Behavior Genetics, Vol. 9, No. 5, 1979

A Note on the Heritability of Memory Span

Arthur R. Jensen[1] **and Daniel Q. Marisi**[2]

Received 26 Apr. 1979—Final 25 Jun. 1979

The heritability of performance on a digit span memory test was estimated by means of the twin method, using the correlations of teenage monozygotic and dizygotic twins. The broad heritability is estimated as $h^2 = 0.44$*, which when corrected for attenuation is increased to 0.52. The values of* h^2 *determined by the same method for height and weight are 0.96 and 0.73, respectively. The MZ data show no evidence of a first-order genotype × environment interaction.*

KEY WORDS: memory span; digit span; heritability; MZ and DZ twins; height; weight.

INTRODUCTION

By means of the twin method, we have attempted to estimate the broad heritability of scores on a memory span test which has been frequently used as a measure of level I ability in Jensen's research on the level I–level II theory of mental abilities (e.g., Jensen, 1973, 1974). Level I ability is essentially the capacity to receive or register stimuli, to store them, and to later recognize or recall the material with high fidelity. It is characterized by the lack of any need of elaboration, transformation, or manipulation of the input in order to arrive at the output, as contrasted with level II ability, which is characterized by transformation and mental manipulation of the stimulus input prior to making the response. Digit span memory, particularly *forward* digit span, exemplifies level I ability (Jensen and Figueroa, 1975). To provide a more reliable measure of digit memory than is provided by the very brief digit span subtests of the Stanford-Binet and Wechsler intelligence tests, Jensen and Rohwer (1970) devised the Memory for Num-

[1] Institute of Human Learning, University of California, Berkeley, California 94720.
[2] Department of Physical Education, McGill University, Montreal, Quebec, Canada.

0001-8244/79/0900-0379$03.00/0

bers test. It is an auditory forward digit span test (administered by tape recorder) which has high reliability and broader factorial validity in the level I–level II system than the traditional memory span test. The Memory for Numbers test gains this advantage by means of repeated measurements and by varying procedural factors in order to reduce task-specific variance in the total scores.

Previous estimates of the heritability of digit span and short-term memory have been summarized by Pezzullo *et al.* (1972). In general, the outcomes of these older studies are virtually uninterpretable, because of the meager or unknown reliabilities of the memory tests, the small samples of twins used, and the theoretically inadequate methods of heritability estimation from measurements on monozygotic (MZ) and dizygotic (DZ) twins. Pezzullo *et al.* (1972) devised a tape-recorded memory span test with a split-half reliability above 0.90 and administered it to 37 pairs of MZ and 28 pairs of DZ twins, of ages 10–15 years. (Pezzullo *et al.* mistakenly reported their *N*'s as 28 MZ pairs and 37 DZ pairs.) They unfortunately estimated the heritability by means of Holzinger's (1929) formula, which has been shown to be conceptually incorrect (Jensen, 1967, 1976) since it does not correspond theoretically to what is meant by heritability in quantitative genetics (Jensen, 1975). Moreover, Pezullo *et al.* used twins ranging in age from 10 to 15 years and did not partial out the effect of age on the twin correlations. Since memory span is age dependent, the correlation between twins could therefore be significantly inflated by the common variable of age. Since the article by Pezzullo *et al.* does not contain the information required to obtain a proper estimate of heritability, Dr. Pezzullo has kindly provided us with the original data of their study so that we could apply the very same analysis that was used in our study, thereby permitting a direct comparison of their results with ours.

METHOD

Subjects

A questionnaire was sent to a number of junior and senior high schools in the San Francisco area to identify all the like-sexed twins enrolled and to request that they volunteer as subjects in this study. A total of 35 MZ twin pairs and 35 like-sexed DZ twin pairs volunteered. Their ages ranged from 11 to 18 years (MZ mean 14 years 7 months, SD 1 year 6 months; DZ mean 15 years 5 months, SD 1 year 6 months).

