INTELLIGENCE 4, 41–49 (1980)

# Level I and Level II Abilities in Asian, White, and Black Children*

ARTHUR R. JENSEN AND ARLENE R. INOUYE
*University of California, Berkeley*

All the Asian-American, white, and black children in grades 2 through 6 in a California school district were given a battery of tests including measures of IQ, scholastic achievement, and short-term memory. Factor analysis of the tests yielded two main factors identified as Level I (memory) and Level II (general intelligence) in Jensen's system. The three ethnic groups were compared with one another on uncorrelated Level I and Level II factor scores. At every grade level, bivariate means of the three groups occupy distinctly different quadrants in the factor space. Asians and whites differ on Level I (A < W) but not on Level II. Asians and blacks differ on Level II (A > B) but not on Level I. Whites and blacks differ (W > B) on both Levels I and II, but the white–blaek difference on Level I is less than one fourth as large as the white–black difference on Level II. A similar pattern of group differences is found for scores on tests of memory and nonverbal IQ. Scholastic achievement shows much smaller correlations with the Level I than with the Level II factor.

Previous studies have compared various racial, cultural, and socioeconomic groups with reference to Jensen's distinction between Level I and Level II abilities in California school populations comprised predominantly of white, black, and Mexican-American pupils (Jensen, 1973, 1974; Jensen & Figueroa, 1975). These references quite extensively define the meanings of Levels I (association) and II (transformation). Briefly, Level I involves rote learning and primary memory ability requiring minimal transformation or mental manipulation of the informational inputs prior to recall of the material; Level II involves transformation, mental manipulation, or reasoning; Level I is epitomized by the forward digit span test, Level II by the *g* factor common to all tests of general intelligence.

The present investigation extends this comparative study of Level I and II abilities to Asian-Americans (Chinese and Japanese) in a California school district comprised of approximately 10 percent Asians, 50 percent whites, and 40 percent blacks. Previous studies (listed above) have found highly significant interactions of Level I and II abilities with ethnic groups. In general, the ethnic groups (white, black, Mexican) differ from one another,

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

on the average, much less (or even hardly at all) in Level I ability than in Level II ability, on which these groups differ quite markedly, usually by about one standard deviation or more.

Asian-Americans, to our surprise, show just the opposite effect when compared with a white sample on Levels I and II. Other large-scale studies (e.g., Coleman et al., 1966) have found negligible mean differences between Asian and white pupils on tasks of scholastic aptitude and achievement, which are largely measures of Level II ability. So there was no expectation of an appreciable Asian–white difference on Level II. But there also seemed to be no reason to expect that Asian children would perform significantly below white children (with whom they are on a par scholastically) in Level I ability. However, there was the rather surprising finding that when the Wechsler Intelligence Scale for Children (WISC) was standardized in Japan, the only WISC subtest on which the Japanese standardization samples scored *lower* than the American white standardization samples was the Digit Span test. Overall, however, the Japanese scored about 6 to 7 points *higher* than the American norms on the WISC Full Scale IQ (Lynn, 1977). Since digit span memory typifies Level I ability and the WISC Full Scale IQ reflects Level II ability, the Japanese results reported by Lynn are consistent with the present findings.

## METHOD

### *Subjects*

The *S*s were all the children enrolled in grades 2 through 6 in a California school district who were present on the days the several tests were administered. Table 1 shows the numbers of white, black, and Asian-American pupils in each grade. Data on the few Mexican-Americans and

TABLE 1
Number of White, Black, and Asian-American Subjects in Each Grade

| | Total Tested | | | Complete Battery | | |
|---|---|---|---|---|---|---|
| Grade | W | B | A | W | B | A |
| 2 | 665 | 584 | 95 | 596 | 529 | 82 |
| 3 | 631 | 474 | 82 | 594 | 457 | 76 |
| 4 | 649 | 487 | 101 | 551 | 456 | 84 |
| 5 | 501 | 473 | 95 | 475 | 452 | 85 |
| 6 | 728 | 500 | 105 | 682 | 467 | 99 |
| Total | 3174 | 2518 | 478 | 2898 | 2361 | 426 |

other minorities in this school district are excluded from this study. Virtually all the *S*s are American born. The *N*s listed under Total Tested are the total numbers of pupils who were present on any (or all) of the testing days; the *N*s listed under Complete Battery are those for whom scores were obtained on all of the tests. Factor analyses were based on the Total Tested (hence slightly unequal *N*s for the various tests), while the ethnic group comparisons of mean factor scores are based only on *S*s who took the complete battery of tests (92 percent of all *S*s tested).

