

INTERACTION OF LEVEL I AND LEVEL II ABILITIES WITH RACE AND SOCIOECONOMIC STATUS¹

ARTHUR R. JENSEN²

University of California, Berkeley

Two levels of mental abilities were hypothesized to interact with socioeconomic status and/or race such that (a) socioeconomic-status differences were greater for Level II than for Level I abilities and (b) the correlation between Levels I and II and the regression of Level I upon Level II were greater in upper- than in lower-socioeconomic-status populations. These hypotheses were borne out by the present data, consisting of Level I measures (digit-span memory) and Level II measures (Large-Thorndike Intelligence Tests, verbal and non-verbal) obtained on all white and black pupils in Grades 4-6 in one school district. The largest effects were attributable to differences between the white population and the low-socioeconomic-status-black group.

The present study tests Jensen's Level I-Level II theory of mental abilities in a total school population. The theory has been tested heretofore only with specially selected samples from the population.

The theory and related evidence have been presented in detail elsewhere (Jensen, 1968, 1969, pp. 109-117, 1970a, 1970b, 1973, pp. 193-293; Jensen & Rohwer, 1968). Briefly, the theory involves two types of mental abilities, Level I and Level II, and their interaction with population (socioeconomic status and/or race) differences. Level I ability consists of rote learning and primary memory; it is the capacity to register and retrieve information with fidelity and is characterized essentially by a relative lack of transformation, conceptual coding, or other mental manipulation intervening between

information input and output. Level II ability, in contrast, is characterized by mental manipulation of inputs, conceptualization, reasoning, and problem-solving; it is essentially the general intelligence (*g*) factor common to most complex tests of mental ability and standard tests of intelligence. Level I abilities are best measured by rote-learning tasks: serial learning, repeated trials of free recall of a number of successively presented familiar uncategorized objects, pictures, or nouns, and tests of short-term memory, such as digit span. Level II ability is best measured by tests of general intelligence that have a high general intelligence loading and especially those of the nonverbal, fluid-intelligence, culture-fair variety.

An interesting point about Level I and Level II abilities is their interaction with socioeconomic status and race, as has been shown in the articles cited previously. The first studies showed mainly that in groups of children selected for low Level II ability (IQs of 60-80), the low-socioeconomic-status children (white or black, although socioeconomic status and race are confounded in some studies) obtain markedly higher scores on Level I tests (usually approaching children with average IQs of 90-110) than are obtained by the middle- or upper-socioeconomic-status children with

¹ This research was supported by the Office of Economic Opportunity under Contract OEO 2404 and by grants to the University of California by the Berkeley Unified School District and the Sterling Morton Charitable Trust.

The author is grateful for the cooperation of the Berkeley Unified School District and for the able assistance of Wade Egbert, who supervised the testing, and Carol Treanor and Barbara Cerny, who conducted the computer analyses of the data.

² Requests for reprints should be sent to Arthur R. Jensen, Institute of Human Learning, University of California, Berkeley, California 94720.

the same low IQs. On Level I tasks, middle-socioeconomic-status children with low IQs perform more in accord with their low IQ, while low-socioeconomic-status children perform more like children of average IQ. This finding suggests a lower correlation between Level I and Level II ability of low-socioeconomic-status than in middle-socioeconomic-status populations. Also, it means that, in general, groups differing in socioeconomic status should differ less in Level I ability than in Level II ability. Thus it was suggested that if Level I ability could be made more important in the educative process, there might be a chance of diminishing the present large differences in scholastic performance associated with socioeconomic status and racial group differences in Level II ability, which is known to correlate highly with scholastic achievement in the prevailing system of education.

The earlier studies were based on a 2×2 analysis of variance design: high (or middle) socioeconomic status versus low socioeconomic status and high IQ (100–120) versus low IQ (60–80) thus forming four groups. Typically there were equal numbers of subjects (20–40) in each of the four groups. The low-IQ groups were often selected from classes for the educable mentally retarded with average IQs slightly below 70 (since 75 is the cutoff for admission to educable mentally retarded classes in California public schools). Because of the difficulty of matching low- and high-socioeconomic-status groups for high IQ, the “high-IQ” groups were usually only slightly above average (i.e., IQ of 105–110). The socioeconomic-status difference in Level I (learning) ability for the low-IQ groups was always highly significant, but the low- and high-socioeconomic-status groups of high IQ (i.e., about an IQ of 105–110) usually did not differ significantly, although the crossover or disordinal type of interaction usually appeared.

The present Level I–Level II hypotheses can be stated in their simplest form as follows:

1. Social classes do not differ, on the average, in Level I ability, but differ on Level II ability. (Another way of stating this is that Level I ability is not correlated with socioeconomic status and Level II ability is

positively correlated with socioeconomic status.)

2. The regression of Level I upon Level II ability is greater (i.e., steeper slope of the regression line) in upper- and middle-socioeconomic-status populations than in low-socioeconomic-status populations. A less general corollary of this is that the correlation between Level I and Level II is greater in upper- and middle-socioeconomic-status populations than in low-socioeconomic-status populations. It is less general because restriction of the range of talent can affect the size of the correlation coefficient; whereas the slope of the regression line remains the same even if the distribution on one or both of the variables is truncated and the variance is thereby reduced. The correlation is lowered, therefore, but the slope of the regression line remains unchanged. The slope of the regression line (of Level I on Level II), therefore, is a more stable and fundamental datum. Thus, a proper test of the hypothesis should involve testing the difference in the regression of Level I upon Level II in low- and middle-socioeconomic-status groups.