Zygosity Diagnosis

Twins were classified as MZ or DZ on the basis of the specially devised questionnaire and zygosity diagnosis procedure described by Nichols and

Bilbro (1966), which takes account of blood type and various physical characteristics. In addition, all twins were given the phenylthiocarbamide (PTC) taste test. About 40% of the population cannot taste PTC in the concentration used in this study, and any taste discrepancy between twins quite definitely rules out monozygosity. Errors in zygosity classification, of course, will diminish the estimated heritability.

Memory for Numbers Test

The Memory for Numbers test takes about 30 min to administer. Series of from four to nine simple random numbers (exclusing zero, repetitions, and normal numerical order) are presented by a tape-recorded male voice at the rate of one digit per second. The sound of a "bong" indicates the start and the termination of each series. Subjects are allowed 13 sec between series to write as many of the digits as they can recall in the correct order on a specially prepared answer sheet. There is one series at each length of from four to nine digits, in each of three experimental conditions: (1) immediate recall, in which subjects write their recall immediately after the "bong" which comes 1 sec after the last digit of each series; (2) repeated series, in which subjects hear the same series repeated three times (separated by a "click") before writing the digits; and (3) delayed recall, in which subjects must delay their recall until 10 sec have elapsed after the last digit in the series, the end of the 10-sec interval being signaled by a "bong." To ensure that all subjects understand the instructions and procedure, each experimental procedure is preceded by a short practice test consisting of three different digit series each of three digits. In each case the procedure for the practice test is identical to the particular experimental condition that follows it. The subject's score on the Memory for Numbers test is the total number of digits recalled in the correct position, with a maximum possible score of 117.

All subjects were tested by the second author as part of a larger study (Marisi, 1972).

Heritability Estimation

Broad heritability was estimated by using the simplest heritability formula which is commonly used in quantitative genetics (e.g., Falconer, 1960). It is $h^2 = 2(r_{MZ} - r_{DZ})$, where h^2 is the estimated broad heritability, and r_{MZ} and r_{DZ} are the intraclass correlations for MZ and DZ twins. The rationale, assumptions, and limitations of this method of estimating heritability have been explicated in detail elsewhere (Jensen, 1976). It is a conservative estimate of heritability in that it assumes random mating for the trait in question, so the formula underestimates the broad heritability when there is in fact assortative mating for the trait. However, since we have no precise information on assortative mating for memory span, it seems preferable to assume random mating and obtain a lower-bound estimate of the broad heritability. If there is dominance deviation for the trait, the above formula will slightly overestimate the broad heritability. There is no knowledge of dominance deviation for memory span. Thus the broad heritability will be underestimated by the above formula to the extent that there is assortative mating and slightly overestimated to the extent that there is dominance. These biases, if both exist, tend to cancel each other to some extent.

An estimate of the standard error of h^2, derived from the formula given by Loehlin *et al.* (1975, p. 288).

$$SE_{h^2} = 2\{[(1 - r_{MZ}^2)^2/N_{MZ}] + [(1 - r_{DZ}^2)^2/N_{DZ}]\}^{1/2}$$

Since the raw memory span scores are age dependent, it is essential to statistically control the effect of age in calculating the intraclass correlation between twins. This can be done by analysis of covariance, in which the total variance is partitioned into four components associated with (1) the regression of differences between twin pairs on age, (2) the regression of differences within twin pairs on age, (3) the residual differences between pairs, and (4) the residual differences within pairs.

The rationale of this analysis by means of which age is partialed out of the twin intraclass correlation is as follows. The aim is (1) to remove from the variance among twin pair means (i.e., the variance between pairs) that portion that is attributable to the correlation between (a) age and (b) the twin pair means, and (2) to remove from the variance within twin pairs that portion attributable to the correlation between (a) age and (c) the twin differences. Following a simple one-way analysis of variance, the between-pairs sum of squares (SS_B) can be analyzed into two parts: a part due to regression of pair means on age (which is $r_{ma}^2 \times SS_B$, where r_{ma} is the Pearson correlation between pair means and age) and the residual, i.e., $SS_B(1 - r_{ma}^2)$. Similarly, following a simple one-way analysis of variance, the within-pairs sum of squares (SS_W) can be analyzed into two parts: a part due to regression of twin differences on age (which is $r_{da}^2 \times SS_W$, where r_{da} is the Pearson correlation between twin absolute differences and age) and the residual, i.e., $SS_W(1 - r_{da}^2)$. In each instance the regression uses up 1 df. In no case, with the present sample sizes, did a nonlinear component of the regression of between-pairs and within-pairs effects on age even approach statistical significance; so only the linear regression of effects on age is partitioned in the analysis of covariance, from which the twin intraclass correlations are derived.

