A Research Strategy for Studying[1] Cognition and Intelligence Through Study of Individual Differences in Span Ability

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

The purpose of this paper is to describe the strategy I have been using to study intelligence and to illustrate briefly the type of research the strategy has generated. My research is a combination of psychometric research and experimental analysis of behavior in the immediate memory span task. Immediate memory span tasks are of interest because they appear as subtests on the WAIS, the WISC, and the Stanford-Binet IQ tests.

There are four general areas of research in my program: the continuing development of a span psychometric, the experimental analysis of behavior in span tasks, the search for validity for the span task, and the exploration of the utility of the span task as a generalized research task for investigating behavioral phenomena other than span.

In order to develop a psychometric test, one should have a clear concept of the trait to be measured. The face validity of a test is a judgment of what the test appears to be measuring and span tests have been ascribed a rather broad range of face validities.

Staats (1968) speaks of the span tests as measures of the ability to form rote associations. Jensen (1970) has a similar concept of span and uses span tests as a measure of Level I intelligence. Hilgard (1951) described span tests as measures of the number of items which can be learned in one trial. Miller (1956) describes the span phenomenon as a limitation on our ability to process information. Humpstone (1918) describes span as the ability to distribute attention over a series of discrete perceptions, and similarly, in the same paper, the ability to grasp a number of different elements in one operation of attention. Leaming (1922) wrote that span tests measure the number of discrete units over which the individual can successfully distribute his attention and still organize them into a working unit. Ellis (1963) speaks of span tests as good measures of the Hullian construct, stimulus trace. Murdock (1974) cites span phenomena as critical data to be explained by theories of short term memory.

No one of these diverse face validities appears to be more correct than any other so it is unclear just what basic processes are being measured. Even though there is disagreement on the specific basic processes measured by a span test, there is virtually complete agreement on the operational face validity of a span test; that is, a span test is easily recognized as a span test.

The essential operations of a memory span task include presenting a series of stimuli and requiring that the subject respond in some way to each stimulus. If we score the subject as right or wrong on each stimulus string, we note that the number of items in a stimulus string is a

----

[1]This paper was presented at the Tenth Annual Gatlinburg Conference on Research in Mental Retardation, Gatlinburg, TN, March 9-11, 1977. The present electronic version was scanned from an original on 4-14-08. References were revised to modern format.
[2]306 West Union St, Morganton, NC 28655. Bruce@BruceBachelderPhD.com www.BruceBachelderPhD.com
[3]No longer affiliated with Western Carolina Center. Independent practice from 1980 through retirement in 2004.

potent variable. When performance is plotted as a function of increasing size of stimulus string there is a pronounced threshold effect. If stimulus strings are much above or below span threshold, performance is 100% in error or 100% correct, respectively. In view of this observation the staircase procedure for measuring psychophysical thresholds was appropriated for use as a span test. In this method the subject first attempts a simple subspan stimulus string, say two items in length. Then whenever the subject responds perfectly, his next string is one item longer. Whenever he errs, the next string size is one item shorter. This staircase procedure is continued for a fixed number of strings. The string sizes presented increase rapidly to a threshold point, then oscillate around the threshold. The span score is the mean string size presented in the plateau portion of the protocol. This method has excellent operational face validity as a measure of span ability, and the same specific test materials and procedures can be applied directly to a broad range of subjects from young children to adults and from severely retarded subjects to bright normal subjects. Using this method, we measure word spans in a range from about 1.5 to 8.1 words.

Experimental analysis of behavior in the staircase span task can lead both to an understanding of the effects of experimental variables and to improvement of the face validity of the span task. For example, I have investigated the effects of Grouping and Rate of Presentation on the staircase digit span. Only the Rate variable will be discussed here. The subjects were 42 borderline adults and 42 mildly retarded adults. Spans were measured at each of four rates ranging from one digit every 2 sec to 4 digits per sec. Rate and IQ interacted in their effects on digit span. Slow rates of presentation increased the digit spans of the higher-IQ group, but had no significant effect on the digit spans of the lower-IQ group. In fact, among the lower-IQ group there was a definite trend in the opposite direction. Slow rates appeared to decrease the digit spans of the lower-IQ subjects. Subsequent research has confirmed this trend among retarded subjects. At the higher presentation rates, rate had little effect in either IQ group. Another way to describe these data is that at high presentation rates the only relevant variable was span ability. For this reason, the present research program has standardized on rapid presentation rates, usually 2 items per sec.

The next experiment illustrates the search for criterion-related validity. A test has criterion-related validity if it has predictive utility. The criterion task for this experiment was the free learning task. The subjects were 10 borderline and mildly retarded adolescents and 10 moderately retarded adolescents. Each subject took 16 free learning trials. Half learned high frequency words and half learned low frequency words, but these data will not be reviewed here.