*S*s were tested in three one-hour sessions in intact classrooms, under highly uniform, standardized conditions, by a team of trained psychometrists.

### *Tests*

The tests are the same as those used in the other studies cited above, so the test results are directly comparable across the various studies. The following tests, which are described in detail elsewhere (Jensen, 1974), were used:

Lorge-Thorndike Intelligence Test
Stanford Achievement Tests (form appropriate for grade level)
Figure Copying Test (Grades 2, 3, 4 only)
Making Xs Test (a measure of test taking motivation)
Listening Attention Test (a test of ability to understand and comply with test instructions)
Memory for Numbers Tests:
(a) Immediate recall test
(b) Delayed recall test
(c) Learning by repetition test

The first three tests listed above were intended as Level II tests, which had been established by factor analysis in previous studies. The three Memory for Numbers Tests have been previously established as good measures of Level I ability. The Making Xs and Listening Attention tests are intended as control variables. They assess degree of effort, cooperativeness, and understanding of test instructions in the test situation. Past studies indicate that they show individual differences and slight age differences, but only negligible ethnic group differences. Including these control tests in a factor analysis helps to yield slightly more pure Level I and Level II factors.

## RESULTS

### *Factor Analysis of Tests*

First, it was established that within each ethnic group within each grade, a factor analysis (i.e., varimax rotation of the principal components having eigenvalues greater than 1) of the intercorrelations among the tests yielded two factors clearly interpretable as Level I and Level II. In all three groups at every

grade level, the three Memory for Numbers Tests had their only significant loadings (with values of .80 to .90) on one factor identified as Level I. The factor identified as Level II had its largest loadings (mostly in the .80s and .90s) on Lorge-Thorndike IQ (both Verbal and Nonverbal) and the several Stanford Achievement Tests [with highest loadings on Paragraph Meaning (reading comprehension) and Arithmetic Concepts (thought problems)].

Since the same factor structure emerged in each ethnic group at each grade level, it was justified to factor analyze the tests in the same way for the combined groups in order to derive orthogonal (i.e., uncorrelated) Level I and Level II factor scores on which to compare the groups.

### *Factor Scores*

At each grade, all the test intercorrelations in the combined ethnic groups were subjected to varimax rotation of the principal components having eigenvalues greater than 1. Included with the test variables were chronological age (in months) and age squared. Thereby, when the principal components are orthogonally rotated to approximate the varimax criterion of simple structure, age is partialed out onto a separate factor, leaving the remaining factors uncorrelated with age. By including age squared as well as age, both the nonlinear (quadratic) and linear components of any correlation that age may have with the other variables are partialed out of the two main factors of interest, viz, the Level I and Level II factors. The Making X's test produced a separate factor after rotation, which in effect partials out the Making X's variance from the other factors, leaving the Levels I and II factors free of the speed and persistence in test taking that is measured by Making Xs. [The results of the analyses as well as the test means and standard deviations for grades 5 and 6 have been presented elsewhere by the second author (Note 1).]

Factor scores on the Level I and Level II rotated factors (in the combined ethnic groups) were then derived for every subject. This was done separately within each grade, so as to yield five independent replications of the group comparisons. The factor scores are scaled to a mean of 0 and standard deviation of 1 for the combined ethnic groups within each grade. These factor scores represent "pure" (i.e., uncorrelated with each other or with other factors in this battery) measures of Level I and Level II abilities.