The age-corrected twin intraclass correlation (r_i) is calculated from the residual mean squares for between pairs (MS_B) and within pairs (MS_W), as follows: $r_i = (MS_B - MS_W)/(MS_B + MS_W)$. The correlation is significantly greater than zero if $(MS_B/MS_W) = F$ is significant for the given degrees of freedom. The r_i thus obtained for a sample of twin pairs of differing ages approximates the expected values of r_i for a sample of twin pairs all of the same age.

When a particular method of heritability analysis is applied for the first time to a variable about whose genetics little or nothing is known, it seems wise to apply exactly the same analysis to physical variables measured in the same twin samples, because we already have considerable knowledge about the heritability of these measures. Heritability estimates of the physical measurements, as R. A. Fisher (1918) was the first to suggest, can serve as a "sheet anchor" for comparison of the results of applying the same method of analysis to a little known trait—in the present case, memory span. Therefore, height and weight measurements were secured (which, like memory span, are age dependent) on all the twins, and we have applied parallel analyses to these physical measurements.

RESULTS AND DISCUSSION

Memory for Numbers Test

The results of the analysis of covariance of the Memory for Numbers scores are shown in Table I.

The MS of 13.33 for the regression of MZ twin pair means on age is anomalously very much smaller than the corresponding MS of 274.65 for DZ twins, because (1) the total variance of the DZ twin scores is 1½ times larger than the total variance of MZ twin scores and (2) the correlation between age and twin pair means is greater in the DZ sample than in the MZ sample (+0.16 vs. −0.04). These differences between the MZ and DZ twins have no obvious explanation and are most likely sampling fluctuations.

Table I. Analysis of Covariance of Memory for Numbers Test

	MZ twins			DZ twins		
Source of variance	df	MS	F	df	MS	F
Between pairs (residual)	33	231.94	6.04[a]	33	303.78	2.98[a]
Within pairs (residual)	34	38.40		34	101.99	
Regression on age (between)	1	13.33	<1	1	274.65	2.69
Regression on age (within)	1	337.86	8.80[a]	1	244.41	2.40

[a] $p < 0.01$.

The age-corrected twin intraclass correlations are $r_{MZ} = 0.716$ and $r_{DZ} = 0.497$. A conventional test of the significance of a genetic effect is $F = MS_{WDZ}/MS_{WMZ}$, i.e., the ratio of the within-pairs variance for DZ and MZ twins. In this case, $F = 2.65$, which with 33 and 34 df is significant beyond the 0.01 level. The h^2 is 0.437, SE = 0.303 ($t = 1.44$, one-tailed $p < 0.10$).

Since the internal consistency reliability of the test scores can be determined using the Hoyt (1941) method, we can correct the obtained h^2 for attenuation. The reliability of the memory scores is 0.84. The corrected h^2 is 0.437/0.84 = 0.52.

Comparison with Results of Pezzullo *et al.*

It is interesting to compare these results with those obtained from a parallel analysis of the twin data on a test of memory span used in the study by Pezzullo *et al.* (1972), shown in Table II. The twin correlations are $r_{MZ} = 0.764$ and $r_{DZ} = 0.414$. The ratio of within-pairs variance for DZ and MZ twins is $F = 2.37$, $p < 0.02$. The $h^2 = 0.700$, SE = 0.342 ($t = 2.05$, one-tailed $p < 0.03$). This value of h^2 is not significantly different ($t < 1$) from the $h^2 = 0.437$ obtained in our study.