The regression of Level II upon Level I has not been a part of the theory and cannot be inferred from the theory unless certain assumptions are made, assumptions for which at this point there seems to be no real theoretical basis. The lines of regression of Level II upon Level I can be determined only if we assume a precise value of the correlations between Levels I and II in low- and middle-socioeconomic-status groups. The theory posits no precise values, for no specific value exists for the general case. The correlations are merely population parameters, which may vary according to the populations sampled and the method of classifying individuals by socioeconomic-status group. The theory only posits that the regression coefficient (i.e., slope) of Level I on Level II is *greater* (i.e., steeper slope) in middle-socioeconomic-status than in low-socioeconomic-status populations. The posited difference thus is *directional* rather than precisely quantitative. If the variances are assumed to be the same in the middle- and low-socioeconomic-status groups (an assumption that is independent of the theory),

then the correlation between Levels I and II will be greater in the middle- than in the low-socioeconomic-status group. Assuming this to be the case, then, the regression lines (of Level II upon Level I) should gradually converge. But the convergence would be very gradual, and assuming realistic values of the Level I–Level II correlations in the low- and middle-socioeconomic-status groups, there would be no point within ± 3 sigmas of the mean of a normal distribution (which includes 99.74 % of the population) of Level I scores at which low- and middle-socioeconomic-status groups matched on Level I ability would be equal in Level II ability or where the low-socioeconomic-status group would exceed the middle-socioeconomic-status group in Level II ability. If the means of the low- and middle-socioeconomic-status groups were assumed to differ by 1 sigma and if the Level I–Level II correlations were .6 and .4 in the middle- and low-socioeconomic-status groups, respectively, we would have to match subjects from the two socioeconomic-status groups for Level I scores at least 5 sigmas below the common Level I mean in order for the low-socioeconomic-status subjects, on average, to equal or exceed the middle-socioeconomic-status subjects on Level II ability. But any subjects who were 5 sigmas below the mean on Level I ability would be in the range of severe mental defect, at the imbecile or idiot level, where the deficit is more likely due to a major gene or chromosomal anomaly or to organic damage, rather than to the normal variations in the polygenic and environmental determinants of mental variation that operate in the bulk of the population. For most of the normal population, the regression lines of Level II upon Level I for the two socioeconomic-status groups would be practically parallel. Estimating the point of convergence at 5 sigmas below the mean assumes linearity of regression all the way down into the range of severest mental defect, and since the causal factors in that range are different than for the rest of the distribution, such an assumption is quite unwarranted. Within reasonable boundary conditions for the operation of the theory, the lines of regression of Level II upon Level I should be pictured as almost parallel, with

such slight convergence that the lines would not come together within the range of abilities normally found in the public schools. These hypotheses are depicted graphically in Figure 1.

As can be seen, the angles between the regression lines are different for the lower- and middle-socioeconomic-status groups (angles l and m), being smaller for the middle socioeconomic status. (The cosine of the angle between the regression lines is equal to the correlation coefficient when sigma is the same on both variables or the scores are standardized with the same sigma.)

3. The third element of the theory concerns the hierarchical relationship between Level II and Level I ability. The development of Level II ability, as well as Level II performance itself, is seen as having some functional dependence upon Level I ability, but the reverse is not true. For example, initial learning of the information and cognitive skills involved in Level II performance may depend in part upon short-term memory and its consolidation, which are Level I processes. Thus an individual with superior Level I ability will in the long run show better Level II performance than will a person with the same genetic and environmental potential for Level II ability but with

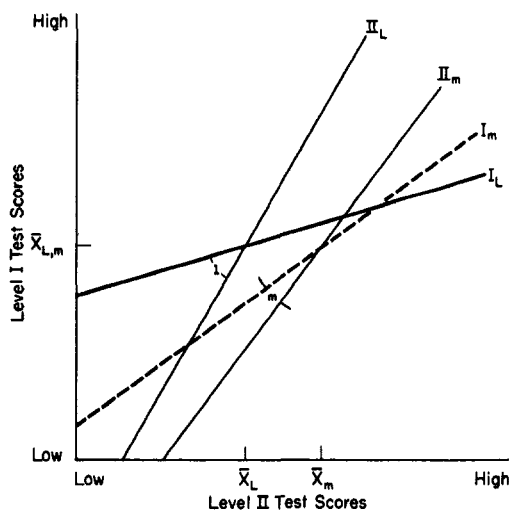


FIGURE 1. Hypothetical regression lines for relationship between Level I and Level II abilities in middle- and lower-class populations (angles l and m).

poorer Level I ability. Also, it seems reasonable to suppose that some short-term memory can be involved in solving Level II problems, such as Raven matrices items or the mental arithmetic subtest of the Wechsler scale, in which information must be retained in memory (i.e., Level I) while mental operations are being performed on it (i.e., Level II). A relatively pure Level I test, such as digit-span memory, on the other hand, can hardly be seen as depending upon the processes of abstraction, generalization, and conceptualization that are called for in Level II tests.

Another way of stating the hierarchical relationship between Levels I and II, that is, the functional dependence of Level II upon Level I ability, is to say that Level I is necessary but not sufficient for the development and operation of Level II ability. A consequence of this hierarchical formulation would be that one would seldom if ever find individuals with very high Level II ability who have very low Level I ability. The reverse, however, would not be uncommon, that is, persons with high Level I ability but low Level II ability. (In fact, quite extreme idiot savants of this type are known to exist.) As Matarazzo (1972, p. 204) has noted in connection with the clinical use of the Wechsler intelligence test, a low score on memory span for digits is highly related to general mental retardation, while a high score on digit span is not highly indicative of superior general intelligence. Matarazzo states,

Ordinarily, an adult who cannot repeat at least four or five digits forward is [in about 9 cases out of 10] either organically impaired or mentally retarded. Nevertheless, mental retardates sometimes do well on the Memory Span Test [p. 205].

This observation suggests a hierarchical (or necessary-but-not-sufficient) type of relationship between Levels I and II.

