Journal of School Psychology, Vol. 31, pp. 385-405, 1993
Pergamon Press Ltd. Printed in the USA.
0022-4405/93/$6.00 + .00

Correlates of Some Curious Regressions on a Measure of Intelligence

Lloyd G. Humphreys
University of Illinois

David Lubinski
Iowa State University

Grace Yao
University of Illinois

In the course of a methodological investigation of regressions of means of selected tests, computed in stratified levels of a general intelligence composite, on the mean levels of intelligence within the strata we encountered curious phenomena at the low end of the distribution of intelligence. At about −2.00 *SD* units of that distribution, the regressions of the vocabulary means for both boys and girls turned abruptly upwards, while those for mechanical reasoning dropped sharply. To investigate these curious regressions, we formed high and low vocabulary groups for each sex within the low-scoring subsample of general intelligence, and the means of the four groups on a series of cognitive and self-report scales were obtained. In this sample the high school boys and girls who are high in vocabulary relative to their low scores on the intelligence composite have lower means than their low-low controls on a set of cognitive tests epitomized by visualization in three dimensions, but the set also includes verbal tests in which the tasks are ambiguous and require inference and hypothesis formation in order to obtain a solution. Correct answers are not directly present in long-term memory as they are for the academic (including mathematics) and nonacademic information tests, on which the two high vocabulary groups are substantially superior. The high and the low vocabulary groups also differ from each other in health histories, personality traits, vocational interests, and biographical data scores, the high groups being consistently closer to the norm for their high school class. Sex differences are minimal in the incidence of the deficit and in its correlates. A genetic explanation for the problem is plausible, but the locus is not on the 23rd pair of chromosomes.

In the course of computing correlations between distributions of sample means (Lubinski & Humphreys, submitted for publication) of selected cognitive tests from Project TALENT (Flanagan et al., 1962), we encountered some unusual regressions of narrower tests on a measure of general intelligence in the large national sample of 10th-grade high school students. The initial exam-

Editorial services for this contribution were carried out by Dr. Raymond S. Dean.
Received November 21, 1991; final revision received June 25, 1992.
Address correspondence and reprint requests to Lloyd G. Humphreys, Department of Psychology, 603 East Daniel, Champaign, IL 61820 or to David Lubinski, Department of Psychology, Iowa State University, Ames, IA 50011-3180.

ples that attracted our attention are shown in Figure 1. The sample size for each sex is between 40,000 and 45,000, producing smooth data.

The intelligence measure is a composite of reading comprehension, arithmetic reasoning, and abstract (figural) reasoning with effective weights of approximately .50, .25, and .25, respectively (Wise, McLaughlin, & Steel, 1979). The Vocabulary test is a measure of general, as opposed to specialized, knowledge of the meaning of words. The Mechanical Reasoning test is a pictorial test of the operation of mechanical principles. It has been called a

