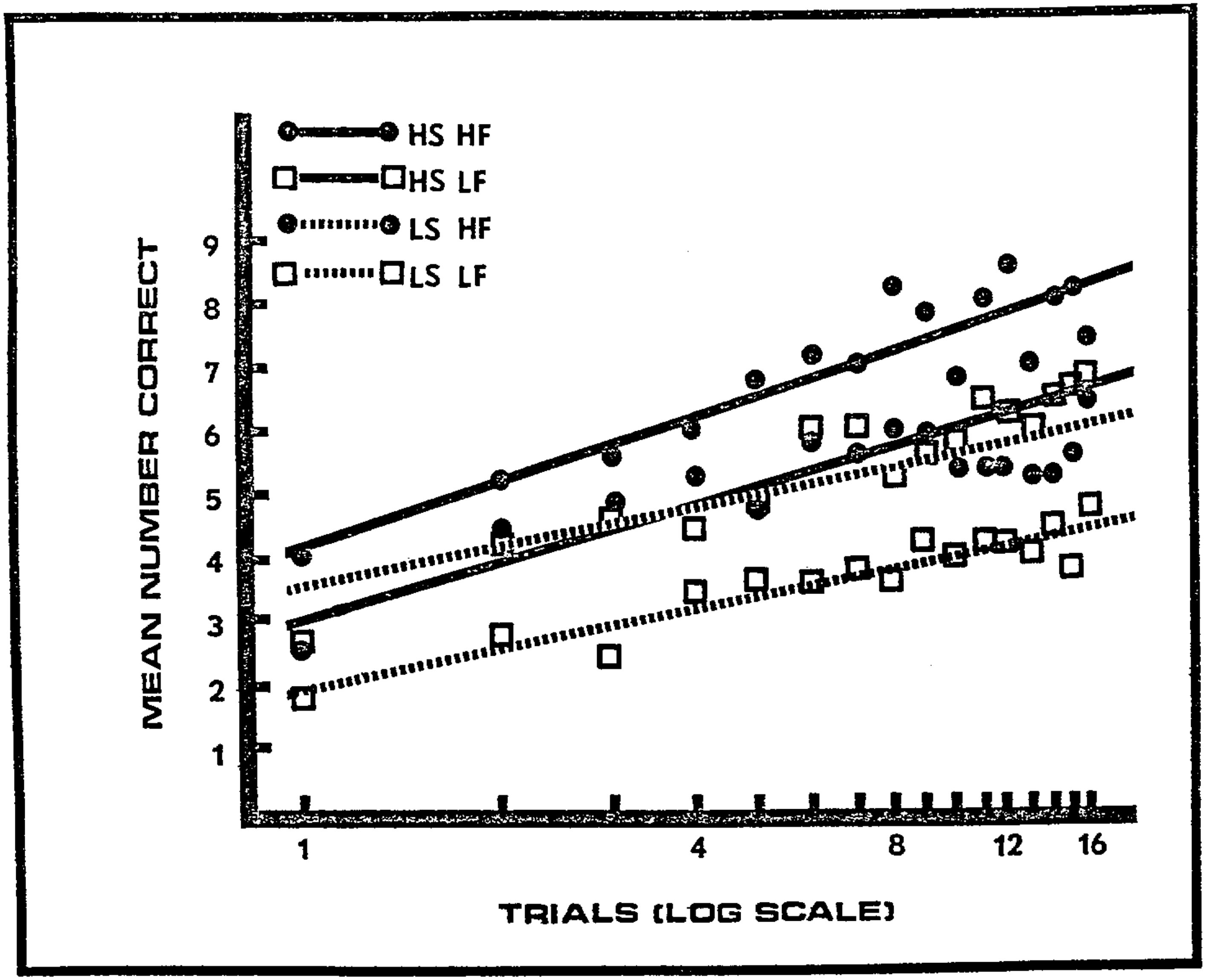

Fig. 2. The mean number correct in the free recall task as a function of $\log_{10}$ Trials, Span (high span versus low span), and Word Frequency (high frequency versus low frequency). Figure 2 presents the same information as Figure 1 except Trials are plotted on a logarithmic scale.