If this is in fact the case, the dispersion (i.e., the standard error of estimate) of Level I scores about the line of regression of Level I upon Level II should show a gradual and regular decrease in going from lower to higher scores on Level II. Thus the relative magnitudes of the standard error of estimate of Level I scores for low and high scores on

Level II provide a test of the hypothesis of hierarchical dependence of Level II upon Level I. That is, if the hypothesis is true, we should expect to find a larger dispersion of Level II scores in the lower range of Level I scores than in the higher range of Level I scores.

The purpose of the present study was to test each of these three main hypotheses (described under 1, 2, and 3) derived from the Level I–Level II theory of mental abilities.

METHOD

Subjects

The 2,612 subjects in this study consisted of virtually all the white ($n = 1,489$) and black ($n = 1,123$) children enrolled in regular classes of the fourth, fifth, and sixth grades from all 14 elementary schools of the Berkeley Unified School District in California.

The small percentage of children who were absent on the particular day that their class was tested are not included in this study. Also, test data on all children not classified in the school records (and according to their own parent(s)) as either white or black were excluded from the present study. (These excluded subjects, mostly Orientals, comprised about 10% of the total school population.)

The adult white population in this district is largely of middle or upper-middle socioeconomic status; the three largest employers (mostly of whites) are the university, the Lawrence Radiation Laboratory, and a large pharmaceutical firm, all of which employ workers with better than average education and socioeconomic status for the white population as a whole. The adult black population is predominantly lower-middle to low socioeconomic status, comprised largely of semi-skilled and unskilled workers, although it is a somewhat higher-socioeconomic-status group than the black populations in the surrounding communities, with fewer unemployed and on welfare.

All tests were group administered to the regular classrooms by a staff of testers (3 whites and 3 blacks) who were specially employed and trained for this purpose. The white and black testers were assigned to classes at random. In any given class, the Level I and Level II tests were always administered by different testers on different days never more than one week apart. Thus the correlations between the Level I and Level II tests would not be systematically affected by any individual tester biases.

Tests

Control tests. Two different control tests were used, one in each of the two testing sessions. The

main purposes of the control tests were to set a good test-taking attitude in the class, emphasizing attention and effort while at the same time lessening tension and test anxiety by giving subjects tasks they could perform successfully simply by being attentive and trying their best.

The listening-attention test was given just before the Level I test (memory for numbers). The listening-attention test measures the child's ability to attend to and follow orally given directions paced at two-second intervals by means of a tape recording. The child is presented with an answer sheet containing 100 pairs of digits in sets of 10. The child listens to a tape recording that speaks one digit every two seconds. The child is required to put an X over the one digit in each pair that has been heard on the tape recorder. The purpose of this test is to determine the extent to which the child is able to pay attention to numbers spoken on a tape recorder, to keep his place in the test, and to make the appropriate responses to what he hears from moment to moment. Low scores on this test indicate that the subject is not up to validly taking the memory for numbers test, which follows immediately. High scores on the listen-attention test indicate that the subject has the prerequisite skills for taking the digit-span (memory for numbers) test. The listening-attention test thus is intended as a means for detecting students who, for whatever reason, are unable to hear and to respond to numbers read over a tape recorder. The test itself makes no demands on the child's memory, only on his ability for listening, paying attention, and responding appropriately—all prerequisites for the digit-memory test that follows.

The speed and persistence test (making Xs) was always given just before the Level II tests (Lorge-Thorndike IQ). The making Xs test is intended as an assessment of test-taking motivation. It gives an indication of the subject's willingness to comply with instructions in a group-testing situation and to mobilize effort in following those instructions for a brief period of time. The test involves no intellectual component, although for young children it probably involves some perceptual-motor skills component, as reflected in other studies by increasing mean scores as a function of age between Grades 1-5. Individual differences among children at any one grade level would seem to reflect mainly general motivation and test-taking attitudes in a group situation. Children who do very poorly on this test, it can be suspected, are likely not to put out their maximum effort on ability tests given in the same group situation, and to that extent, their ability test scores are not likely to reflect their real level of ability.

The making Xs test consists of two parts. On Part 1 the subject is asked simply to make Xs in a series of squares for a period of 90 seconds. In this part the instructions say nothing about speed. They merely instruct the child to make Xs. The maximum possible score on Part 1 is 150, since there are 150 squares provided in which the child

can make Xs. After a two-minute rest period the child turns the page of the test booklet to Part 2. Here the child is instructed to show how much better he can perform than he did on Part 1 and to work as rapidly as possible. The child is again given 90 seconds to make as many Xs as he can in the 150 boxes provided. The gain in score from Part 1 to Part 2 reflects both a practice effect and an increase in motivation or effort as a result of the motivating instructions, that is, instructions to work as rapidly as possible.

Level I test. Previous studies have indicated that one of the most unambiguous and reliable Level I measures is digit-span memory. A specially devised test of such memory, which has much higher reliability than the short digit-span tests included in such general test batteries as the Stanford-Binet and the Wechsler, is the author's memory for numbers test. It has three parts. Each part consists of six series of digits going from four digits in a series up to nine digits in a series. The digit series are presented on a tape recording on which the digits are spoken clearly by a male voice at the rate of precisely one digit per second. The subjects write down as many digits as they can recall at the conclusion of each series, which is signaled by a bong. Each part of the test is preceded by a short practice test of three digit series in order to permit the tester to determine whether the child has understood the instructions, etc. The practice test also serves to familiarize subjects with the procedure of each of the subtests. The first subtest is labeled immediate recall. Here the subject is instructed to recall the series *immediately* after the last digit has been spoken on the tape recorder. The second subtest consists of delayed recall. Here the subject is instructed not to write down his response until 10 seconds have elapsed after the last digit has been spoken. The 10-second interval is marked by audible clicks of a metronome and is terminated by the sound of a bong, which signals the subject to write his response. The delayed recall condition invariably results in some retention decrement. The third subtest is the repeated series test, in which the digit series is repeated three times prior to recall; the subject then recalls the series immediately after the last digit in the series has been presented. Again, recall is signaled by a bong. Each repetition of the series is separated by a tone with a duration of one second. The repeated series almost invariably results in greater recall than the single series. This test is very culture fair for children in second grade and beyond, who know their numerals and are capable of listening and paying attention, as indicated by the listening-attention test. The maximum score on any one of the subtests is 39, that is, the sum of the digit series from four through nine. Only the total score (i.e., the sum of the scores on the three subtests) is used in the present study.