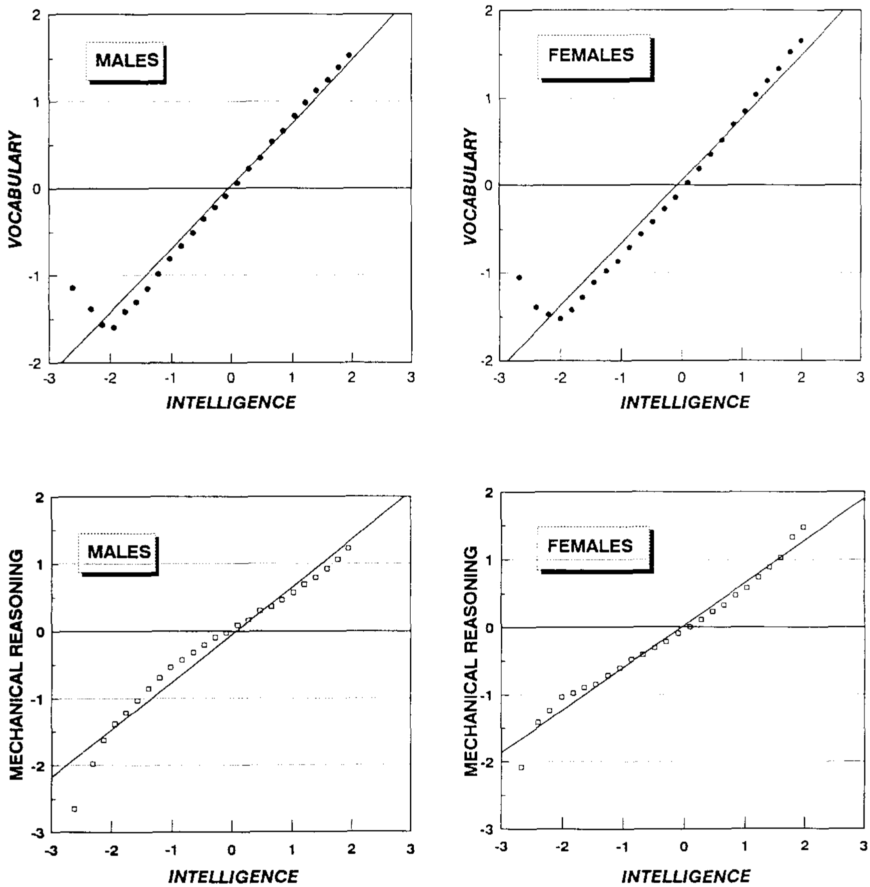

Figure 1. These scatter plots were based on standardized bivariate means. For each plot, the Intelligence Composite was systematically parsed above and below the mean, for each gender, by approximately 0.20 standard deviation units (raw score means and standard deviations being 158.19 and 53.65 for the males, 158.55 and 52.22 for the females). The *Y* means were computed at each interval on the *X* axis for Mechanical Reasoning and Vocabulary. Total sample sizes: females 43,165; males, 43,629.

measure of "barnyard" physics. The two tests are very different in content and in the operations required of examinees. The anomalies in their regressions are also very different and are highly significant statistically. The area of the vocabulary anomaly in the scatter plot represents a negative correlation with a typical measure of intelligence. This is surprising, because knowledge of words has long been used in intelligence tests.

Our only theory at the outset was that psychological phenomena are lawful. The regressions depart little from linearity through 98% of the range of scores on the predictor, but break sharply from the straight line at about −2.00 *SD* units. This was entirely unexpected and aroused our curiosity. Serendipitous findings that are highly replicable are always interesting and at times may be quite important. A substantive a priori theory is not a requirement, and research can be justified post hoc by the data (see Skinner, 1950). The development of a theory requires as a first step a dependable, systematic description of the correlates of the phenomenon.

It is significant for the regression of vocabulary on the measure of intelligence that the students who are below the inflection of the regression line have variances in vocabulary that are three times those in the rest of the distribution of intelligence. This suggests that there may be two quite different groups represented in the lower 2% of the distribution of the composite. The lower portion contains persons who represent the extension of the linear regression, while the upper portion contains students who are not expected to be so far below the mean in *general* intelligence. In what follows, we shall look for clues concerning the source of the deficit of the aberrant cases on the intelligence composite.

PROCEDURE

Subjects

The subjects were obtained from the Project TALENT Data Bank (Flanagan et al., 1962; Wise et al., 1979). A nationwide probability sample of high school students, Grades 9–12, with approximately 100,000 students in each grade, were tested in 1960. TALENT is a well-known data bank containing a large number of cognitive and self-report measures (see below). Data for the present study were based primarily on the 10th-grade cohorts of boys and girls.

In the first step, we selected everyone who fell below the point of inflection of the vocabulary regression, which was approximately two standard deviations below the mean of the intelligence composite. It so happens that this is also the approximate point of inflection of the less dramatic regression of mechanical reasoning, also previously shown in Figure 1. This decision produced 743 10th-grade boys and 826 girls.

Defining the Groups

For purposes of dividing these groups into those who "belonged" low and those who did not, we arbitrarily set the dividing score for low and high vocabulary groups near the median. The cut was 8 and below for boys, 7 and below for girls. Mean standard scores for the high and low groups of each sex based on own-sex distributions in the full range of 10th-grade talent indicate the heterogeneity of the vocabulary scores. These means and the sample sizes on which they were based are, respectively, as follows: −0.63 and 329 for high boys, −0.73 and 397 for high girls, −2.03 and 414 for low boys, and −1.98 and 429 for low girls. Because sex differences in vocabulary are small, the raw score equivalents of these own-sex standard scores are also closely equivalent. Seemingly, the regression problem is not sex-linked.

Dependent Variables

With four groups defined, the next step involved obtaining means on other variables in the Project TALENT Data Bank. There are more than 50 cognitive tests, 10 personality and 17 interest self-report scores, and 18 scores developed from biographical data items that make these data attractive as a resource for descriptive purposes. We shall report practically all of these scores either individually or in composites such as English (5 components) and nonacademic information (22 components). The small number of cognitive tests omitted were redundant in outcome and characteristics with the ones reported.

RESULTS FROM COGNITIVE TEST DATA

Correlations Among the Defining Tests

The correlations in the full 10th-grade sample of Vocabulary and Mechanical Reasoning measures with each other, with the components of the intelligence composite, and with the composite itself represent background information relevant to the analyses that follow. These correlations are shown in Table 1. The boys' sample is below the diagonal, the girls' above.

There is a prominent general factor in the five experimentally independent measures, Reading Comprehension having the most in common with the rest. The weight given to reading in the composite seems appropriate. The correlation of vocabulary with reading is quite high, given the relatively modest reliabilities of the short tests administered in Project TALENT (Flanagan et al., 1964) in spite of the reversal at the low end of the intelligence distribution.

