

"CONSTANCY" OF THE IQ

BY ROBERT L. THORNDIKE

Teachers College, Columbia University

It is the purpose of this article to review recent experimental evidence on the stability of intelligence test performance by the same individual over a lapse of time. The content of the review will be limited in two ways:

(1) It is proposed to consider chiefly those reports which have appeared since 1930 or 1931. The literature previous to that time has been covered by Nemzek (75), Foran (24, 25), and others. Earlier studies will be referred to only as this becomes necessary to understand recent trends of investigation.

(2) It is proposed to limit this report primarily to those investigations in which the same individuals have been tested more than once. Although studies of the resemblance of relatives, the intelligence level of occupational groups, etc. provide evidence concerning the relative weight of hereditary, as opposed to environmental, factors in producing differences in intelligence test performance, and consequently bear on the stability of the performance of the individual, it is not our purpose to try to cover this broader literature here.

Certain general trends seem to appear in the type of experimentation reported within the last eight years, and it might be profitable to consider these briefly before going on to a survey of findings. In the period now being studied, the general fact of a rather high positive correlation between test and retest after an interval of time in children of school age seems to be fairly well taken for granted, and relatively few studies were found which were concerned merely with demonstrating this point. Most studies try to push this finding forward in one or another of several directions, of which the following are rather typical: (a) attempts to get retest information on younger and younger children; (b) efforts to get retest data over longer periods of time; (c) efforts to determine the effect upon test constancy of particular types of environmental manipulation.

GENERAL STUDIES OF TEST-RETEST CORRELATION IN ADULTS

Several studies of college students at the beginning and end of their college course have been carried out, primarily with the purpose

of studying gains during the years at college. These studies have agreed in finding that those students who remain through a college course tend to gain in intelligence test score. Test-retest correlations are also reported. Between scores on the Thorndike Intelligence Examination for High School Graduates, Part I, before and after a two-year course at normal school Masters and Upshall (63) report correlations of .78 and .80 for two groups. Wolcott (111) finds a correlation of .81 between scores as freshmen and as seniors on the Thorndike examination. McConnell (66) reports a correlation of .83 between freshman and senior scores on the American Council examination. For this same test, given before and after a four-year interval of college, Livesay (55) reports a correlation of .88 for total score and correlations of .69 to .82 for individual subtests.

Testing reformatory inmates with an average age of about 20 years on the Kuhlmann-Anderson test, Hales (32) reports a correlation of .89 when the average interval between test and retest is about five years.

Two studies have been made of small groups of adults over a 10-year interval. Davidson (17) reports a correlation of .89 for a group of 50 insurance company employees tested with a brief group intelligence test. Garrison (28) retested a group of students, whose average age at time of first testing was 25 years, with the Yerkes point scale, using a revised scoring. The correlations were .58 for a group of 32 men and .76 for a group of 41 women.

In view of the limitations placed by the reliability of the tests used and the limited range to be expected in some of the groups, these studies offer confirming evidence of the stability of intelligence test performance in adults in fairly stable surroundings, even over quite a period of years.

STUDIES OF SCHOOL-AGE CHILDREN

In the decade of the twenties, reports of correlations and mean differences between intelligence tests and retests given after various lapses of time were quite numerous for general school-age populations. In the thirties, however, relatively few such studies have appeared. Studies of too recent a date to be included in Nemzek's review (75) are reported below.

Lauderbach and Hause (47). McCall multi-mental test given to 150 students in Grades 4 to 6. Retested after an interval of 11 months. Test-retest correlation of .791.

Lincoln and Wadleigh (53). Otis Primary, National A, and

Terman A given in successive years at approximately one-year intervals. The 154 children tested were in Grade 3 at the time of the first test. The median of all the differences in IQ was reported as 7.29 points. Thirty-seven per cent of the changes were 10 points or more.

Miller (69). The data reported by Hirsch (37) were reanalyzed, using only those cases who took all tests and converting the scores on successive tests into scores with the same mean and standard deviation. The median difference was found to increase progressively from 5.4 IQ points at a one-year interval to 7.3 points at a five-year interval. The corresponding correlation dropped from approximately .85 to .79.