### *Groups' Means on Level I and Level II Factor Scores*

The main results can be seen most clearly by plotting the three ethnic groups' bivariate means of the Level I and Level II factor scores for each grade, as shown in Figure 1. The results are strikingly consistent across

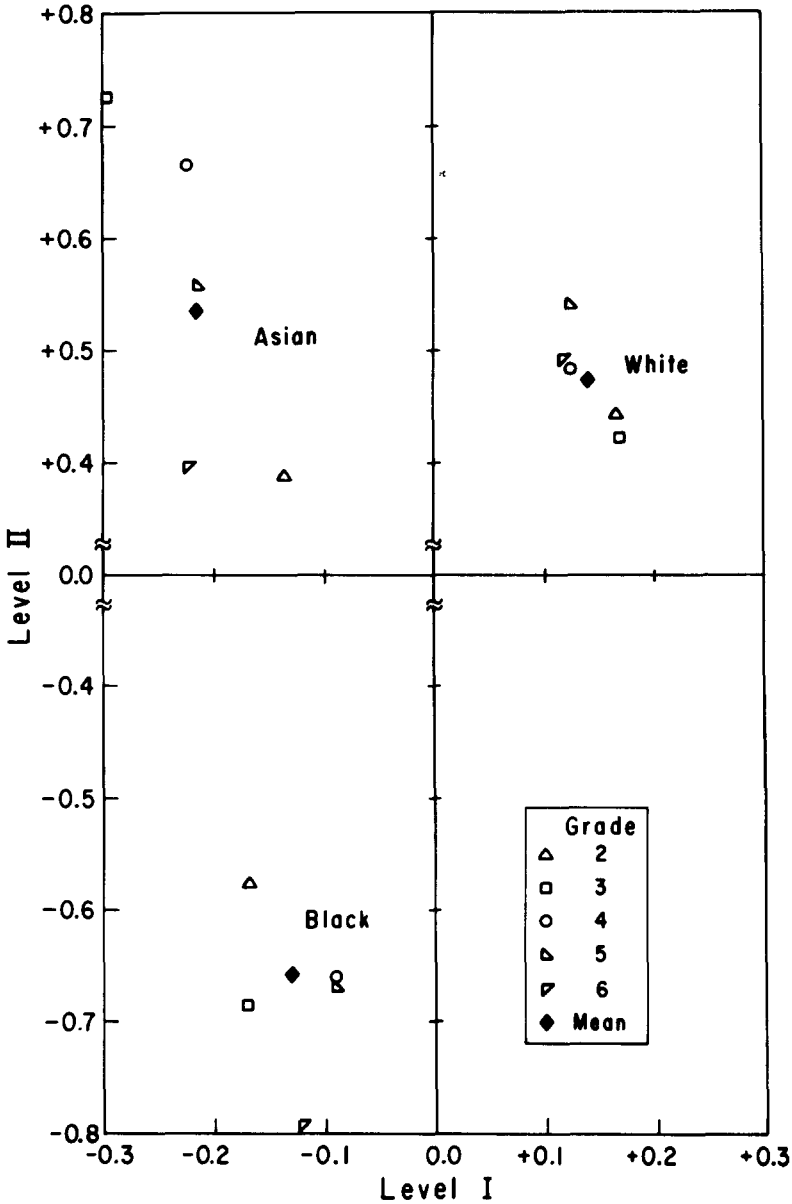


FIG. 1. Bivariate means of Level I and Level II factor scores of whites, blacks, and Asian-Americans in each of five grades. (Within each grade, the factor scores of the combined ethnic groups are scaled to a mean of 0, *SD* = 1.)

TABLE 2
Significance of Main Effect of Race in ANOVAs of Level I and Level II

| Grade | *df*[c] | *F* for Race[a] | | Contrasts[b] *Not* Significant at $p < .05$ | |
|---|---|---|---|---|---|
| | | Level I | Level II | Level I | Level II |
| 2 | 1204 | 18.92 | 205.71 | B-A | W-A |
| 3 | 1124 | 20.25 | 301.48 | B-A | * |
| 4 | 1088 | 9.03 | 334.83 | B-A | W-A |
| 5 | 1009 | 8.15 | 369.85 | B-A | W-A |
| 6 | 1245 | 11.76 | 473.55 | B-A | W-A |

[a]All *F* ratios are significant ($p < .001$).