Differences in Means

As we make the numerous comparisons of boys and girls as well as the high and low vocabulary groups, we shall not report p values. All means will be in standard scores obtained in the own-sex distributions of all 10th-grade boys

Table 1
Intercorrelations[a] of Vocabulary, Mechanical Reasoning, and Intelligence Scores

	1	2	3	4	5	6	$\overline{X}$	S_x
1 Vocabulary		490	739	592	490	742	11.18	3.99
2 Mechanical Reasoning	513		511	486	526	590	8.32	3.46
3 Reading Comprehension[b]	713	481		635	581	936[c]	29.54	9.99
4 Arithmetic Reasoning[b]	572	479	605		515	796[c]	7.66	3.36
5 Abstract Reasoning[b]	459	537	517	484		780[c]	8.54	3.01
6 Intelligence	727	582	930[c]	784[c]	739[c]		161.98	50.48
$\overline{X}$	12.71	12.31	28.84	8.35	8.84	164.16		
S_x	3.74	4.00	10.45	3.42	2.95	50.78		

[a]Girls are above, boys below, the diagonal; decimal points are omitted.
[b]Components of the intelligence composite.
[c]Spuriously high part-whole correlations.

and girls. Standard deviations of the selected groups are less than 1.00 and the *N*s of each group are approximately 400. Thus the standard error of a mean is about 0.05 and the standard error of a difference between sexes within groups or between groups within sexes is about 0.07. The same guidelines apply to sex differences within groups, but interpretation of these differences requires recognition of the own-sex metric. All that a reader has to do is divide a given difference in means by 0.07 to obtain a basically accurate feel for the stability of our findings. A *t* accurate to two decimal places is not required.

Table 2 contains standard score means in the own-sex metric of the high and low groups for each sex for selected cognitive tests. Use of the own-sex metric allows boys and girls to be compared without assuming constancy of means from 1960 to the present. We do assume that relations among psychological measures are more robust to cultural change and cultural differences than mean levels of performance. Because the distributions of these tests are not normal, comparison of the absolute size of the standard score means from one test to another is not meaningful. Differences in standard scores between high and low groups are more interpretable, and the signs of the differences cannot be explained away by manipulation of the units of measurement.

Because the low and high vocabulary groups also had to be low on the intelligence composite, the expectation is for low scores for both groups on the components. The low vocabulary group has higher means than the high group on Reading Comprehension and Abstract Reasoning, while Arithmetic Reasoning shows a difference in the opposite direction. The weights of these tests in the intelligence composite determine the means of the composite for the vocabulary groups. Statistically significant differences in opposite directions almost precisely balance to produce approximately equal intelligence test means.

High group superiority is shown in academic and nonacademic information tests, and in English mechanics, Mechanical Information, two verbal memory tests, and two mathematics tests. All of these tests contain items that have straightforward answers with a minimum of ambiguity. Examinees merely have to know the answer. Minimal inference is required.

In contrast, the cognitive tests on which the low vocabulary groups are superior require inference and perception of relations among novel stimuli. The last three tests in the list contain figural stimuli. The three preceding are answered by words, but in the Disguised Words and Creativity tests the words have to be inferred or hypothesized and compared in order to answer choices that provide minimal cues. Word Functions in Sentences requires examinees to recognize similar functions served by designated words in different sentences that have no overlapping content or meaning.

A Possible Confound

There is complete consistency with respect to sign for the results for the two sexes, and numerical differences are small. Interpretation of the size of the

Table 2
Means and Differences in Means of the High and Low Vocabulary Groups on Various Cognitive Tests[a,b]

	Boys			Girls		
	High	Low	Diff.	High	Low	Diff.
Tests defining groups						
Vocabulary	−0.63	−2.03	1.40	−0.73	−1.98	1.25
Reading Comprehension	−2.18	−1.96	−0.22	−2.32	−2.13	−0.19
Abstract Reasoning	−2.47	−2.13	−0.34	−2.28	−2.01	−0.27
Arithmetic Reasoning	−0.94	−1.60	0.66	−0.99	−1.47	0.48
High group superiority						
Literature	−0.68	−1.52	0.84	−0.80	−1.54	0.74
Social Studies	−0.77	−1.71	0.94	−0.80	−1.58	0.78
Mathematics Information	−0.76	−1.10	0.34	−0.67	−0.96	0.29
Introductory Math	−0.79	−1.22	0.43	−0.86	−1.30	0.44
Physical Science	−0.71	−1.43	0.78	−0.64	−1.11	0.47
Mechanical Information	−0.55	−1.90	1.35	−0.64	−1.38	0.74
English	−1.07	−2.22	1.15	−1.13	−2.28	1.15
Nonacademic Information	−0.77	−1.84	1.07	−0.96	−1.99	1.03
Memory–Sentences	−0.44	−1.03	0.59	−0.57	−1.15	0.58
Memory–Words	−0.60	−0.92	0.32	−0.64	−1.16	0.52
Low group superiority						
Disguised Words	−1.55	−1.38	−0.17	−1.56	−1.48	−0.08
Word Functions	−1.32	−0.93	−0.39	−1.34	−1.09	−0.25
Creativity	−1.61	−1.32	−0.29	−1.54	−1.28	−0.26
Mechanical Reasoning	−2.14	−1.78	−0.36	−1.58	−1.26	−0.32
Visualization: 2-dimensions	−1.70	−1.22	−0.48	−1.40	−1.15	−0.25
Visualization: 3-dimensions	−1.87	−1.26	−0.61	−1.68	−1.14	−0.54