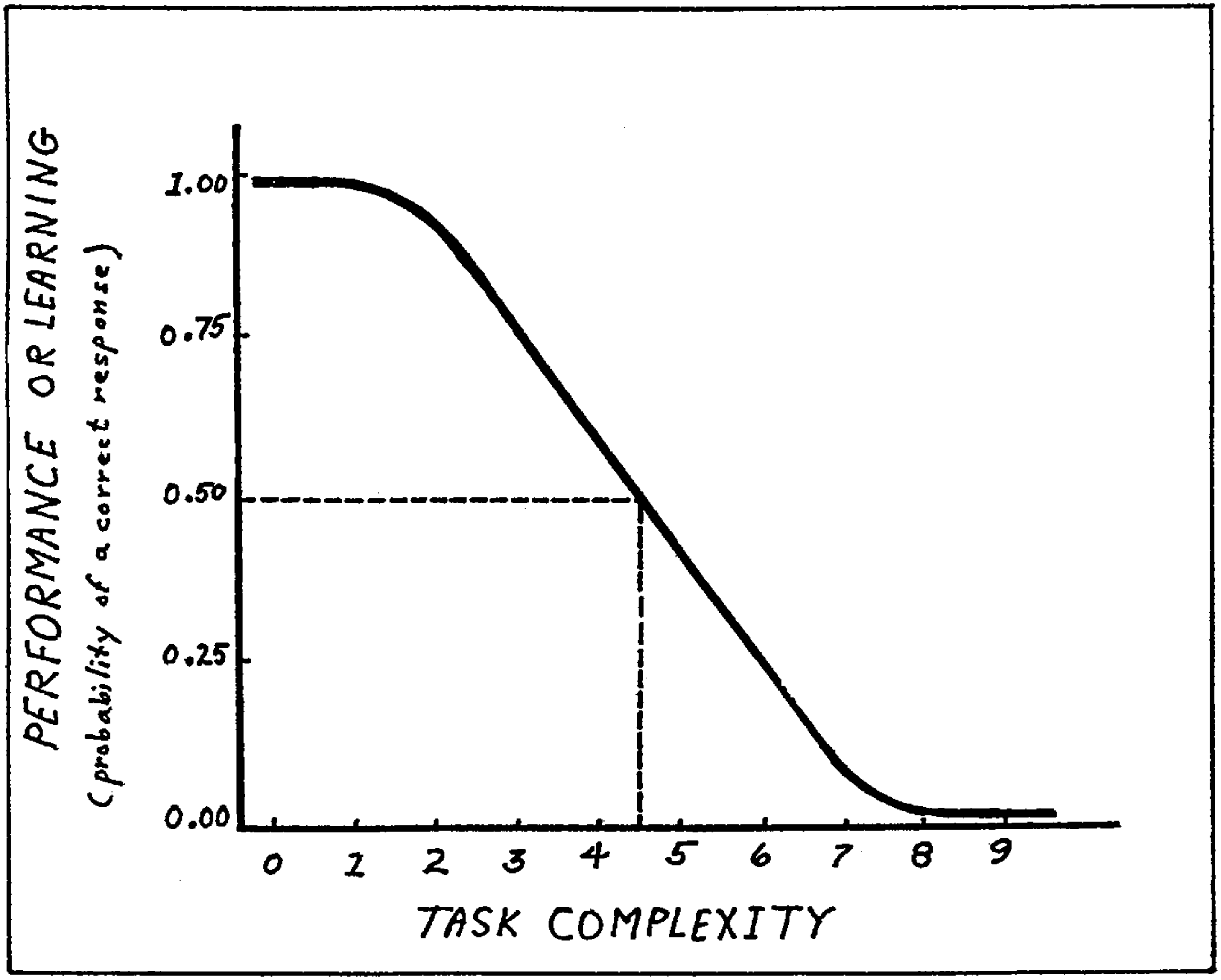

Fig. 3. The theoretical relation between performance or learning and task complexity. The dotted lines indicate a measured span ability of 4.5.

