Gareth Stack - Lab Group 2

Date of practical - 03/11/03 / Date of Submission - 28/11/03

Measurement of forward and backward digit span recollection, a correlate of I.Q

ABSTRACT

This experiment recorded forward and backward digit span recollection, a measure of short term memory which has been demonstrated to correlate with I.Q (Jensen, 2000). Thirty undergraduate students, 9 male and 21 female of mixed age, were tested individually with 2 sets of 16 nominal strings, ranging from 3 to 8 digits in length. Participants repeated the first set of strings in the order they had heard them, and reversed the order of the second set. Their scores on each set of strings were measured and collated. This experimenter recorded a score of 26. The group mean was calculated to be 18.73. Using kurtosis and skew analysis, the normality of the groups distribution was confirmed and the groups SD calculated to be 4.60. A z score of 1.58 was calculated by this experimenter. This was transformed to an I.Q o 123.73. Additional data more precisely correlating I.Q with digit span, and a wider sample of digit span SD and mean's would be required to infer the significance of the data recorded.

TABLE OF CONTENTS

INTRODUCTION

The Victorian Francis Galton developed the inferential statistical techniques of regression analysis and correlation, and his belief that ability was to a large extent innate and hereditary led him to popularise the theory that a general factor of intelligence existed, which was at least partially heritable (Tredoux, 2003). Elaborating on Galton's statistical techniques Charles Spearman (1904) first developed a test for a (g) or general factor underlying intelligence. Spearman discovered that the degree of uncertainty in the correlation of two variables could be calculated and used to more accurately infer the relationship between them. Spearman used this formula (r = r / var1 * var2) to infer that g was an existent measurable quality, rather than a summation or approximation of abilities or aptitudes (Spearman, 1904, as cited in Plucker 2003). More sophisticated tests such as the Binet-Simon, Stanford-Binet, and Wechsler-Bellevue tests (Wechsler Intelligence Scale for Children, and Wechsler Adult Intelligence Scale) have been developed to test as wide a collection of capacities as necessary in order to summarise g, or Intelligence Quotient (I.Q) (Plucker, 2003).

Digit Span Tests, which assess the ability of a participant to recollect a string of aurally presented nominal digits, are a component of David Wechers WISC (1949) and WAIS (1951) IQ tests (Flynn, 1984). They are designed to measure the capacity of participants short term memory store. Arthur Jensen (1971) discovered a high correlation between backward digit span tests, in which participants must recite digits read to them in reverse order, and g or general intelligence, as measured on more complete IQ tests (Jensen 1971, as cited in Jensen 2000).

There are three prominent explanations of how information disappears from short term memory. D.E Broadbent suggested that digits decay in STM by rehearsal through a series of selective filters (Conrad & Hille, 1958). In contrast Atkinson and Shiffrin theorised a limited capacity (7+-2 chunk) short term memory from which units of information could be displaced (Atkinson & Shiffren, 1968, cited in Passer & Smith 2001). Finally, suppression through interference of the phonological loop (proposed as a component of

Baddeleys & Hitch’s model of working memory, 1974), may inhibit short term memory and reduce digit span specifically (Passer & Smith, 2001).

This experiment attempted to measure forward and backward digit span as a measure of short term memory, which has been demonstrated to correlate with I.Q (Jensen, 2000).

METHOD

Design

Participants were tested in a single trial on a one to one basis. Each participant was read two sequences of 14 strings of numbers. Each sequence contained 7 pairs of strings of each ordinal length from 3 to 9 digits. For the initial sequence participants were awarded 1 point for each string correctly repeated. On hearing each string of the second sequence, participants were awarded 1 point for each string correctly repeated in reverse order. No digit string contained repeating numbers

Participants

Nine male and twenty one female undergraduate psychology students of University of Dublin, Trinity College, participated in the experiment as part of a course in 'Practicals Methodology and Statistics. Each participant constructed their own nominal strings as directed by the handout and was tested by a fellow student, whom they then tested. Although a more specific breakdown is unavailable, participants were of Irish, British, and Romanian nationality; and included both mature and post secondary undergraduates.

Procedure

The experiment conducted was a measurement study. Participants were presented with a series of strings of increasing length, from 3 to 9 digits and asked to repeat each string immediately after it was presented. Participants were instructed to listen carefully to the numbers presented, which were read out by the experimenter at the rate of one per second. Participants were initially asked to repeat verbatim the first set of strings presented, directly after hearing each one. Subsequently participants were requested to invert a second set of strings, repeating each back to front after it had been read. The number of strings correctly recollected on each of the two tests was recorded. The test being halted either when the participant had failed to recollect two strings of any one length, or had reached the end of 7 pairs.

RESULTS

Each student collated their experimental scores and reported them. From these the class mean score was calculated to be 18.73. This experimenters result was 26, with a score of 13 each on the forward and backward digit span tests.

The kurtosis and skew of the class results were assessed in SPSS, and found to be within +-2, indicating a normal distribution. Hence a standard deviation of 4.60 and a standard (z) score of 1.58 could be calculated.

This experimenters Z score was transformed using a new standard deviation of 15 and a new mean of 100 (the method used to calculate I.Q on tests such as the WAIS), to produce an I.Q of 123.73. It should be noted that this does not represent a calculation of g, but a single subsection of many standardised I.Q tests, which has been found to correlate with I.Q (Jensen, 2000).