[a]Test scores were standardized in each sex separately.
[b]$S_{\bar{x}} \approx 0.05$; $S(\bar{x}_1 - \bar{x}_2) \approx 0.07$.

differences is ambiguous without confidence in the equality of the units of measurement, but there is another, more troubling sense in which there is also complete consistency. All of the tests with negative signs for differences were printed in the same test booklet administered in a half-day session of the two days of test administration. This possible confound was discovered just prior to formation of a confident conclusion, but there was a logical basis for the confound that blunts its importance at the outset. There were valid a priori reasons for assigning tests to booklets and half-day sessions. They were not placed at random.

Academic information tests were assigned to one booklet in which the many different items were scattered at random in a "test" administered under a single time limit. The nonacademic information tests were assigned to a second booklet, similarly arranged and administered. A third booklet contained two tests of immediate memory, numerical and mathematics tests, and the tests of English mechanics, each independently timed. Nothing in these three booklets required elaborate directions or numerous practice problems. Two of the three contained no graphics, while graphics were both simple and trivial in amount in the third. No test posing a problem in printing appeared in the first three booklets. A fourth booklet contained the tests that were most difficult to print and to administer. They had complex directions, multiple practice problems, and much graphical material. Some of the tests were entirely figural in content. However, outcome data to which we now turn will allow a more definitive evaluation of the confound hypothesis.

Speeded Tests

There were also three highly speeded tests, all composed of relatively simple items, that were included in the same test booklet with the "low group superiority" tests. A fourth speeded test, a measure of speed of numerical operations, appeared in the booklet with the mathematics tests. None of these speeded tests were included in the initial analyses because the problem appeared to be highly cognitive, and these tests were known to have small correlations with tests highly loaded on the general factor of intelligence.

The mean correlations in Table 3 show that there is indeed little general intelligence variance in the four speeded tests. Based on the data in Table 2, this table substitutes Three-Dimensional Spatial Visualization for Mechanical Comprehension as the key test for the high vocabulary groups. The score on the four tests that penalized wrong answers has the highest mean correlation with the two cognitive tests, while number of attempts is almost completely independent. Object Identification has somewhat higher correlations with Visualization than with Vocabulary, while Numerical Operations has the reverse pattern. All correlations, however, are small. These speeded tests measure primarily other characteristics of examinees than general intelligence.

Table 4 contains the relations of each of the three scores on the four tests

Table 3
Mean Correlations[a] of the Speeded Tests and Types of Scores With Vocabulary and Spatial Visualization

	Tests				Scores		
	Numbers	Tables	Names	Objects	Rights	Attempts	Formulas
Vocabulary							
Girls	33	11	07	13	16	08	22
Boys	30	08	02	08	14	00	22
Visualization							
Girls	12	12	08	22	16	10	22
Boys	17	00	05	19	14	02	21

[a]Decimal points are omitted.

Table 4
Means of Three Different Scores on Four Speeded Tests for High and Low Boys and Girls[a]

	Boys			Girls		
	High	Low	Diff.	High	Low	Diff.
No. right						
Numerical Operations	−0.85	−1.64	0.79	−1.01	−1.78	0.77
Table Reading	−1.00	−0.76	−0.24	−1.17	−0.92	−0.25
Name Checking	−1.28	−0.54	−0.74	−1.09	−0.40	−0.69
Object Identification	−1.78	−0.88	−0.90	−1.78	−0.90	−0.88
No. of attempts						
Numerical Operations	−0.48	−0.96	0.48	−0.78	−1.28	0.50
Table Reading	−0.52	−0.02	−0.54	−0.58	−0.18	−0.40
Name Checking	−1.10	−0.26	−0.84	−1.00	−0.09	−0.91
Object Identification	−1.66	−0.48	−1.18	−1.63	−0.49	−1.14
Formula scores						
Numerical Operations	−0.75	−1.40	0.65	−0.93	−1.78	0.85
Table Reading	−0.81	−1.04	0.23	−1.28	−1.28	0.00
Name Checking	−1.06	−0.87	−0.19	−0.91	−0.95	0.04
Object Identification	−1.58	−1.10	−0.48	−1.74	−1.19	−0.55

[a] $S_{\bar{x}} \approx 0.05$; $S(\bar{x}_1 - \bar{x}_2) \approx 0.07$.

with the vocabulary groups. The Numerical Operations test shows differences favoring the high vocabulary group in the number of right answers scores that are of approximately the same size as those for Introductory Mathematics in Table 2, while Table Reading, Name Checking, and Object Identification show, in that order, increasing superiority of the low vocabulary groups.