[b]Post hoc contrasts between each pair of group means tested for significance ($p < .05$) by the Scheffé (1959) method.

[c]This is the degrees of freedom for the error term (denominator of the *F* ratio). In every case, there are two *df* for the race main effect (numerator of the *F* ratio).

*All three contrasts (W-B, A-W, A-B) are significant at $p < .01$.

grades. The bivariate means of the three ethnic groups cluster entirely in different quadrants of the factor space. The greater dispersion of the grade means for Asian-Americans is probably a result of their considerably smaller *N*s (and consequently larger sampling error) than in the white and black groups. The interesting fact, however, is that the bivariate means of Asians, whites, and blacks are very clearly separated in this Level I–Level II factor space. Asians and whites are above the general mean on Level II; and Asians and blacks are below the general mean on Level I.

The results shown in Figure 1 were tested for significance by analysis of variance, followed by post hoc Scheffé (1959) contrasts performed on each of the three pairs of groups. The ANOVAs and post hoc contrasts are summarized in Table 2. In every grade, the overall group differences are statistically significant beyond the .001 level, on both Levels I and II. The pair-wise contrasts are of greater interest. The contrasts are tested at the .05 level of significance, since with such large *N*s, mean differences which are *not* significant at the .05 level are quite small and negligible. It can be seen in Table 2 that on Level I ability the difference between blacks and Asians (B-A) is consistently nonsignificant ($p < .05$) in every grade, whereas the blacks and Asians both differ very significantly from the whites. On Level II, Asians and whites do not differ significantly ($p < .05$), except at grade 3 ($p < .01$); and at all grades Asians and whites both differ very significantly from blacks. Table 3 compares the overall mean differences (in standard deviation units) between the groups. The results are statistically quite clean cut; the mean group differences are either significant well beyond the .001 level or fall far short of significance at the .05 level. The consistency of results across the five replications (i.e., grades 2 through 6) lends further weight to the conclusion that the pattern of observed differences between these ethnic groups,

TABLE 3
Overall Mean Differences (in $\sigma$ Units) Between Groups on Levels I and II Orthogonal Factor Scores

| Contrast | I | II |
|---|---|---|
| White–black | 0.27* | 1.13* |
| White–Asian | 0.36* | –0.06 n.s. |
| Asian–black | –0.09 n.s. | 1.19* |

*$p < .001$, two-tailed test.
n.s. is nonsignificant ($p > .05$).

whatever the causes of the differences may be, represents a genuine phenomenon.

The question arises whether raw scores on typical Level I and Level II tests themselves, rather than the derived orthogonal (i.e., uncorrelated) factor scores, would present a very different picture of the group comparisons. They would present a different picture, of course, to the extent that the Level I and Level II tests are correlated with one another. Both ability factors may enter into performance on any actual test, although a given test may depend predominantly upon one or the other type of ability. To gain some idea of the effect of correlation between raw scores on typical Level I and II type tests on the group comparisons, we can compare the groups on the two tests in the present battery that are conceptually perhaps the most representative of Level I and Level II abilities: Level I—immediate recall forward memory span for digits; and Level II—Lorge–Thorndike Nonverbal IQ. The average (across-grades) Pearson correlation between these two tests in the combined ethnic groups is 0.39. Table 4 shows the overall mean differences between the ethnic groups, expressed in the units of the white group's standard deviation. Table 4 may be compared with Table 3, in which the same comparisons are made in terms of uncorrelated factor scores. It can be seen that although the absolute values of the corresponding differences vary between Tables 3 and 4, the pattern of values is similar (the Pearson correlation between the six