Class and Experimenters Statistics on Digit Span Test

N of participants	30
Mean score	18.73333
Standard Deviation	4.593349
Kurtosis	-1.29706
Skew	0.183772
Experimental Score	26
Experimental z Score	1.581998
Transformed z-score	123.73

DISCUSSION

During this experiment several confounding variables potentially relevant to the study of intelligence arose. The importance of a controlled environment and a comprehensively described and uniformly conducted testing procedure was demonstrated by the wide variation of scores throughout the class. Conrad & Hillie (1958) noted that differences in the speed that digits are read out by experimenters on a digit span test can greatly influence the number of digits recollected, with a quicker reading pace leading to greater digit span recollection. Additionally, in a classroom environment the distractive potential of 15 experiments being conducted simultaneously is potentially large. Although the test is in part a measure of attention, relative differences in concentrative ability may potentially have been exaggerated by variations in sensitivity to noise.

The class results obtained in this digit span experiment were normally distributed, making it possible to produce standard (z) scores, which could have been compared to additional subtests to produce an overall measurement of intelligence (I.Q). This was done by assessing the group mean, and the degree to which scores varied over the group, known as standard deviation (SD).

By subtracting the mean of the distribution of norm scores from a participants score the amount by which the participant deviated from the groups average was estimated; indicating the amount by which the participant scored above or below the average. By then dividing this number by the standard deviation of the distribution of norm scores (the average amount by which the group deviated from the mean), it was possible to determine to what extent an individual participant's score represented a departure from the scores of the group.

Thus subtests could have been compared directly, as the z score results for each subtest would have measured how well each participant did relative to the group, rather than absolutely; producing a score adjusted for the relative difficulty of each test.

Because each z score was adjusted by the mean and standard deviation of the group tested, recording the mean and SD of differing age groups separately would allow a direct comparison of z scores between participants of varying age.

This experiment was carried out on a selection of Psychology students, pre-selected by exam success and psychometric testing (in the case of mature students). If the scores recorded were compared with a wider measure more representative of the population in general, the mean would likely be lower and the standard deviation higher, as a result of the greater amount of variation within society. The effect of this upon the z-score is impossible to quantify without exposure to a larger set of sample data.

It is important to note that many factors other than innate intelligence may have played a role in the test scores attained. This experimenter used the same set of test strings for both forward and backward digit span tests, creating a possible carryover effect. The participant tested by this experimenter was not a native English speaker, necessitating a mental translation of each number back and forth to carry out the memorisation task. Finally the limit of a sequence of non-repeating numbers in the Arabic system to 9 digits set an arbitrary limit on the highest digit span achievable.

While psychometrics retain a central role in psychological research, and play an ever increasing role in hiring procedures, scholastic streaming and selection, as well as educational impairment and giftedness testing, the I.Q paradigm is not unchallenged as a measure of intelligence. Two central criticisms seek to redefine I.Q as a limited depiction of human cognitive ability, literally "what intelligence tests measure" (Boring, 1923 quoted in Passer & Smith, 2001). The first criticism, that underlying intelligence is not a single generalisable factor but a gestalt of uncorrelated aptitudes and abilities, has existed as long as Spearman’s vague definition of task specific (s) intelligence. L.L.Thurstone's (amongst others) multiple factor intelligence theory broke conceptual intelligence into 7 specific uncorrelated abilities, from spatial visualisation to perceptual speed (Plucker, 2003). The second criticism of a specific g factor underlying all intelligence comes from Howard Gardiner’s theory of multiple intelligences (1993), which questioned whether a single quality called intelligence exists, or is a socially constructed summation of cognitive faculties which our society finds socially desirable. Gardiner has proposed 8 uncorrelated intelligences, many of which are un-testable by traditional methods (e.g.: musical intelligence, naturalist intelligence, kinaesthetic intelligence).

REFERENCES

Binet, A., (1905) New Methods for the Diagnosis of the Intellectual Level of Subnormals. Retrieved 19 Nov, 2003 from http://psychclassics.yorku.ca/Binet/binet1.htm

Conrad, R., & Hille, B. A., (1958). The decay theory of immediate memory and paced recall. Canadian Journal of Psychology, 12, 1-6.

Flynn, J. R., (1984) The mean IQ of Americans: Massive gains 1932 to 1978. Psychological Bulletin, 95, 29-51.

Gardner, H. E., (1993) Multiple Intelligences: The theory in Practice. New York: Basic Books.

Jensen, A. R., (2000) A "Simplest Cases" Approach to Exploring the Neural Basis of G, Psycoloquy, 11, 36-

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

Passer, M.W., Smith, R.E., (2001). Psychology: Frontiers and applications. McGraw-Hill Education

Plucker, J. A., (Ed.). (2003). Human intelligence: Historical influences, current controversies, teaching resources. Retrieved 28 Nov, 2003 from http://www.indiana.edu/~intell

Tredoux, G., (2003) Galton.org. Retrieved 21 Nov, 2003 from http://www.mugu.com/galton/index.html

APPENDIX

This experimenters results on Digit Span test

	Digit Forwards		Digit Backwards	
String Length	Test Digits	Response	Test Digits	
3 digits	729	correct	729	correct
3 digits	517	correct	517	correct
4 digits	8931	correct	8931	correct
4 digits	4623	correct	4623	correct
5 digits	71826	correct	71826	correct
5 digits	54237	correct	54237	correct
6 digits	925748	correct	925748	correct
6 digits	132495	correct	132495	correct
7 digits	4681732	correct	4681732	correct
7 digits	1596834	correct	1596834	correct
8 digits	49261783	correct	49261783	correct
8 digits	31627948	correct	31627948	correct
9 digits	501386724	incorrect	501386724	incorrect
9 digits	305174628	correct	305174628	correct

Digit Span Class Results	
14	18
22	14
23	12
19	18
24	23
17	16
22	26
25	26
26	15
20	14
23	15
12	18
16	16
18	24
14	12