The formula score that best measures the general factor of intelligence among these four groups shows the largest advantage of the high groups to lie in Numerical Operations and the smallest disadvantage in speed to lie in Object Identification. The other two tests are neutral with respect to group differences on the score that penalizes incorrect answers.

Number of attempted items represents the purest measure of speed. The differences between the high and the low vocabulary groups are of the same sign and in the same rank order as for the number of correct answers. Also, the advantage of the high group on Numerical Operations is smallest and their disadvantage on Object Identification is greatest.

The differences between the two vocabulary groups for Name Checking and Object Identification are greater for all three scores than for the cognitive tests in Table 2. This is the result primarily of the performance of the low vocabulary group on these tests, as they are only about as far below the mean of the norm group as one might expect from the low cognitive content of the tests. In contrast, the means of the high group remain at about the size of their means in Table 2.

An Attribute Is Added

If we tentatively accept the possibility that the results in Table 2 are not artifacts (by some unknown mechanism) of placement in separate test booklets, it becomes necessary to add another attribute to those that separated the high and low groups on the highly cognitive measures: The high group cannot perform rapidly when stimuli require fine perceptual discrimination, even when discriminations can be made with little error, given unlimited time. A deficit in immediate visual memory may be involved. We believe that, on a priori grounds, judges would select Object Identification as the speeded test most likely to show a deficit in the high vocabulary group similar to that of Three-Dimensional Visualization and that Numerical Operations would have the opposite pattern. The evidence seems quite compelling that more is involved than test booklet placement, but additional checks can be made on the confounding hypothesis.

Additional Checks on the Confounding

Did something unique occur during the 10th-grade administration that led to these results? No, the regressions are highly similar in the 12th-grade cohort of Project TALENT. Was careless test administration at a small number of sites responsible? No, the members of the high and low groups are widely scattered in terms of identification number, state of residence, rural-urban location, community size and ethnicity. The proportions in the high and low groups do differ significantly from one variable to another, but the differences are supported by other data. For example, the groups that are low in both vocabulary and intelligence are disproportionately Southern and black—a typical test outcome. Finally, did a small proportion of students in many different high schools in both the 10th and 12th grades, with approximately equal proportions of both sexes, simply decide not to work at capacity on one test booklet? This is highly doubtful. One would expect to find a sex difference if there were rebellion against the testing and the test administrator. Girls in Project TALENT also had a consistently lower rate of omitted items than the boys.

Our conclusion to this section of our report is that the confound with the test booklets was coincidental, that other spurious explanations are of low probability, and that the performance differences are real.

RESULTS FROM SELF-REPORT QUESTIONNAIRES

If the cognitive differences are indeed real, analysis of the self-report measures available in TALENT may provide some insights. The two vocabulary groups can be compared with the entire 10th-grade sample as a baseline. This is accomplished again where applicable with the use of own-sex standard scores. There are, of course, large sex differences on many of the measures to be reported, but means of most self-report scores are a good deal more labile with cultural change than are the means on tests of abilities. There has been ample occasion for change in means since 1960, but relations involving self-report scores are also much more robust to cultural change than are the mean levels.

Health Questionnaire

There are 43 questions concerning health, physical well-being, and physical wholeness in TALENT's Student Information Blank. Based on the evidence presented thus far, the high group might have had severe visual problems, but this is not the case. These groups report about the same incidence of visual problems, including myopia, of 10th-grade students in general and are well below a mathematically gifted group in this regard (Lubinski & Humphreys, 1992). Nothing stands out on the remaining self-report items, including hearing problems and inabilities to use arms, hands, and legs fully. The high group has a failure to answer rate somewhat greater than the norm (20% vs. 10%) but rate of failure to respond is inversely correlated with intelligence (Humphreys, Davey, & Kashima, 1986).

In contrast to the high group, the low vocabulary group responds to these questions at a substantially lower rate, about 50% on average. Although Lubinski and Humphreys (1992) corrected data for differences in response rate, their mathematically gifted and socially privileged groups required little correction. Correcting proportions for the present groups is questionable. One can conclude that the low and high groups are composed, at the very least, of different kinds of people. One can also conclude, if there is organic involvement in the performance of the high group, that the test performance is more symptomatic than any item or combination of items in TALENT's health questionnaire.

Personality

Table 5 contains a comparison for the high and low groups of scores from the personality questionnaire. It is significant in the interpretation of these data that the measuring instrument did not have a forced-choice format, and the intercorrelations define a general factor. The high group of boys has average to slightly below average scores on these scales, but the low group is substantially below the mean on most. The median difference between the groups is a third of a standard deviation, and the low group is slightly higher on only one scale, Leadership. The results are similar for the girls. The high group tends to be slightly below average, and the low group is even lower than the low group of boys. The median difference for the girls is almost half a standard deviation, but again the low girls have a Leadership mean slightly above the norm and superior to that of the high group. The high vocabulary groups are close to the norm on everything, but the low groups are near normal on only two scores. To use a currently popular concept, high and low vocabulary groups have distinctly different self-concepts.