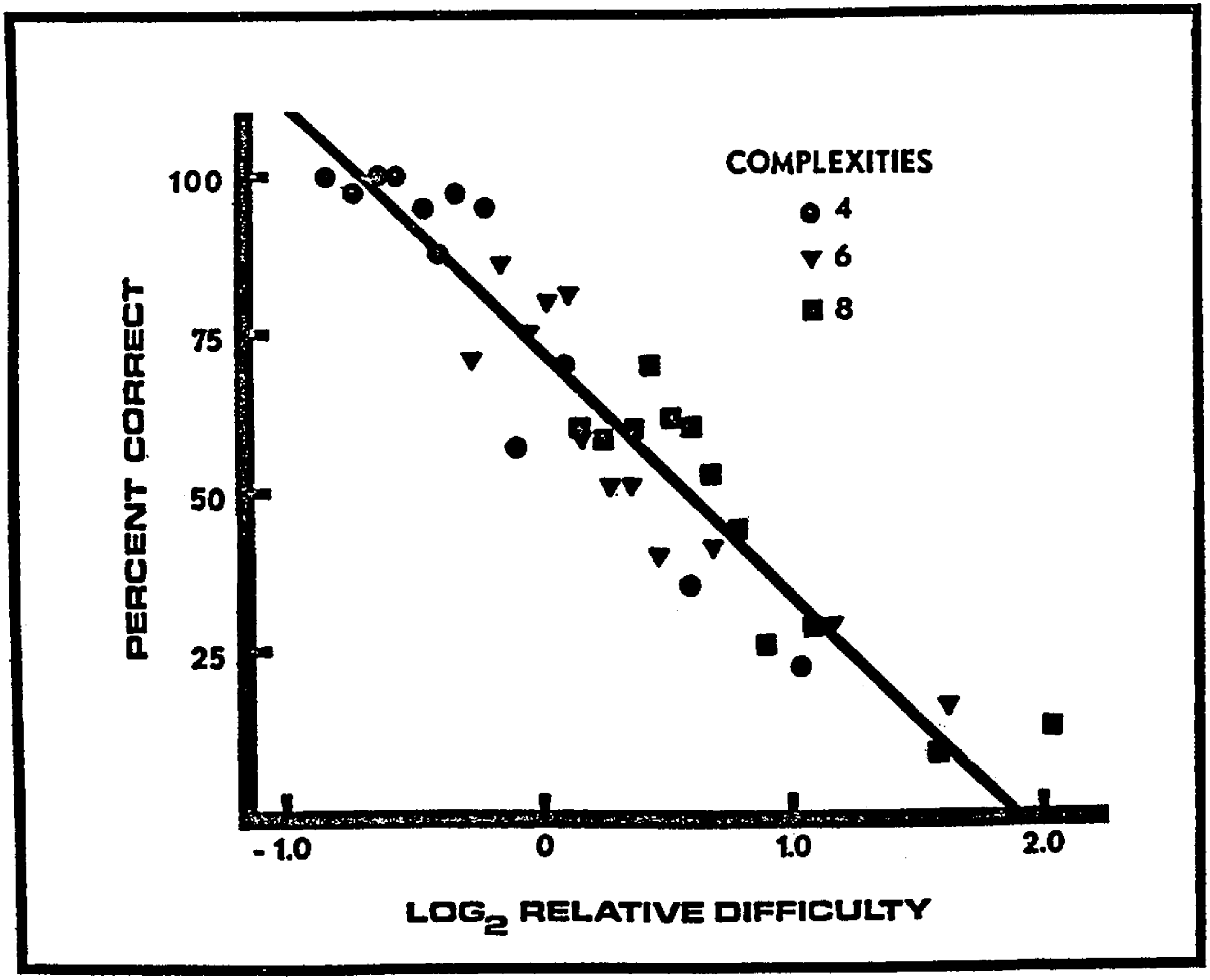

Fig. 4. The percent correct in a unidimensional absolute judgment task as a function of $\log_2$ relative difficulty.