Level II tests. Level II was measured by the Lorge-Thorndike Intelligence Test (Level 3, Form B), which has two parts, verbal and non-verbal. This is a nationally standardized group-

administered test of general intelligence. In the normative population, which was intended to be representative of the nation's school population, the test has a mean IQ of 100 and a sigma of 16. The test is primarily a measure of reasoning ability; it has a high general intelligence saturation when factor analyzed with other mental ability tests, so it is deemed a good Level II test, especially the nonverbal part, which is based on pictorial problems and depends not at all upon reading skill or scholastic knowledge.

RESULTS

Control Tests

On the listening-attention test there was no significant difference between the white and black groups in any grade. The mean number correct (out of 100) was above 98 for all groups in every grade, and the 25th, 50th, and 75th percentile score was 100 in each group at each grade. Since a perfect score on this test is 100, it is evident that the vast majority of subjects were motivated to do their best in the test situation and were capable of correctly hearing the numerals as spoken over the tape recording and of properly following directions and registering their responses on answer sheets. Practically all subjects obtained a perfect score. At this age level, there is no appreciable difference between the grades or between whites and blacks on the listening-attention test. Since the correlation between the listening-attention test and either the memory test or the Lorge-Thorndike Intelligence Tests is not significantly greater than zero in the white or black group, it is clear that no significant amount of the variance in these tests is attributable to differences in the kinds of sustained attentiveness and willingness to comply with instructions that are assessed by the listening-attention test.

On the speed and persistence test (making Xs), the black group scored significantly higher than the white group on both the first and second try and on the gain score (i.e., the difference between second try — first try), and these differences are fairly consistent across the three grades. These results, like those for the listening-attention test, indicate that at least equally good cooperation and effort in the test situation were put forth by the black subjects as by the white subjects. The lower quartile scores should be

a most sensitive indicator of children who are not cooperating or putting out much effort, and at every grade the performance of black subjects equals or exceeds that of the white subjects. These results contradict the common notion that black children have a slower "personal tempo" or are less cooperative or more lackadaisical in a test situation. The correlation between making Xs and the memory for numbers and Lorge-Thorndike Intelligence Tests are close to zero in both racial groups.

Mean White-Black Difference in Memory (Level I) and Intelligence (Level II) Tests

The hypothesis in its most simple and extreme form states that low- and middle-socioeconomic-status groups differ in Level II but not in Level I ability. Table 1 shows the raw score means on the Level II and Level I tests in the white and black groups and (in the last column) shows the group difference in terms of the total within-groups variation. We see that although the white-black difference is highly significant both on the memory and on the intelligence tests, the difference on the intelligence tests is more than twice the difference on the memory test. It is thus unclear whether this finding disproves or supports the hypothesis. It would seem to disprove the "no difference on Level I" aspect of the hypothesis, and yet

TABLE 1
RAW SCORE MEANS AND STANDARD DEVIATIONS ON INTELLIGENCE (LEVEL II) AND MEMORY TESTS (LEVEL I) AND MEAN WHITE-BLACK DIFFERENCES IN SIGMA UNITS

| Variable | White (n = 1,489) | | Black (n = 1,123) | | $(M_W - M_B)/\sigma$ |
|-----------------|----------------------|-------|----------------------|-------|----------------------|
| | M | SD | M | SD | |
| Age (in months) | 131.23 | 10.89 | 132.61 | 11.24 | -.12 |
| Intelligence | | | | | |
| Verbal | 69.85 | 12.56 | 46.24 | 16.88 | 1.62 |
| Nonverbal | 63.12 | 10.83 | 43.47 | 14.50 | 1.57 |
| Memory | | | | | |
| Immediate | 23.33 | 6.41 | 18.75 | 6.61 | .70 |
| Repeat | 26.89 | 5.81 | 23.40 | 6.56 | .57 |
| Delay | 24.25 | 5.76 | 20.29 | 6.73 | .64 |
| Total | 74.48 | 15.58 | 62.45 | 16.82 | .75 |

Note. Sigma is the square root of the combined within-groups variance.

the results are consistent with the hypothesis in that the white-black difference on the memory test is very much less than the difference on the intelligence test. Since the theory also posits a correlation between Level I and Level II and a higher correlation in higher- than in lower-socioeconomic-status groups, we should expect a Level I difference between the present white and black groups if the IQ of the white group is above the white mean for the general population or if the present black group is below the general black mean IQ. In the general population, the groups differ by only about one sigma or 16 IQ points, while in our Berkeley population the difference is considerably greater. In terms of the Lorge-Thorndike national norms, the results of the present white group are verbal IQ, $M = 118.4$, $SD = 15.7$; nonverbal IQ, $M = 120.24$, $SD = 14.6$. The results of the present black group are verbal IQ, $M = 92.8$, $SD = 13.9$; nonverbal IQ, $M = 95.4$, $SD = 15.5$. The consequences of this difference between the groups used in the present study and the averages for the general United States population can be more easily discussed in the next section in connection with the regression of memory scores upon intelligence scores.

Regression of Memory upon Intelligence

The hypothesis predicts a steeper slope of the regression line of Level I (memory) scores upon Level II (intelligence) scores in the white group than in the black group. Figure 2 shows the relevant regression lines for the Lorge-Thorndike nonverbal scores. The graphical results are practically identical for the verbal scores.