Interest Tests

The outcomes for the interest inventory scores in Table 6 are a little unexpected after the data in Table 5. The high group of boys is close to the mean in everything other than a small elevation in trades and labor. The low group expresses more interest than the highs in everything except sports, hunting-fishing, and farming. The low group shows the pattern of unrealistic interests typically found in students of low intelligence. The pattern of interests for the girls in the high group is similar to that for the high boys, but there is little similarity of the sexes in the low group. For several categories expected scores based on sex differences in the normative population are reversed. Thus low boys are relatively high in literary-linguistic, social service, and office work, while girls are relatively high in physical science, public service, and mechanical-technical occupations.

It appears that the high vocabulary boys and girls who were at the low end of the distribution of the Intelligence Composite have expressed interests

Table 5
Differences Between High and Low Groups on Personality Questionnaire Scores[a]

	Boys			Girls		
	High	Low	Diff.	High	Low	Diff.
Sociability	−0.21	−0.74	0.53	−0.23	−0.84	0.61
Sensitivity	−0.22	−0.56	0.34	−0.23	−0.83	0.60
Impulsive	0.09	−0.08	0.17	0.03	−0.19	0.22
Vigor	−0.27	−0.59	0.32	−0.28	−0.59	0.31
Calmness	−0.26	−0.69	0.43	−0.19	−0.69	0.50
Tidiness	−0.13	−0.57	0.44	−0.22	−0.71	0.49
Culture	−0.14	−0.51	0.37	−0.22	−0.74	0.52
Leadership	0.14	0.26	−0.12	0.07	0.19	−0.12
Confidence	−0.21	−0.66	0.45	−0.10	−0.55	0.45
Mature	−0.23	−0.52	0.29	−0.18	−0.48	0.30

[a] $S_{\bar{x}} \approx 0.05$; $S(\bar{x}_1 - \bar{x}_2) \approx 0.07$.

Table 6
Differences Between High and Low Groups on Interest Inventory Scores[a]

	Boys			Girls		
	High	Low	Diff.	High	Low	Diff.
Physical science	−0.27	−0.03	−0.19	−0.04	0.33	−0.37
Biological science	−0.19	0.19	−0.38	−0.08	0.16	−0.24
Public service	0.01	0.30	−0.29	0.16	0.43	−0.27
Literary-linguistic	0.05	0.52	−0.47	−0.06	0.04	−0.10
Social service	0.03	0.63	−0.60	−0.01	0.13	−0.14
Art	0.02	0.27	−0.25	−0.16	−0.20	0.04
Music	0.18	0.54	−0.36	0.11	0.33	−0.22
Sports	−0.12	−0.15	0.03	−0.06	−0.09	0.03
Hunting-fishing	−0.08	−0.24	0.16	0.01	0.11	−0.10
Business management	0.03	0.45	−0.42	0.17	0.41	−0.24
Sales	0.05	0.48	−0.43	0.15	0.41	−0.26
Computation	0.04	0.52	−0.48	0.11	0.44	−0.33
Office work	0.15	0.79	−0.64	0.17	0.18	−0.01
Mechanical technical	0.07	0.32	−0.25	0.11	0.67	−0.56
Trades	0.30	0.80	−0.50	0.24	0.90	−0.66
Farming	0.02	−0.01	0.03	−0.12	0.07	−0.19
Labor	0.36	0.91	−0.55	0.22	0.93	−0.71

[a] $S_{\bar{x}} \approx 0.05$; $S(\bar{x}_1 - \bar{x}_2) \approx 0.07$.

reasonably in line with their measured abilities. Although they felt better about themselves, as pictured by their personality scores, they are not sanguine about their vocational prospects.

Background Scores

Biographical information scores are presented in Table 7. Both groups and both sexes are below average in socioeconomic status, but the low groups are substantially lower. As a matter of fact, their means are about what one would expect from a group selected to be more than two standard deviations below the mean in intelligence. There is little difference between highs and lows in their high school curriculum, and the high groups report fewer solid courses and lower grades. The low groups received less guidance from out-of-school sources, which is in line with their lower socioeconomic status. Somewhat surprisingly, the high groups evaluated reading competence at a somewhat higher level, but this might be true when they have unlimited time. The data in Table 4 showed that, in general, they did not work as fast as the low groups. The largest difference between the groups occurs in number of organizational memberships. Both compensate for low scores in abilities and grades by participating in organizations, but this tendency is substantially more marked in the low groups. Differences between the high and the low groups occur in both directions and are similar for the two sexes. The swings in signs of differences, however, are almost entirely a function of the reports made by the two low groups. High groups again tend to have means close to the norm. They are more nearly average in all of their self-report scores, which is more in line with their vocabularies than with their Intelligence Composite results.

DISCUSSION

Several sources of evidence indicate that the unusual regressions of vocabulary and mechanical reasoning that initiated this investigation are psychologically important phenomena. Parallel results have also been reported in the learning disability literature for verbal and spatial tests for two groups defined by a completely different set of operations. These are the data of Dean, Schwartz, and Smith (1981).