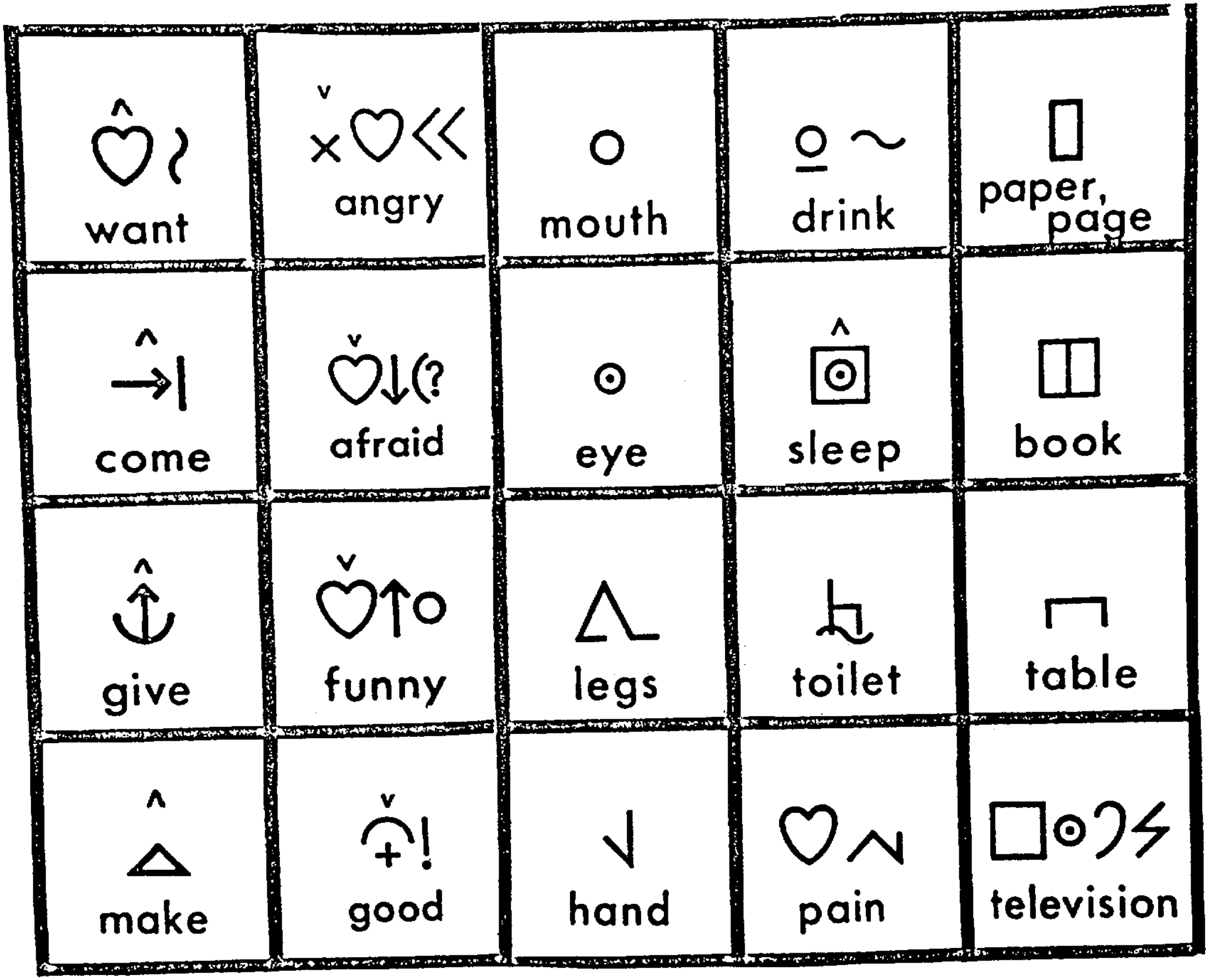

Fig. 5. A detail of a Bliss symbol communication board.