Lithauer and Klineberg (54). The test-retest correlation for the Stanford-Binet for 120 orphans first tested shortly after entrance to the Hebrew Orphan Asylum, and retested after a median interval of 14 months, is reported to be .76.

Traxler (102). Otis Self-Administering IQ's were obtained for students at the University of Chicago High School. Retests were obtained after intervals of one, two, and three years. Correlations were .756, .673, and .676, respectively. For the same students, the retest correlation for the Binet with an interval of one to four years is reported to be .762.

Seagoe (87). Tests were given to school children in Grades 1, 3, 5, and 7. Tests were Detroit First-Grade, Detroit Primary, National Intelligence, and Terman Group, respectively. Retests were available for those children who remained in the system long enough to be caught by two or more tests. Successive tests were at two-year intervals. Correlations were reported as follows: Detroit First-Grade *vs.* Detroit Primary, .642; Detroit Primary *vs.* N.I.T., .700 and .727; N.I.T. *vs.* Terman Group, .797 and .869. Over a four-year interval, Detroit Primary correlated with Terman Group .765.

Lorge (57). About 160 boys, a random sample of 860 who had been tested in the eighth grade, were retested after an interval of a little over 10 years. Both raw correlations and correlations corrected for estimated test unreliability were presented. These were:

	Raw	Corrected
Thorndike-McCall Reading Scale.....	.57	.74
I.E.R. Arithmetic Test.....	.60	.79
Stenquist Assembly Test.....	.66	1.06
I.E.R. General Clerical Test.....	.63	.75

The results of these studies fit in quite well with the general trend of earlier investigations, and the size of the correlations reported here

is comparable to the average for the Stanford-Binet and for group tests reported in Nemzek's summary (75).

The somewhat lower correlations reported in the study by Lorge (57) direct attention to the factor of interval between test and retest as one factor influencing the stability of performance from test to retest. This has been considered by R. L. Thorndike (101) in a synthesis of previous studies. Putting together the data from those studies in which the interval between test and retest was fairly uniform, a curve was fitted to all the data by the method of least squares. The correlation for an immediate retest, estimated from this curve, falls close to .90, and the correlation for a test after a five-year interval is estimated to be about .70. Further data on the question of time interval is contributed by the study of R. R. Brown (8). In one comparison of 58 cases who had a first retest after an interval of less than two years and a second retest after an interval of five years or more, he found the correlation between initial test and first retest to be .86 and between initial test and second retest to be .61. It seems apparent that one factor which influences the accuracy of prediction from an intelligence test at school age is the length of time over which we endeavor to predict.

STUDIES OF GIFTED CHILDREN FIRST TESTED AT SCHOOL AGE

The performance upon adult tests at maturity of gifted children first tested while in school has been studied by Hollingworth and Kaunitz (39) and by Lorge and Hollingworth (58). In the first of these studies, 116 children who had received Binet IQ's of 135 or higher when they were about eight years old were tested with the Army Alpha approximately 10 years later. Of these children, who had been chosen to represent the top centile of the child population, 82% scored in the top centile on the Army Alpha according to the Army norms. The remainder were in the high centiles. No individual regressed to, or nearly to, average. Lorge and Hollingworth report additional test material on a group of 21 individuals who, as children, had a median IQ of 168. On the CAVD intelligence test, 20 of the 21 were found to surpass the median of law school freshmen. Those who had, as children, received IQ's of about 140 were found to define approximately the seventy-fifth percentile of college graduates, taking it the country over.

Whether the IQ's of gifted children tend to increase or decrease has been studied by Lincoln (49, 50, 51, 52) and by Cattell (10).

Lincoln finds that, when children are classed on the basis of their initial test, those with high initial tests tend to decline. Cattell finds that, when children are classified on the basis of the average of their initial and final tests, children in the highest classifications tend to have higher final tests than initial tests. These results are in no sense contradictory, the first illustrating the universal tendency of any fallible measure to regress toward the mean of the sample upon repetition, and the second being evidence of a tendency for the variability of IQ's in the population to increase with increasing age.