A Different Route to Similar Results

Dean et al. (1981) divided their subjects into learning disabled-bilateral and learning disabled-coherent (their terms) patterns, with 38 preadolescent males in each group. The tests used were the Wide Range Vocabulary Test and the Raven Matrices (most similar to TALENT's abstract reasoning). The so-called normally lateralized group was approximately 1.5 standard deviations lower than the bilateral group on the verbal measure, but was about 0.5 standard

Table 7

Differences Between High and Low Groups on Biographical Information Composites[a]

	Boys			Girls		
	High	Low	Diff.	High	Low	Diff.
Socioeconomic status	−0.46	−0.80	0.34	−0.52	−1.01	0.49
Curriculum	−0.44	−0.40	−0.04	−0.33	−0.33	0.00
Solid courses	−0.09	0.15	−0.24	−0.37	−0.11	−0.26
Grades	−0.14	0.11	−0.25	−0.38	−0.10	−0.28
H.S. guidance	0.21	0.36	−0.15	0.22	0.24	−0.02
Non-H.S. guidance	−0.17	−0.32	0.15	−0.23	−0.55	0.32
Study habits	−0.48	−0.44	−0.04	−0.48	−0.73	0.25
Writing skills	−0.09	−0.03	−0.06	−0.13	−0.29	0.16
Reading skills	−0.24	−0.43	0.19	−0.23	−0.49	0.26
Extra reading	−0.09	0.14	−0.23	−0.16	−0.20	0.04
Variety of organizations	0.34	1.19	−0.85	0.31	1.15	−0.84
Participation	0.29	0.96	−0.67	0.20	0.79	−0.59
Variety of hobbies	0.01	0.35	−0.34	−0.12	0.23	−0.35
Participation	0.04	0.43	−0.39	−0.02	0.22	−0.24
Sports	−0.27	−0.73	0.46	−0.28	−0.54	0.26
Leadership	0.14	0.08	0.06	−0.07	−0.15	0.08
Social life	0.32	0.06	0.26	0.07	−0.38	0.45
Work	0.13	0.02	0.11	0.12	0.26	0.14

[a] $S_{\bar{x}} \approx 0.05$; $S(\bar{x}_1 - \bar{x}_2) \approx 0.07$.

deviation higher on the spatial measure. The interaction was significant at well below the .05 level in spite of the small samples.

The measures that separated the laterality groups had been validated in a previous study involving independent samples of learning disabled and normal children. Lateral preference patterns were obtained on the Dean Laterality Preference Schedule (Dean, 1978). A low score on an item in the schedule indicates right preference, a high score left preference. The items were factor-analyzed, and item scores were summed for all items that loaded on a given rotated factor, of which there were six. Three scores separated the groups significantly in univariate analysis. These were visually guided fine motor-hand, fine motor-foot, and ear preference. When the six scores were used in a multivariate analysis, substantial group separation was obtained on a single dimension defined by positive weights for fine motor-hand and ear and by negative weights for general handedness and eyedness. Intercorrelations of the six scores were not reported, but it is clear from their near-zero validities that general handedness and eyedness suppressed common variance in the six scores.

Dean et al. (1981) described their discriminant score as a measure of mixed dominance or bilaterality. It does describe a mix, but it is a special kind of mix. Their maximum bilaterality is found only when visually guided fine motor and ear preferences are for the left and general handedness and eyedness are for the right. The opposite pattern, which is equally bilateral, produces a minimal score on their discriminant scale. (A negative weight is given to the high score, a positive weight to the low, in the reverse pattern of factor scores.) This would seem to require that a more specific name be assigned to the function measured by their discriminant.

The Contribution of Our Data

Our data could not be related to the laterality issue, but they do provide a great deal of dependable information that paints a consistent picture of two kinds of low-achieving children. These two groups of children differ substantially from each other in many more ways than in their performance on vocabulary and spatial visualization tests. High vocabulary groups perform relatively well on a gamut of unspeeded cognitive tests that require retrieving from memory definitive answers to unambiguous questions. They perform more poorly than low vocabulary groups on a gamut of unspeeded cognitive tests that require inferences from or about complex, novel, ambiguous questions or problems. Although tests of spatial visualization are prominent in this second set, the cluster includes tests that appear to be highly verbal.

There is also a substantial split in results between a speeded test of the four basic numerical operations and a second speeded test that requires identification of complex, shaded, figural drawings as same or different. The examinee is required to select the one different drawing within a set of five. A good performance on Numerical Operations requires speedy retrieval from memory of a truly definitive answer to a question of essentially zero ambiguity (What does 2 + 2 equal?). A good performance on Object Identification requires rapid, systematic checking of the many aspects of the figures on which the minute differences provided by the test constructor could occur. Short-term visual memory of small detail does seem to be a feasible work method for these items, but a holistic memory for the figure is not sufficient. The items are meaningless and the differences do not leap off the page at the viewer. Under time pressure examinees may not have confidence that they have found the right answer. Attempting to describe verbally the figural items would clearly result in poor performance.