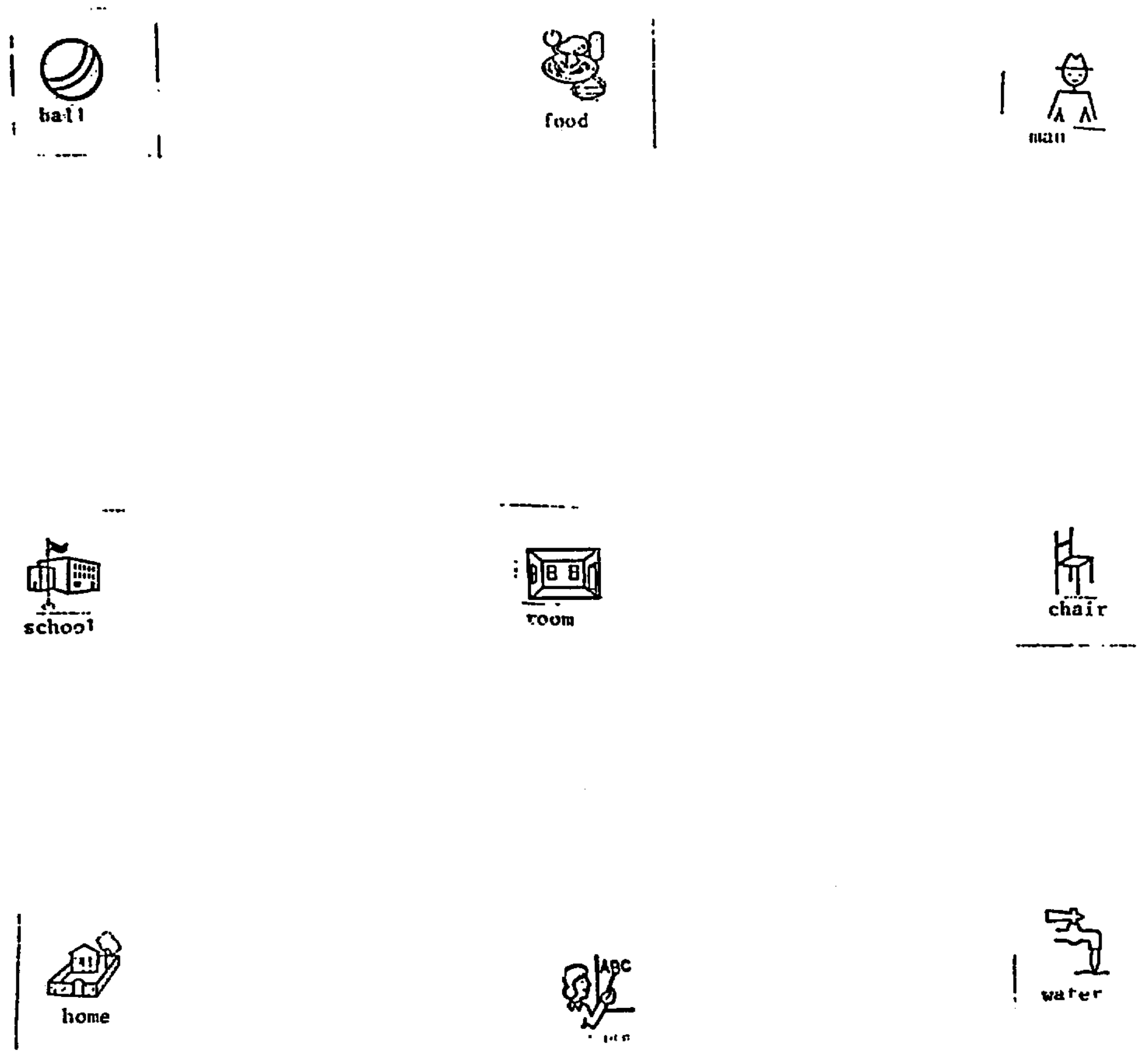

Fig. 6. A photograph of the span testing symbol board using Rebus symbols.