Nemzek (73, 74) analyzed data previously reported by Carroll and Hollingworth and by Lamson, and found correlations between Stanford-Binet test and retest with an interval of one or two years of .53 to .72. For data on the Herring-Binet the correlation was .73. For the Stanford-Binet the mean change ranged from 7.85 to 9.37 points, and for the Herring-Binet it was 9.06 points. Nemzek concluded that the IQ was more variable for gifted children, which is in accord with other results on this question. McNemar (67) has pointed out that the greater variability of high IQ's is inherent in the use of a ratio and appears as an inevitable consequence if the variability in performance at a given mental age level is constant.

RETARDED OR OTHERWISE HANDICAPPED GROUPS

Studies of the mentally retarded by Arthur (2), Engel (21), Hoakley (38), Parker (77, 78), and Woodall (112) confirm, in general, previous findings of a progressive decline in the IQ's of these groups. Woodall finds an increase in IQ for retarded individuals after the age of 16, indicating that mental growth has not completely stopped at that age. The greater stability of the IQ for defective groups again becomes apparent in these studies. As mentioned above, it must be remembered that this is implicit in the use of such a ratio as the IQ.

Schott (84) reports that adult neuropsychiatric cases show markedly greater variation between tests than do normal children.

Miller (68) suggests that a pseudointellectual deficiency may be produced by emotional maladjustment, and that this is remediable with careful treatment.

Arthur (3) finds that "an examination with the Kuhlmann-Binet given by an experienced psychologist to a kindergarten, first, or second grade child from a non-English-speaking home can yield a rating with a high degree of reliability and predictive value as measured by achievement on the Kuhlmann-Anderson 5 to 7 years later,

if the child has had as much as a year in the English-speaking school environment."

Gildea and Macoubrey (30) and O'Neill (76) have endeavored to analyze the factors associated with large changes in IQ between test and retest. Lack of a control group in one study, and limited statistical analysis in both, preclude any definite conclusions. The pattern of factors involved seems to be quite complex in any event.

PREDICTIVE VALUE OF INFANT AND PRESCHOOL TESTS

Following upon the success of the intelligence tests for school-age children, efforts have been made to develop tests for younger and younger children. The years covered by this review have witnessed a growing crop of studies of the constancy and predictive value of indices derived from these early tests. Evidence on the predictive value of tests at these early years has been presented by Bayley (6), Cunningham (15), Driscoll (19), Furfey and Muehlenbein (27), Gesell and others (29), Hallowell (33), Herring (35), Honzik (40), Hubbard (41), Kawin (43), Mowrer (70), Nelson and Richards (71, 72), Stutsman (98), Symmes (99), Updegraff (103), and Wellman (105, 106, 107). These studies cover a number of different tests and differ quite a bit in the age and other characteristics of the population studied. Consequently, any effort to lump the studies together and generalize from them may well be questioned. However, we have tried to do this in a rough sort of way. Taking the test-retest correlations from those studies in which they were available, and in which the tests had been given at fairly definitely specified ages, it was possible to tabulate them according to age at first test and interval between test and retest. This has been done with correlations from the studies by Bayley (6), Cunningham (15), Driscoll (19), Furfey and Muehlenbein (27), Herring (35), Honzik (40), Kawin (43), and Nelson and Richards (71, 72), and from the earlier study by Goodenough (31). A table was prepared, and each correlation coefficient was entered in the appropriate cell. The average correlation coefficient in each cell was then determined, as a rough estimate of the degree to which it is possible to predict over that interval at that age. These average coefficients are presented in Table I.