Raven's Progressive Matrices

Pezzullo *et al.* administered Raven's Standard Progressive Matrices test, a highly *g*-loaded test of abstract reasoning or level II ability, to the same twin samples. Since they failed to take account of age in their analysis, their data were reanalyzed by the method that was applied to the memory scores. The results are shown in Table III. The twin correlations are $r_{MZ} = 0.817$ and $r_{DZ} = 0.369$. The F ratio for the DZ/MZ within-pair variances is 5.91, $p < 0.01$. The $h^2 = 0.897$, SE = 0.344 ($t = 2.60$, one-

Table II. Analysis of Covariance of Memory Span Scores from Study by Pezzullo *et al.* (1972)

	MZ twins			DZ twins		
Source of variance	df	MS	F	df	MS	F
Between pairs (residual)	35	1451.63	7.47[a]	26	1110.85	2.41[b]
Within pairs (residual)	36	194.43		27	460.58	
Regression on age (between)	1	26835.14	138.02[a]	1	5904.86	12.82[a]
Regression on age (within)	1	818.68	4.21[b]	1	22.43	<1

[a] $p < 0.01$.
[b] $p < 0.05$.

Table III. Analysis of Covariance of Raven's Progressive Matrices Scores from Study by Pezzullo *et al.* (1972)

	MZ twins			DZ twims		
Source of variance	df	MS	F	df	MS	F
Between pairs (residual)	35	71.55	9.92[a]	26	92.29	2.17[b]
Within pairs (residual)	36	7.21		27	42.58	
Regression on age (between)	1	916.48	127.11[a]	1	1520.51	35.71[a]
Regression on age (within)	1	0.57	<1	1	199.93	4.69[b]

[a] $p < 0.01$.
[b] $p < 0.05$.

tailed $p < 0.01$). This is surely one of the highest heritabilities for a mental test ever reported in the literature.

Height and Weight

Tables IV and V show the same analysis applied to the height and weight measurements obtained in the present twin samples.

For height the twin correlations are $r_{MZ} = 0.973$ and $r_{DZ} = 0.494$. The F ratio for the DZ/MZ within-pair variance is 32.39, $p < 0.01$. The $h^2 = 0.958$, SE $= 0.256$ ($t = 3.74$, one-tailed $p < 0.01$).

For weight the twin correlations are $r_{MZ} = 0.976$ and $r_{DZ} = 0.611$. The F ratio for the DZ/MZ within-pair variances is 22.96, $p < 0.01$. The $h^2 = 0.730$, SE $= 0.212$ ($t = 3.44$, one-tailed $p < 0.01$).

Thus the broad heritabilities of height and weight, as estimated by the present method, are quite in accord with expectations (*cf.* Fisher, 1918; Bulmer, 1970).

Table IV. Analysis of Covariance for Height (Inches)

	MZ twins			DZ twins		
Source of variation	df	MS	F	df	MS	F
Between pairs (residual)	33	20.65	73.75[a]	33	26.81	2.95[a]
Within pairs (residual)	34	0.28		34	9.07	
Regression on age (between)	1	136.50	487.50[a]	1	802.05	88.43[a]
Regression on age (within)	1	1.46	5.21[b]	1	6.73	<1

[a] $p < 0.01$.
[b] $p < 0.05$.

Table V. Analysis of Covariance of Weight (Pounds)

Source of variance	MZ twins df	MS	F	DZ twins df	MS	F
Between pairs (residual)	33	678.51	83.46[a]	33	773.67	4.14[a]
Within pairs (residual)	34	8.13		34	186.68	
Regression on age (between)	1	18720.10	2303.59[a]	1	16190.41	
Regression on age (within)	1	64.44	7.93[a]	1	1087.46	

[a] $p < 0.01$.