A statistical test of departure from linearity was applied to all the regressions and none was found to depart significantly (at the .10 level) from linear regression. Though the linearity of the regression appeared to extend throughout the entire range of scores for both racial groups—a total range of more than 100 IQ points beginning at about an IQ of 50—the regression lines shown in Figures 2 and 3 were drawn to extend only over the range of scores that permits an unequivocal test of departure from linearity. (The ns at the very extremes of the distributions [less than the upper and lower

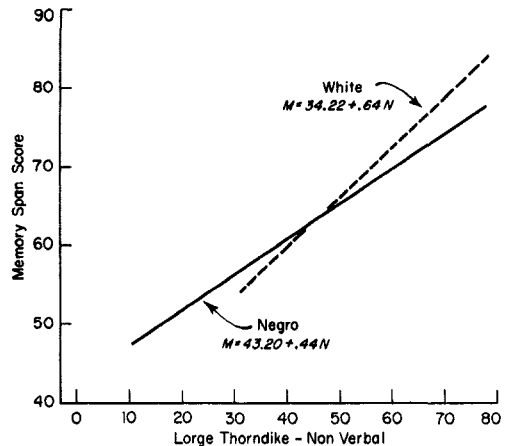


FIGURE 2. Regression of memory scores upon Lorge-Thorndike nonverbal intelligence raw scores in white and black groups.

2.5%] are too small and scattered to permit statistical confidence of linearity in the regions beyond approximately the middle 95% of the distributions.)

We see in Figure 2 (and in verbal scores as well) that the regression is greater for whites than for blacks. For the Lorge-Thorndike verbal scores, the regression, b , is .58 for whites and .42 for blacks. A test³ of the significance of the difference between the two slopes is highly significant ($t = 4.10$, $p < .001$). For the Lorge-Thorndike nonverbal scores, the white-black difference in slope is also highly significant ($t = 4.35$, $p < .001$).

According to the norms provided by the Lorge-Thorndike test manual, the point on the scale of raw scores at which the regression lines for the present white and black groups cross over is equivalent to a Lorge-Thorndike IQ of approximately 100, both for the verbal and the nonverbal tests. In the range of intelligence below an IQ of 100, the black children, on the average, surpass white children in memory scores; in the range above an IQ of 100, the white children surpass the black children in memory performance. The crossover in the above-average IQ range is clearly not a statistical

³ The standard error of the regression (see Guilford, 1956) of Y on X , b_{yx} , is

$$\sigma_{b_{yx}} = \frac{\sigma_y}{\sigma_x} \sqrt{\frac{1 - r_{xy}^2}{N}}$$

artifact as was originally believed when only small sample data based on selected groups were available. These results mean that, on the average, the white child below an IQ of approximately 100 has a poorer memory span than his black counterpart in IQ, and the white-black difference increases, in favor of the black child, the lower the IQ. In terms of national IQ norms, the approximately 80% to 85% of black children who fall below an IQ of 100 would, on the average, surpass in memory span the 50% of white children who fall below an IQ of 100. If we assume that the white and black regressions in the general United States population are the same as those in the present data and if the general white and black IQ means are 100 and 85, respectively, then, according to the regression equations in the present data, we should expect the white and black populations (which differ 1 sigma in IQ) to differ by only about .3 sigma to .4 sigma in memory span, in favor of whites. That is to say, the present data do not support the hypothesis of no white-black difference in Level I (here measured by the memory test), but the data do indicate a much smaller racial difference in memory than in IQ. This conclusion would, of course, not hold if the relative slopes of the regression lines for the two races are not about the same in the general population as in the Berkeley school population. The rather atypical nature of the Berkeley population with respect to mean Lorge-Thorndike IQ, especially in the white population, should make us wary of generalizing to the general population or to the populations of other communities with markedly different demographic and socioeconomic-status features than Berkeley.

Regression of Intelligence upon Memory

These regression lines present a very different picture from that of the regression of memory upon intelligence. As seen in Figure 3, the slopes of the regression lines for whites and blacks are parallel (the regression coefficients do not differ significantly), and they are separated by approximately 1.6 sigmas on the intelligence scales. (The results are virtually identical for the verbal scores.) Thus there is no point on the scale of memory

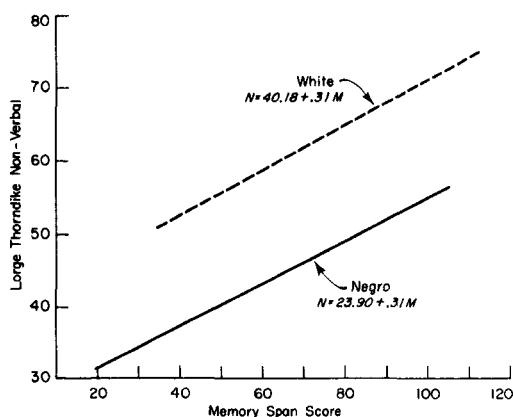


FIGURE 3. Regression of Lorge-Thorndike non-verbal raw scores upon memory scores in white and black groups.

scores at which equated groups of whites and blacks obtain equal intelligence scores. The picture is close to the hypothetical regression lines depicted in Figure 1. It would seem to be consistent with the hypothesis that Level I is necessary but not sufficient for the development and functioning of Level II. Why should white and black children with precisely the same memory performance differ by 1.6 sigmas on both the verbal and non-verbal intelligence measures? When matched for intelligence, on the other hand, whites and blacks are considerably more alike in memory, and they average just about the same in memory performance when matched on intelligence in the vicinity of an IQ of 100. In other words, it appears that if subjects have the intelligence, they have the memory; while if they have the memory, they do not necessarily have the intelligence.