In any interpretation of Table I the diversity of tests studied,

the diversity of final tests being predicted, and variations in the range of ability in the groups being studied must be borne in mind. Furthermore, no effort was made to apply differential weights, taking into account the different sizes of populations studied, to the correlations upon which Table I is based. Finally, some of the figures reported in Table I are based upon only one correlation or the work of only one experimenter, while others are based upon a number. However, it seems clear from these results that the adequacy of prediction of future mental development is a combined function of the age at which the test is given and of the length of time over which

TABLE I

COMPOSITE OF TEST-RETEST CORRELATIONS FROM SEVERAL STUDIES OF INFANT AND PRESCHOOL GROUPS

Age at Earlier Test	Interval Between Test and Retest							
	Less than 4 mos.	4-9 mos.	10-15 mos.	16-21 mos.	22-29 mos.	30-41 mos.	42-53 mos.	Over 53 mos.
Under 4 mos.	.57	.33	.10	-.03	-.09			
4-9 mos.	.77	.53	.49	.23	.16	.46	.00	
10-15 mos.	.78	.66	.50	.45	.33			.55
16-21 mos.	.76	.68	.51	.44	.38	.41	.25	.33
22-29 mos.	.82	.74	.68					.43
30-41 mos.	.87	.68	.66	.49	.57	.57	.56	.66
42-53 mos.	.81	.65	.72	.71	.66	.63	.63	.41
54-65 mos.			.76		.73			

we try to predict. Before the age of four months, it seems impossible to get satisfactory correlations between tests given no more than a month apart. From this point on, however, correlations of .75 and better seem to be obtainable for such short-time forecasting. But it is not until the individual is about four years old that predictions can be made over a span of as long as a year with a correlation of .70 or better. It does not seem possible to make very reliable *long-time* predictions from any of these early tests.

Honzik (40), working with tests between the ages of 21 and 84 months, suggested that the size of correlation might be thought of as a linear function of age at the time of the first test divided by age at the time of the second test. She found a correlation of .92 for one group and a correlation of .78 for another between this ratio and the correlation coefficient between test and retest. Apparently, within the range of ages and intervals that she studied, the relation did hold in an approximate sort of way. Whether it would hold as well over a greater range of ages is, of course, open to question. English and

Killian (22) present some evidence to suggest that age is not a factor in determining the amount of variation between tests between the ages of 4 and 16.

At least two alternative explanations of the limited value of these early tests in predicting school-age intelligence test performance may be suggested. On the one hand, it may be argued that at this early age the individual is very susceptible to environmental influences, and so his intellectual development is likely to be markedly accelerated or retarded, depending upon the environment to which he is exposed. Low correlations between early and late tests would be thought of as due to the differential effect of different environments. On the other hand, it may be suggested that the aspects of mental functioning studied by early tests are rather different from those incorporated in school-age tests. This seems at least superficially true, as we consider the test materials. It may be that in the early tests we are not testing quite the same type or aspect of "intelligence" that we test in school children, possibly because the type of function which we study in our later tests is not as yet present or susceptible to testing. The low correlations would, in this case, be explained as due to a considerable shift in the aspect of intellectual functioning which we were testing.

EFFECT UPON INTELLIGENCE OF CHANGE IN PHYSICAL ENVIRONMENT

Studies of the effect of diet upon learning and intelligence are reviewed by Fritz (26), who concludes that "contrary to what would be generally supposed, there is very little experimental evidence at present to indicate that diet markedly affects intelligence or capacity to learn." Fritz indicates that the results of Maurer and Tsai, showing the effect of vitamin B deficiency upon maze learning in rats, represent an exception to this generalization.

Balken and Maurer (5) report a preliminary study of the effect of increasing the vitamin B intake of malnourished children. A variety of mental tests were given to 46 children from homes of very low economic status in which vitamin B deficiency was probable. The children were given supplementary feedings of vitamin B for 15 weeks, and tests were repeated at the end of this time. Gains were found upon each test, but these were generally small. No control group was used, so it is impossible to determine whether there would have been any residual gain after practice effects had been eliminated.

Poull (80) matched 41 pairs of records from the testing at New

York City Children's Hospital. One member of each pair was originally malnourished, as judged by the physician's statement, and subsequently became well nourished. The control group was uniformly well nourished. The group whose nutritional status improved showed an average gain in IQ of approximately 10 points, while the control group showed no change.

Studies of the effect of diseased tonsils and adenoids by Richey (81) and of intestinal protozoa by Schell (83) gave negative results.