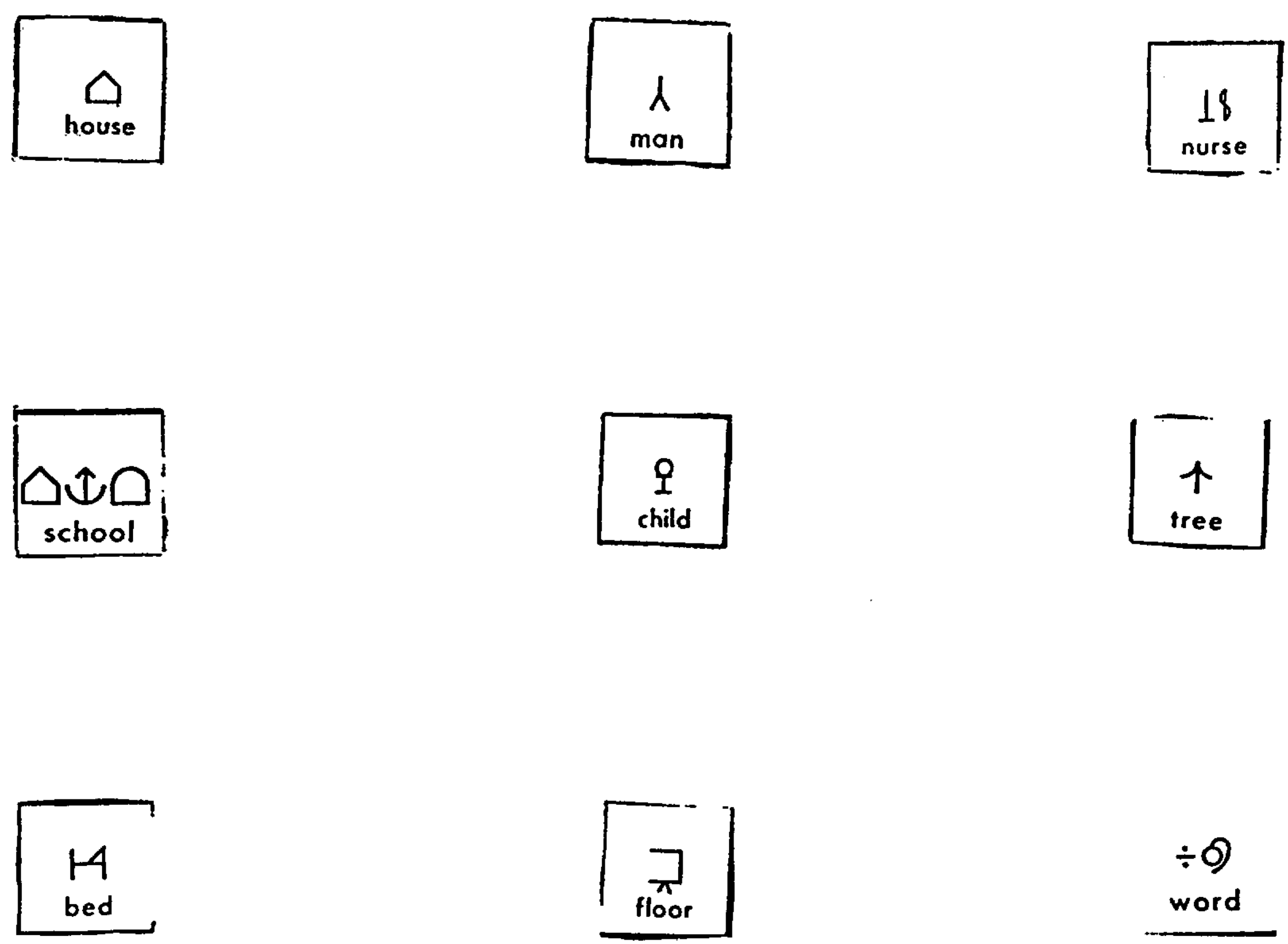

Fig. 7. A photograph of the span testing symbol board using Bliss symbols.