Dispersion of Memory Ability as a Function of Intelligence

If it is true that intelligence depends upon memory but that the reverse does not hold, we should expect the dispersion of memory scores to show a systematic decrease going from low to high levels of intelligence. To test this hypothesis, the standard error of estimate of memory scores (i.e., the standard deviation of memory scores around the regression line of memory upon intelligence) was examined for systematic change over the full range of Lorge-Thorndike Intelligence Test scores, verbal and nonverbal. The

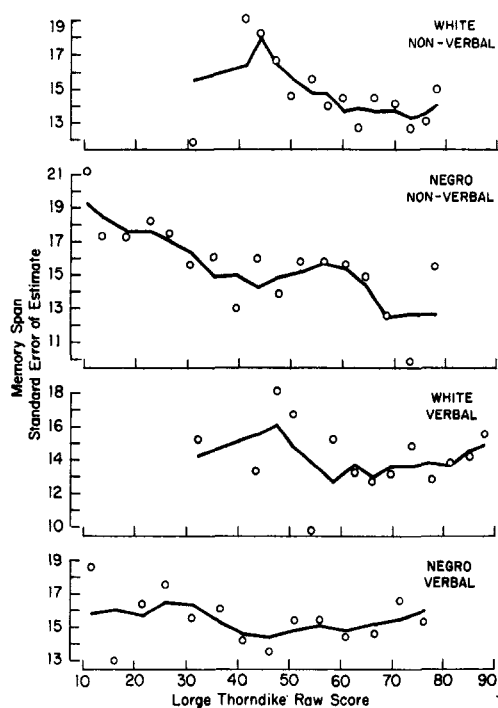


FIGURE 4. Memory score dispersion (standard error of estimate) as a function of Lorge-Thorndike verbal and nonverbal raw scores in white and black groups.

results are shown in Figure 4. Since the standard errors of estimates (indicated by circles) are rather erratic, their trend is better indicated by a moving average (the line going through the data points). For the nonverbal test the trend is clearly in accord with the hypothesis; that is, the standard error of estimate of the memory scores systematically decreases with increasing nonverbal intelligence. Bartlett's test for homogeneity of variances and a test of trend are both significant ($p < .01$) both for whites and for blacks. The results for the verbal test, however, yield only a faint suggestion of a decreasing standard error of estimate, and the trend is nonsignificant.

Thus the prediction based on the hypothesized hierarchical relationship between Level I and Level II is borne out by the nonverbal but not by the verbal test. Why should the two tests differ in this way? One can only speculate at this point. A possibility is that while both tests are highly saturated with general intelligence, the nonverbal test

is more a measure of what Cattell (1971, chap. 5) calls "fluid" intelligence and the verbal test is more a measure of "crystallized" intelligence. The hypothesized hierarchical relationship between Level I and Level II may hold only for Level II as measured by tests of fluid intelligence. But this conjecture cannot be tested with the present data and must await a study specially designed for this purpose.

Socioeconomic-Status Differences within Racial Groups

A questionnaire sent to the home of every child in the study, as well as the school records, served as the basis for classifying subjects according to socioeconomic status. Among other items of family background information obtained from the parents was the current occupation (or last job held in case of the unemployed) of the head of the household. Since returns of the parental questionnaire were considerably less than 100%, especially in the black group, and not all the questionnaires that were returned had answered the occupation question, the sample size for the socioeconomic-status analysis was reduced and the remaining subjects cannot be regarded strictly as a random sample of the Berkeley school population because of the self-selection in answering the questionnaire. When parent's occupation was not given in the questionnaire, it was sought in school records, but often without success. If the parental occupation appeared in the school records, it almost invariably was given in the questionnaire and vice versa. Lack of information or ambiguity or doubt in the socioeconomic-status classification of a given occupation was cause for omitting subjects from the present analysis.

Parental occupations were first coded into 82 job description categories. These were then reduced to seven categories in terms of conventional socioeconomic-status rankings of the occupations. But in order to obtain large enough socioeconomic-status samples to allow tests of Level I-Level II correlations and regressions within each sample, these seven categories had to be reduced to three broad socioeconomic-status categories as follows:

High Socioeconomic Status

- 1. High-level administrators, supervisors, college teachers.
- 2. High-level professionals, engineers, physicians, etc.

Middle Socioeconomic Status

- 3. White collar occupations requiring college or technical training.
- 4. Self-employed, technicians, skilled craftsmen.
- 5. Merchants, managers of small business, service workers, contractors.

Low Socioeconomic Status

- 6. Manual workers.
- 7. Nonmanual workers, relatively unskilled, jobs ordinarily requiring less than a high school diploma.

The categories are admittedly crude and somewhat arbitrary, but would undoubtedly correlate highly with any of the various methods of socioeconomic-status classification.

Table 2 gives the means and standard deviations of the three socioeconomic-status groups within each race. The row labeled "total" is based on the subjects who were classifiable. The "population" row consists of all subjects on whom test data were

available, whether they were classifiable by socioeconomic status or not. It can be seen that the total and population values do not differ appreciably in means or standard deviations, which indicates that the subjects who were classified by socioeconomic status are a fairly representative sample of the school population, at least as regards the present test variables.

Though there are the expected differences between each of the socioeconomic-status levels, among whites the largest differences are seen between the middle and low groups, while among blacks the largest differences are between the high- and middle-socioeconomic-status levels. But this difference between the racial groups is of little significance, since whites and blacks are not perfectly matched for occupations *within* the three broad socioeconomic-status categories. The average racial difference (see the last column of Table 2) within each socioeconomic-status level is slightly larger than the high-low socioeconomic-status differences within each race for the Lorge-Thorndike verbal and nonverbal scores. For the memory scores, on the other hand, the high-low socioeconomic-status difference within each racial group is greater than the difference