Approximately the same small proportions of boys and girls show the phenomena in spite of sex differences on all of the tests involved. Some of these differences are substantial, such as the one on the measure of three-dimensional spatial visualization or mechanical reasoning. The proportions of high school students affected are quite small, and the individuals involved have widely scattered demographic characteristics. Self-reports on the health questionnaire indicate that the high vocabulary groups do not have obvious physical handicaps, such as visual problems, that adversely affected performance on a subset of the TALENT tests.

Other self-report scores show that the high and low vocabulary groups are composed of different kinds of people. The low groups tend to rate themselves below average on almost all of the personality traits measured by TALENT, while the high groups are only a little below the various means. The low groups have a differentiated interest pattern, though departing from traditional male–female patterns, while the high groups are not far from the normative mean in anything. The high groups are from families closer to the mean in socioeconomic status, but they report lower grades and fewer solid high school courses than the low groups. The latter are more active in hobbies and extracurricular organizations. Differences between the high and the low groups occur in both directions, but this is largely due to swings in size of means of the low groups. The high groups tend to be a little below average on most self-report scores, but are not far from the average on those scores where they are above the mean.

Some Observations Relevant to a Theory

The cognitive tests on which our groups of high and low verbal students differ have loadings in common only on a general factor in the full range of high school talent. The tests on which the high verbal group is relatively high are verbal, but so are some of the tests on which they are relatively low. The shared verbal content may temper the size of the deficit on the tests Word Functions in Sentences and Disguised Words. Similarly, the four speeded tests measure a common factor in the full range of high school talent, but two of the four show widely different patterns of results for the high and the low verbal groups.

There is a presumption that our groups would show the particular combination of laterality reported by Dean et al. (1981), but this would require a major new research undertaking. Subjects who meet the regression criteria, whether or not they have been diagnosed as learning disabled, should be administered the Dean (1978) laterality scales. A more limited check on our findings can be obtained by inspection of the regression of the verbal on the performance scales of an individual or group test of intelligence that has been administered in a wide range of talent.

If our high verbal group were just a little closer to the mean of the distribution of the socioeconomic index, they would fit the pattern of recessive inheritance of a trait. Seriously retarded children having a known recessive gene are found in families that have a mean parental intelligence score (and socioeconomic status) at the population mean with a standard deviation also of the same size as the population value. If our high verbal groups have a genetic handicap, however, it does not appear to be sex-linked, and it would explain a portion only of students who have a reading disability.

It is conceivable that we could modify our arbitrary criteria for group membership to produce a high verbal group closer to the population mean. Our groups fell below the *approximate* point of inflection of the regression of vocabulary on the intelligence composite. We split the high and low groups at a point that placed the low group *approximately* on the extension of the linear regression line, and at a point that produced approximately *equal Ns* for statistical convenience. Refinement of the criteria might produce even larger differences between the high and the low groups.

On the other hand, the statistical criteria we used cannot be drastically modified. There are not substantial numbers of students with this particular handicap who have intelligence test scores higher than the two standard deviations cut-off. The linearity and steepness of the regression of vocabulary on the intelligence composite, plus the small variances about the regression line, negate this possibility.

ACKNOWLEDGMENTS

This research was supported by the Pioneer Fund and by a grant from the National Science Foundation (MDR8855625).

REFERENCES

Dean, R. S. (1978). *Dean Laterality Preference Schedule.* Madison: University of Wisconsin Press.

Dean, R. S., Schwartz, N. H., & Smith, L. S. (1981). Lateral preference patterns as a discriminator of learning difficulties. *Journal of Consulting and Clinical Psychology, 49*, 227–235.

Flanagan, J. C., Dailey, J. T., Shaycoft, M. F., Gorham, W. A., Orr, D. B., & Goldberg, I. (1962). *Design for a study of American youth*. Boston: Houghton Mifflin.

Flanagan, J. C., Davis, F. B., Dailey, J. T., Shaycroft, M. F., Orr, D. B., Goldberg, I., & Neyman, C. A., Jr. (1964). The American high school student. Final Report, Cooperative Research Project No. 635, U.S. Office of Education. Pittsburgh: Project TALENT Office.

Humphreys, L. G., Davey, T. C., & Kashima, E. (1986). Experimental measures of cognitive privilege/deprivation and some of their correlates. *Intelligence, 10*, 355–376.

Lubinski, D., & Humphreys, L. G. (1992). Some bodily and medical correlates of mathematical giftedness and commensurate levels of socioeconomic status. *Intelligence, 16*, 99–115.

Lubinski, D., & Humphreys, L. G. (Submitted for publication.) Using individual differences measures to predict group performance with near certainty.

Skinner, B. F. (1950). Are theories of learning necessary? *Psychological Review, 57*, 193–219.

Wise, L. L., McLaughlin, D. H., & Steel, L. (1979). *The Project* TALENT *Data Handbook*. Palo Alto, CA: American Institute for Research.