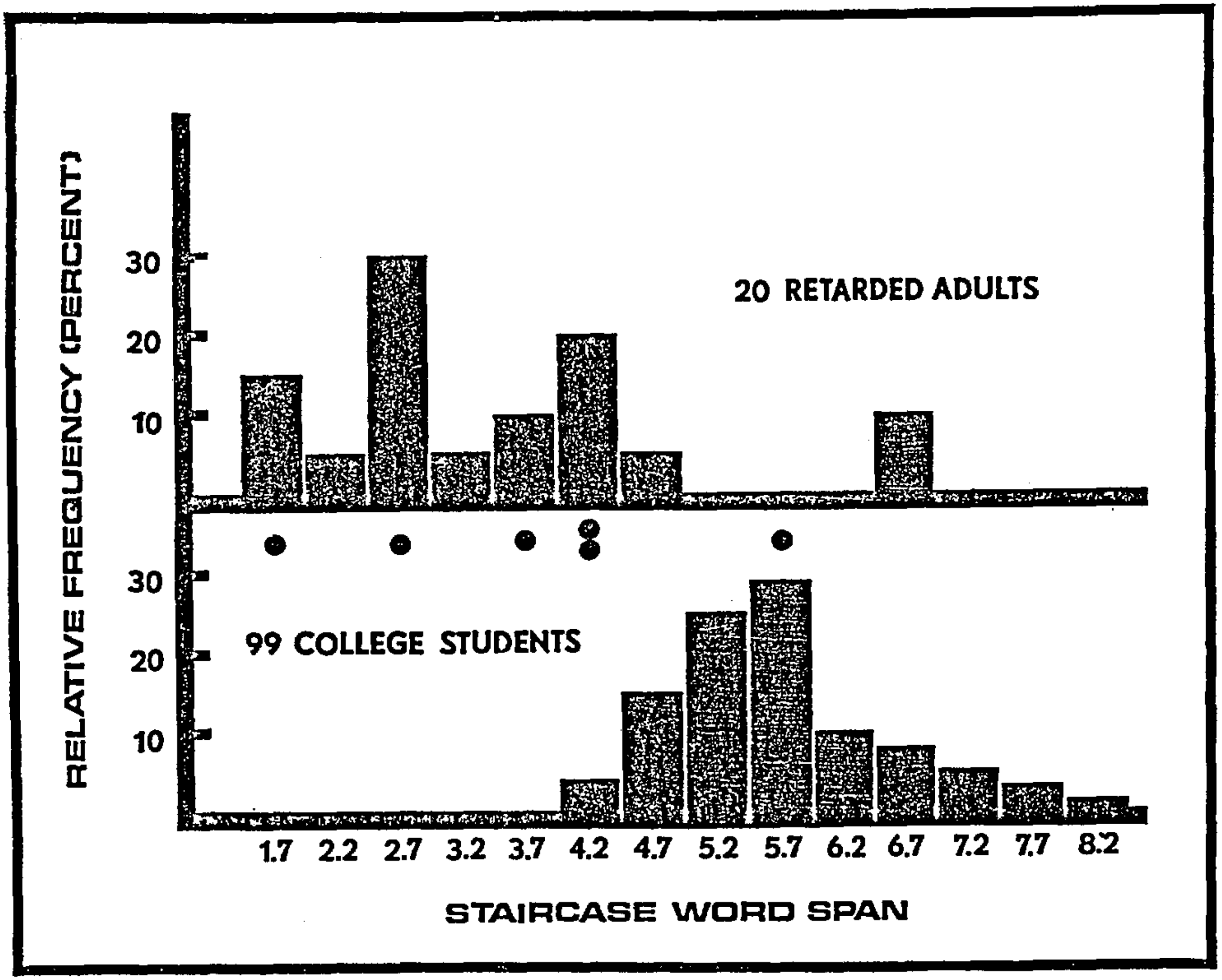

Fig. 8. A comparison of the span abilities of subjects who have served in various research projects in the present research program. The solid circles across the center of the figure denote the symbol board span scores of the physically handicapped individuals who served in the present study.

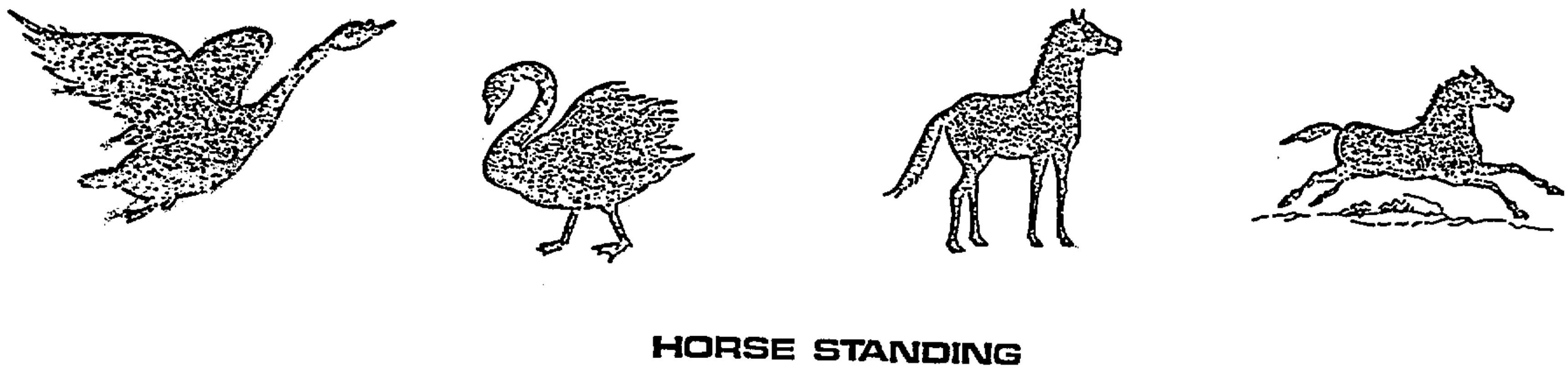

Fig. 9. Examples of two plates from the Assessment of Children's Language Comprehension (ACLC) test. The top plate tests for two critical elements and the bottom plate tests for four critical elements.