Genotype × Environment Interaction

Jinks and Fulker (1970) have proposed a method, based on MZ twins, for determining one kind of genotype × environment interaction, that is, the extent to which environmental effects depend on genotypes. The method detects only first-order interaction, that is, interaction which arises from linear regression of environmental effects on genotypic values. The method is more ideally applied to MZ twins who are reared apart in uncorrelated environments. When applied to MZ twins who are reared together, as in the present twin data, the method detects only the linear regression of within-family microenvironmental effects on the composite of genotype plus between-families environmental effects. Specifically, the method consists of computing the correlation r_{ab} between two sets of variables: a, the sum of the scores for each pair of MZ twins, and b, the absolute difference in scores for each pair of MZ twins. This correlation (with the effect of age partialed out) was obtained for each of the sets of MZ twins shown in Tables I–IV. None of the correlations is larger than its standard error, except for the Pezzullo *et al.* memory scores, for which the correlation (with age partialed out) between MZ sums and differences is -0.35, $p < 0.05$.

DISCUSSION

The estimated broad heritability of 0.437 of the memory test scores in the present study is significantly ($p < 0.10$, one-tailed) greater than zero but is smaller than the h^2 of 0.700 obtained for a similar memory test by Pezzullo *et al.* The difference between these heritabilities, however, does not even approach significance ($t < 1$) in view of the large standard errors of these heritability coefficients. The small sample sizes of these two studies are hardly ideal for making comparisons among heritability coefficients that are all greater than 0.40. In fact, although the several values of h^2 obtained here range from 0.437 for our memory test to 0.956 for height, none of the heritabilities differs significantly from any of the others even at the 10% level of confidence. Although it seems safe to conclude that the abilities measured by the present tests involve heritable factors, it is obvious that much larger samples, including other kinships besides twins, are required for the investigation of more highly specified genetic models.

REFERENCES

Bulmer, M. G. (1970). *The Biology of Twinning in Man*, Clarendon Press, Oxford.

Falconer, D. S. (1960). *An Introduction to Quantitative Genetics*, Ronald Press, New York.

Fisher, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Tr. R. Soc. Edinburgh* **52**:399–433.

Holzinger, K. J. (1929). The relative effect of nature and nurture influences on twin differences. *J. Educ. Psychol.* **20**:241–248.

Hoyt, C. J. (1941). Note on a simplified method of computing test reliability. *Educ. Psychol. Meas.* **1**:93–95.

Jensen, A. R. (1967). Estimation of the limits of heritability of traits by comparison of monozygotic and dizygotic twins. *Proc. Natl. Acad. Sci.* **58**:149–157.

Jensen, A. R. (1973). Level I and level II abilities in three ethnic groups. *Am. Educ. Res. J.* **10**:263–276.

Jensen, A. R. (1974). Interaction of level I and level II abilities with race and socioeconomic status. *J. Educ. Psychol.* **66**:99–111.

Jensen, A. R. (1975). The meaning of heritability in the behavioral science. *Educ. Psychol.* **11**:171–183.

Jensen, A. R. (1976). The problem of genotype-environment correlation in the estimation of heritability from monozygotic and dizygotic twins. *Acta Genet. Med. Gemellol.* **25**:86–99.

Jensen, A. R., and Figueroa, R. A. (1975). Forward and backward digit span interaction with race and IQ: Predictions from Jensen's theory. *J. Educ. Psychol.* **67**: 882–893.

Jensen, A. R., and Rohwer, W. D., Jr. (1970). *An Experimental Analysis of Learning Abilities in Culturally Disadvantaged Children*, Final Report, Office of Economic Opportunity, Contract No. OE02404.

Jinks, J. L., and Fulker, D. W. (1970). Comparison of the biometrical genetical, MAVA, and classical approaches to the analysis of human behavior. *Psychol. Bull.* **73**:311–349.

Loehlin, J. C., Lindzey, G., and Spuhler, J. N. (1975). *Race Differences in Intelligence*, Freeman, San Francisco.

Marisi, D. Q. (1972). The genetic component in motor performance and numeric memory. Unpublished doctoral dissertation, University of California, Berkeley.

Nichols, R. C., and Bilbro, W. C. (1966). The diagnosis of twin zygosity. *Acta Genet.* **16**:327–352.

Pezzullo, T. R., Thorsen, E. E., and Madaus, G. F. (1972). The heritability of Jensen's level I and II and divergent thinking. *Am. Educ. Res. J.* **9**:539–546.

Edited by John C. Loehlin