TABLE 4
Overall Mean Differences (in White $\sigma$ Units) Between Groups on Test of Immediate Memory Span (Level I) and Nonverbal IQ (Level II)

| Contrast | I: Memory Span | II: Nonverbal IQ |
|---|---|---|
| White–black | 0.63** | 1.52** |
| White–Asian | 0.34** | 0.11* |
| Asian–black | 0.30** | 1.40** |

*$p < .02$
**$p < .001$

corresponding pairs of values is +0.96). In both comparisons Asian-Americans differ from whites more on the Level I test (memory span) than on the Level II test (nonverbal IQ), and blacks show the opposite pattern, differing from whites much more on the Level II than on the Level I test. The identical pattern is seen for both correlated test scores (Table 4) and uncorrelated factor scores (Table 3), although the factor scores provide a more clean-cut picture of this pattern.

*Scholastic Achievement*

Another question of interest concerns the degree to which scholastic achievement is loaded on (i.e., correlated with) the Level I factor. The overall mean of the loadings of the various Stanford Achievement Tests on the Level I factor is only +0.22, as compared with their average loading of +0.85 on the Level II factor. Obviously Level I ability is a much less important source of individual and group differences in scholastic achievement than is Level II ability. The Asian-Americans, despite their lower Level I ability perform scholastically as well as whites. The differential pattern of abilities of Asians and whites, however, might account, at least in part, for the commonly observed differences in the scholastic subject matter preferences of Asian as compared with white students. A relatively larger proportion of Asians who go to college seem to prefer subjects with a high Level II (reasoning and mental manipulation) component, such as mathematics, engineering, and the exact sciences, rather than subjects that depend much more on memory and recall of verbal information, as in languages, the humanities, and social sciences.

## DISCUSSION

It would seem surprising if the present findings, which are replicated significantly in five grade levels, were not generalizable to other Asian-American, white, and black samples. We would expect the pattern of differences to be the same, although the absolute magnitudes of the mean differences between the ethnic groups might well be expected to differ somewhat from one school district to another. All three ethnic populations sampled in the present study, especially the whites and Asians, average considerably above the national norms for these groups on the standardized IQ and achievement tests. It is noteworthy that although the Asian-American population in this district is of generally lower socioeconomic status than the white population, as judged from the average occupational status of the parents, the Asian-Americans score as high on IQ and achievement tests as the whites, and average even slightly higher than whites on Level II factor scores.

Why do the Asian-American pupils score below whites and blacks on Level I factor scores? *We simply do not know.* We can only speculate. If the Level I–Level II pattern of racial differences were largely of genetic origin, the explanation would reasonable be sought in differential selective pressures for different cognitive abilities in the evolutionary histories of these groups. If the differing patterns of abilities are of cultural origin, explanations would be sought in the values, motivation, and styles of child rearing that currently predominate in each of these populations. Both avenues at present are likely sources of explanatory hypotheses and empirical exploration.

## REFERENCES

Coleman, J. S., et al. *Equality of educational opportunity.* Washington, D.C.: U.S. Office of Education, 1966.

Inouye, A. R. Differences in factored abilities of three ethnic groups. Unpublished master's dissertation. University of California, Davis, 1978.

Jensen, A. R. Hierarchical theories of mental ability. In B. Dockrell (Ed.), *On Intelligence.* Toronto: Ontario Institute for Studies in Education, 1970.

Jensen, A. R. Level I and Level II abilities in three ethnic groups. *American Educational Research Journal,* 1973, *10,* 263–276.

Jensen, A. R. Interaction of Level I and Level II abilities with race and socioeconomic status. *Journal of Educational Psychology,* 1974, *66,* 99–111.

Jensen, A. R., & Figueroa, R. A. Forward and backward digit span interaction with race and IQ: Predictions from Jensen's theory. *Journal of Educational Psychology,* 1975, *67,* 882–893.

Lynn, R. The intelligence of the Japanese. *Bulletin of the British Psychological Society,* 1977, *30,* 69–72.

Scheffé. *The analysis of variance.* New York: John Wiley, 1959.