TABLE 2

MEANS AND STANDARD DEVIATIONS OF INTELLIGENCE AND MEMORY RAW SCORES OF SOCIOECONOMIC-STATUS LEVELS WITHIN RACIAL GROUPS AND MEAN DIFFERENCE IN SIGMA UNITS

| Test | Socioeconomic status | White | | | Black | | | $(M_W - M_B)/\sigma$ |
|---------------------------|-------------------------|----------|----------|-----------|----------|----------|-----------|----------------------|
| | | <i>n</i> | <i>M</i> | <i>SD</i> | <i>n</i> | <i>M</i> | <i>SD</i> | |
| Lorge-Thorndike verbal | High | 763 | 71.6 | 10.40 | 38 | 58.4 | 13.22 | 1.25 |
| | Middle | 287 | 70.9 | 9.98 | 43 | 50.5 | 15.88 | 1.87 |
| | Low | 215 | 60.3 | 16.42 | 414 | 45.7 | 15.99 | .90 |
| | Total ^a | 1,265 | 69.6 | 12.24 | 495 | 47.2 | 16.19 | 1.66 |
| | Population ^b | 1,489 | 69.9 | 12.56 | 1,123 | 46.2 | 16.88 | 1.63 |
| Lorge-Thorndike nonverbal | High | 763 | 65.3 | 8.75 | 38 | 52.3 | 11.36 | 1.46 |
| | Middle | 287 | 63.6 | 9.61 | 43 | 46.6 | 12.04 | 1.71 |
| | Low | 215 | 55.6 | 14.82 | 414 | 43.0 | 13.63 | .90 |
| | Total | 1,265 | 63.4 | 10.72 | 495 | 44.1 | 13.59 | 1.66 |
| | Population | 1,489 | 63.1 | 10.83 | 1,123 | 43.5 | 14.50 | 1.56 |
| Memory (total score) | High | 763 | 74.7 | 15.23 | 38 | 69.0 | 16.12 | .37 |
| | Middle | 287 | 73.7 | 14.01 | 43 | 63.4 | 16.65 | .72 |
| | Low | 215 | 65.5 | 15.28 | 414 | 60.9 | 16.76 | .28 |
| | Total | 1,265 | 72.9 | 15.34 | 495 | 61.7 | 16.86 | .71 |
| | Population | 1,489 | 74.5 | 15.58 | 1,123 | 62.5 | 16.82 | .74 |

Note. $M_W - M_B/\sigma$, where σ is the square root of the combined within-groups variance.

^a Total of all subjects who were classified by socioeconomic status.

^b The entire school population in Grades 4-6.

between the racial groups. Expressed in units of the white population sigma, the high-low socioeconomic-status difference for whites on the Lorge-Thorndike verbal = .90, on nonverbal = .90, on memory = .59; the corresponding figures for blacks are 1.02, .86, and .52. In both racial groups, the high-low socioeconomic-status difference is almost twice as great for the intelligence tests as for the memory test, which accords with the hypothesis at least in a directional sense, that is, Level II ability is more highly correlated with socioeconomic status than is Level I ability.

Correlations and regressions within socioeconomic-status groups. Table 3 shows the correlation and regression of memory upon intelligence in each of the socioeconomic-status groups by race. The theory predicts higher correlations and regression coefficients in upper- than in lower-socioeconomic-status groups. This is not completely borne out by the data. The white-socioeconomic-status groups show no systematic trends in this respect, but the black-socioeconomic-status groups show the predicted trend, that is, lower correlations and regressions with lower socioeconomic status. The black-high- and middle-socioeconomic-status groups both appear quite different from the black-low-socioeconomic-status group. The differences

of regression coefficients between all 15 possible contrasts of Socioeconomic Status \times Race groups in Table 3 were subjected to *t* tests (two-tailed) to determine their significance. Only three of the contrasts are significant beyond the .05 level (two-tailed), all involving only the nonverbal test: high-socioeconomic-status white - low-socioeconomic-status black ($p < .01$); low-socioeconomic-status white - low-socioeconomic-status black ($p < .01$); high-socioeconomic-status black - low-socioeconomic-status black ($p < .03$).

All the significant differences involve exclusively the low-socioeconomic-status-black group, and the only significant within-race socioeconomic-status difference is between high- and low-socioeconomic-status blacks. The difference in regressions, therefore, appears to involve race more than socioeconomic status, or a combination of race and socioeconomic-status effects, since the low-socioeconomic-status-black group is undoubtedly somewhat below the low-socioeconomic-status-white group in socioeconomic status. The regressions of the high- and middle-socioeconomic-status-black groups do not differ significantly from those of the white groups.

DISCUSSION

The present study examined three main aspects of the Level I-Level II theory of mental ability, namely (a) the relative magnitudes of socioeconomic status and white-black differences in Level I and Level II abilities, (b) socioeconomic status and racial differences in the correlation between Levels I and II and in the regression of Level I upon Level II, and (c) the hierarchical (i.e., necessary-but-not-sufficient) functional dependence of Level II performance on Level I ability.

The *a* theory as originally stated in its simplest form predicts a socioeconomic-status difference in Level II ability but not in Level I ability. This formulation, however, was intended more as an unambiguous basis for a directional prediction than as a precise expectation of reality, for in reality it is, of course, most improbable that there is "no difference" between any two populations in any given trait. So the realistic issue is the

TABLE 3

CORRELATIONS (*r*) AND REGRESSION COEFFICIENTS (*b*) OF MEMORY SCORES UPON LORGE-THORNDIKE VERBAL AND NONVERBAL RAW SCORES IN SOCIOECONOMIC-STATUS LEVELS WITHIN RACIAL GROUPS

| Socioeconomic status | White | | | | Black | | | |
|-------------------------|----------|----------|-----------|----------|----------|----------|-----------|----------|
| | Verbal | | Nonverbal | | Verbal | | Nonverbal | |
| | <i>r</i> | <i>b</i> | <i>r</i> | <i>b</i> | <i>r</i> | <i>b</i> | <i>r</i> | <i>b</i> |
| High | .376 | .551 | .361 | .628 | .552 | .673 | .594 | .843 |
| Middle | .414 | .581 | .324 | .472 | .408 | .428 | .496 | .686 |
| Low | .591 | .550 | .605 | .624 | .391 | .410 | .311 | .382 |
| Total ^a | .464 | .582 | .442 | .632 | .419 | .436 | .365 | .453 |
| Population ^b | .466 | .578 | .443 | .637 | .420 | .419 | .372 | .436 |

^a Total of all subjects who were classified by socioeconomic status (white, $n = 1,265$; black, $n = 495$).