Mateer (64) finds evidence that gains may result from a combination of glandular and educational treatment in hypopituitary cases.

Dawson (18) reports small, but reliable, gains in IQ for a group removed from a slum to a housing project.

EFFECT OF CHANGED SOCIAL OR EDUCATIONAL ENVIRONMENT

We must next consider various investigations which introduce certain special factors into the social or educational life of the child. Familiar patterns, occurring during the period covered by this survey, are (1) placement in a foster home, (2) placement in some sort of institution, (3) experience in a nursery school, (4) exposure to some special type of schooling in the regular grades, either in the general character of the school or through some special type of instruction.

To the earlier foster-child studies there have been added during this period studies by Hinton (36), Leahy (48), Lithauer and Klineberg (54), Schott (85), Skodak (95, 96) [the report by Skodak is taken as most completely representing the Iowa investigation, parts of which have also been reported by Skeels (90, 91, 92)], and Wells and Arthur (109). The last four of these authors report average changes from before placement in a foster home to some time after placement. The studies by Lithauer and Klineberg, Schott, and Skodak agree in finding an average rise of about six points in IQ for retests given within a year or two of placement. In the Skodak study there was some indication of further gain upon a subsequent retest.

The study by Skodak centered its attention upon the intelligence of a group of children (mostly known to be illegitimate) who were placed in foster homes before the age of six months. Each of these children received two tests, the first at an average age of two years, seven months, and the second at an average age of four years, four months. The average IQ was 116.0 at the time of the first test and

111.5 at the time of the second test. The correlation between these two tests, with an average interval of 21 months, is reported as .56. This corresponds quite closely to correlations reported in other studies for this age and interval, as summarized in our Table I.

The average level of IQ at the time of the second test is in close agreement with results reported by Leahy (48) for somewhat older foster illegitimate children. Leahy reports a mean IQ of 110. What judgment we shall make as to the contribution of the improved foster environment to this average IQ depends upon a judgment as to the IQ expectancy from the true parents. For some of her cases Skodak presents data on true-father's occupation, true-parent's education, and true-mother's Binet IQ. Occupationally, the group is clearly low, though just what that means for intelligence in these times is not as clear as it would have been a decade or more ago. The mean mid-parent education is 10 grades, which seems fairly substantial. The mean IQ of mother is reported as 88.7. This, however, is derived from the Stanford-Binet, with 16 years as the divisor, and is established for a group of unspecified age, tested at an unspecified interval after their school experience. An assumption that the genetic expectancy of the foster group was below 100 IQ seems to be open to question—especially as the preplacement IQ's of the group of children tested before and after placement (mentioned earlier), who represented a poorer sample educationally and occupationally, had an average of 98.5.

The relationship of foster-child intelligence to characteristics of the true and foster parents is somewhat vitiated in this study by selective placement. There was a correlation of .30 between mid-parent education for true and foster parent. This indication of selection makes any interpretation of relationship of child intelligence to either characteristics of true parents or characteristics of foster parents somewhat hazardous.

Wells and Arthur (109) studied children one or both of whose parents were classed as feeble-minded. They compared the course of intellectual development for children left in their own homes with that of children placed in certified foster homes. Over a period of about five years the differential between these two groups amounted to about eight IQ points in favor of those in foster homes. For a rather small group, first tested before the age of five years, the differential was approximately 16 points.

Rather in contrast with the other findings, Hinton (36) reports that children transferred to Mooseheart, which is judged to present a

social environment greatly superior to the one these children had previously known, show no reliable change in IQ over a period of five years subsequent to entrance. However, the younger children are reported to gain more than the older ones.

Change in IQ during residence in an orphanage or during residence in an institution for the feeble-minded was studied by Crissey (13, 14). The finding which is emphasized in this study, as in a number of the other studies from the University of Iowa, is that the extreme members of the group tend to move toward the average of the group upon repeated tests. There does not seem to be sufficient appreciation, however, of the significance of errors of measurement in producing this change. It must be borne in mind, as one reads any of these studies, that those members of any group who are at the top of the group in score on a fallible measure, whether the scores be high or low in terms of some more extensive total population, fall at the top of the group, in part because of a preponderance of plus errors of measurement, and will tend, in the absence of other influences, to drop back toward the mean of the group in any test that is not perfectly correlated with the first test. The converse is, of course, true of those at the bottom.