Performance was measured in terms of the total number of correct responses. Span ability correlated .675 with free learning. IQ and free learning correlated almost as well, .569. Partial correlations, however, indicated that span ability is predictive of performance while IQ is not at all predictive. The correlations were .44 and .04, respectively. The multiple correlation between span, IQ, and free learning was .68, the same as span and free learning alone.

The next experiment illustrates the search for construct validity. A test has construct validity if the test scores behave in ways our theoretical expectations demand. Thus one judges construct validity with respect to both theory and data.

In many immediate memory span tests the subjects imitate the stimulus sequences presented by the experimenter. Span ability can therefore be conceived as the ability to imitate verbal stimulus sequences. The next experiment tested that theoretical conception by measuring the correlation between span ability and the ability to imitate sentences. The study also examined the effect of a digit load condition on sentence imitation.

The subjects were 22 mildly and moderately retarded adults who imitated sentences with and without a digit memory load. Each subject served in both conditions. In the load condition each subject first heard from one to three digits, according to the size of his span; then he heard a sentence. He then attempted to imitate the sentence, then the digits. Sentence spans were measured with a staircase procedure.

Sentence spans ranged from 3.1 to 16.1 words and correlated .82 with span ability. The memory load condition reduced the size of sentence spans from 7.76 to 6.66. The results lend construct validity to span as the ability to imitate verbal sequences and seem to suggest that span ability underlies both imitation and short-term recall since the load digits reduced sentence spans.

The concept of span as verbal imitation, while useful, is not generally applicable, even to memory span tests. If stimuli are presented auditorially subjects do imitate them; but if the stimuli are presented visually, it is difficult to describe spoken or written responses as imitation. This observation led to a concept of span as elicitation span in which the stimuli of the stimulus strings are viewed as eliciting each response in the response string. This concept, too, has been validated by research.

The subjects were 20 undergraduates who briefly studied one of two S-R pairs which were of either high association or low association. The high association pairs were car-truck and fox-hound, the low association pairs were car-wheat and fox-chair. The subjects were then tested for these associations in a staircase procedure in which they heard strings of the two words, car and fox, and responded with response sequences of the newly learned associations.

The subjects performed in a typical staircase span fashion. Their span scores quickly reached plateaus at about the levels expected from their standard staircase span scores, although the spans for the low association pairs stayed below the spans for the high association pairs. The experiment provides construct validity for the elicitation span concept because the stimuli functioned as conditioned elicitors of the responses in the response strings.

.

The last experiment illustrates my attempt to use the staircase span task as a generalized research tool. For this purpose it has much to offer. The task can be used directly with virtually all subjects, it correlates with intelligence, there is a clear developmental effect in the span task, individual differences are clearly revealed, and large amounts of variance due to levels of span ability can usually be extracted from the data so as to increase the precision of measurement. The paired associates span task described above is one example of this type of research in which the span task was modified to study paired associates learning.

The staircase span task has also been used to measure attention to irrelevant form cues. The subjects were 10 mildly retarded adolescents. The materials were nine color samples which were either circles or silhouettes of nine familiar objects such as a tree, a shirt, and a car. Each subject was able to name the silhouettes. Each subject performed twice in each of two conditions, a constant irrelevant form cue condition and a variable irrelevant form cue condition. In the former, the color samples were all circles; in the latter, the color samples were the silhouettes with a different form for each color. The standard staircase procedure was used to measure span in each condition. The subject viewed sequences of color samples and produced sequences of color names.

The color spans were smaller under the variable irrelevant form cue condition, and there was no interaction of conditions and subjects.

It appears that this task will be useful in assessing the degree of competition for attention among cues or dimensions, and therefore might be useful as a test of the effectiveness of attention training procedures. The test also offers a convenient way to study cued shift of attention. For example, the subjects might be tested with the identical procedures under instructions to respond with color names or to respond with form names. The ability to shift attention from cue to cue or dimension to dimension under verbal instruction is relevant in classroom tasks where the teacher often tells the students to attend to the specific stimuli or dimensions relevant for the learning task at hand.

References

Ellis, N. R. (1963). The stimulus trace and behavioral inadequacy. In N. R. Ellis (Ed.), *Handbook of mental deficiency*. New York: McGraw-Hill.

Hilgard, E. R. (1951). Methods and procedures in the study of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley.

Humpstone, H. J. (1918-1919). Memory span tests. *The Psychological Clinic: A Journal of Orthogenics for the Normal Development of Every Child, XII,* 196-200.

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.

Leaming, R. E. (1922). Tests and norms for vocational guidance at the 15-year-old performance level. *The Psychological Clinic: A Journal of Orthogenics for the Normal Development of Every Child, 14,* 193-217.

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

Murdock, B. B. (1974). *Human memory: Theory and data.* Potomac, MD: Erlbaum.

Staats, A. W. (1968). *Learning, language, and cognition.* New York: Holt, Rinehart and Winston.