^b The entire school population in Grades 4-6 (white, $n = 1,489$; black, $n = 1,123$).

relative magnitudes of differences between populations in Levels I and II. In accordance with previous findings, it was found that the white and black groups, and to a slightly lesser degree the high- and low-socioeconomic-status groups within each race, differed much more, on the average, in Level II than in Level I ability. The exact size of the differences, of course, depends upon the particular populations being compared and is not regarded as an intrinsic aspect of the theory, the main point of which is that populations can differ in these two classes of ability and that the direction of the difference in socioeconomically stratified populations is such that the higher and lower groups will show a greater difference on Level II than on Level I. The reason for this, according to the theory, is that social mobility in an industrialized society is more dependent upon Level II than upon Level I abilities. In the present study the white-black differences are larger than the socioeconomic-status differences within the racial groups, but the point is ambiguous here since the average socioeconomic-status difference between the races is probably greater than the high-low socioeconomic-status differences within each racial group. The strict criteria for socioeconomic-status classification used here resulted in the inclusion of a peculiarly small percentage of the black population in the high- and middle-socioeconomic-status categories. It would be advisable in future studies to have socioeconomic-status ratings on a continuous scale based on a large number of home background factors, which might reflect more closely the nature of the child's environment than does merely the occupational classification of his parents.

The hypothesized higher correlation of *b* between Levels I and II in the white than in the black group was fully borne out by the data, as was also the predicted higher regression of Level II upon Level I. The effect is largely attributable to the difference between the entire white sample and the low-socioeconomic-status-black group, which constituted the vast majority of the present black sample. The high-socioeconomic-status- and middle-socioeconomic-status-black groups do not differ significantly from

the white population in this respect but differ significantly from the low-socioeconomic-status-black group.

The cause of different Level I–Level II correlations (or regressions) in different populations has not yet been established and at present can only be hypothesized. There are several possible causes of correlation and they are not mutually exclusive: (a) part-whole functional dependence, that is, one behavior may be a subunit of some other behavior, such as shifting gears smoothly and passing a driver's test consisting of driving in traffic with an examiner present; (b) hierarchical functional dependence, that is, one behavior is prerequisite to another or one is functionally dependent upon another, as skill in working problems in long division is dependent upon skill in multiplication; (c) environmental correlation between the behaviors, that is, cultural contingencies may be such that when one behavior is learned another is also likely to be learned, even though there is no functional connection between the two behaviors, for example, a knowledge of baseball and a knowledge of football; and (d) genetic correlation between behaviors due to common assortment of their genetic underpinnings through selection and homogamy and pleiotropism (one gene having two or more phenotypic effects). The rather low degree of correlation between our Level I and Level II tests suggests that there is little functional dependence, and this could be proved conclusively if one could find a group of subjects that reliably showed a zero correlation between Level I and Level II. The fact of quite large and significant differences in Level I–Level II correlations in various populations is also inconsistent with wholly functional or part-whole dependence as a cause of the correlation. Some substantial part of the correlation therefore, must be attributable to other causes. If the cause is common environmental influences on the Level I and Level II tests, it is hard to imagine what these influences might be and why, if they are common, there should be such large group mean differences in Level II ability and not in Level I. The most reasonable hypothesis at this point would seem to be that the correlation is due only slightly to functional dependence of

Level II upon Level I and mostly to a common genetic assortment on both factors, that is, a genetic correlation in the population between two broad classes of ability with different genetic underpinnings. If this were the case, we might find a wide range of correlations in different populations; one conceivably might even find a group in which the correlation is negative. This would tend to rule out pleiotropy and would suggest independent mechanisms under independent genetic control underlying Level I and Level II. Specially designed studies would be required to test such a hypothesis.

The test of the *c* hypothesis of hierarchical dependence of Level II upon Level I yielded significant evidence consistent with the hypothesis in the case of the nonverbal intelligence test but not the verbal. In any case, there does not appear to be evidence of any strong degree of functional dependence between the abilities; quite low or high scores on the one ability are not incompatible with a high or low score on the other, though there is a tendency for low intelligence - high memory to be more frequent than the opposite combination of abilities, especially for nonverbal intelligence.

In the present study, Level I ability was measured by three slightly differing forms of a single type of test—digit-span memory. In other studies different tests have been

used—paired-associate learning, serial learning, and free recall of pictures and objects—all with similar results generally consistent with the formulation of the two-level theory.

REFERENCES

- CATTELL, R. B. *Abilities: Their structure, growth, and action*. Boston: Houghton Mifflin, 1971.
- GUILFORD, J. P. *Fundamental statistics in psychology and education*. (3rd ed.) New York: McGraw-Hill, 1956.
- JENSEN, A. R. Patterns of mental ability and socioeconomic status. *Proceedings of the National Academy of Sciences of the United States of America*, 1968, **60**, 1330-1337.
- JENSEN, A. R. How much can we boost IQ and scholastic achievement? *Harvard Educational Review*, 1969, **39**, 1-123.
- JENSEN, A. R. Hierarchical theories of mental ability. In B. Dockrell (Ed.), *On intelligence*. Toronto, Canada: Ontario Institute for Studies in Education, 1970. (a)
- JENSEN, A. R. 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, 1970. (b)
- JENSEN, A. R. *Genetics and education*. New York: Harper & Row, 1973.
- JENSEN, A. R., & ROHWER, W. D., JR. Mental retardation, mental age, and learning rate. *Journal of Educational Psychology*, 1968, **59**, 402-403.
- MATARAZZO, J. D. *Wechsler's measurement and appraisal of adult intelligence*. (5th ed.) Baltimore, Md.: Williams & Wilkins, 1972.

(Received January 29, 1973)