Crissey also finds a suggestion that those who would be expected to have constant IQ's in an orphanage tend to show a drop when transferred to the home for feeble-minded, and that those who would be expected to drop in the home for feeble-minded fail to do so when transferred to the orphanage. This is interpreted as being due to the difference in stimulation value of the two environments. One must bear in mind, however, the possibility that other factors, which were a symptom of a dropping IQ, caused the transfer to the asylum, or that other behavior, indicating intelligence superior to that demonstrated on the test, may have been a factor in transferring children from the asylum to the orphanage.

A number of studies have brought out the retarding effect of an impoverished environment by comparing IQ's of younger and older children in the same family or district. During the period of this review, studies based upon such data were reported by Asher (4), Ludeman and McAnelly (60), Sherman and Key (88), Skeels and Fillmore (93), and Wheeler (110). Since they fall somewhat outside the focus of interest of this review, no further mention will be made of them.

To the earlier studies of the effect of nursery school training, the period here being reviewed sees the addition of the series of reports

by Wellman and others (12, 105, 106, 107, 108) of children in the University of Iowa preschool, the study by Skeels, Updegraff, Wellman, and Williams (94) of a nursery school project in an Iowa orphanage, and the report by Kawin and Hoefer (44) of a nursery school project near Chicago.

The findings from the Iowa studies may be summarized somewhat as follows:

(1) The Binet IQ of children from generally superior homes rose markedly during a period in nursery school, but did not rise during the summer spent in the general home environment (105).

(2) Performance on the Merrill-Palmer test showed some residual gain, after nursery school attendance, over and above apparent adaptation effects, but the influence was not as marked as for the Binet (108).

(3) The gains in Binet IQ were maintained by a sample of children located and tested after several years of attendance at other than University schools (106).

(4) The gains in Binet IQ were further added to by a sample of children who remained in the University schools and were tested at a later time (106).

(5) Length of attendance at the University schools was related to intelligence test score in high school and at college entrance (107).

(6) Gain from attending preschool was not related to the occupational level of the parent (12).

(7) The greatest gain in preschool was for those who originally received the lowest scores, and the smallest gain, for those who received the highest (105).

(8) In an orphanage preschool, attendance of 200 days or more resulted in some gain in Binet IQ, whereas a control group exposed to the general orphanage environment for this same length of time showed some loss in IQ (94).

The study by Kawin and Hoefer (44) failed to confirm the gains found in the Iowa studies. It was found that the gains in Merrill-Palmer score made by the nursery school group and the paired control group were substantially identical. These gains were attributed to practice.

There seems to be little question as to the genuineness of the immediate gains in Binet IQ in the Iowa preschool studies. However, before we accept the conclusion that gains so produced are permanent and lasting, it might be well to consider certain other explanations of their apparent permanence. The permanence of the

gains is inferred from conclusions (3) and (4) above. Before the lasting rise is definitely assigned to the nursery school experience, there are three other possible factors which should be considered:

(1) There is a possibility of selection. If we admit that there is a considerable element of chance in the score which a young child will make upon an intelligence test, we must agree that those children who test at 100 IQ, for example, cover quite a range in underlying "true" ability. Some have been overestimated in our testing and some underestimated. If there is a tendency for those who have been overestimated—whose "true" IQ is below 100 and who will consequently tend to drop on later tests—to be eliminated from the population which is retested, we may expect the fraction of the original population to whom we give retests to gain in IQ.

(2) Though these children did not gain between sessions of nursery schools, there is still the possibility that the sudden gain during nursery school represents a short-circuiting of a gain that would have appeared gradually over the years as a result of a generally superior home environment. That is, if the people who send their children to nursery school are people who will provide generally good and stimulating homes for their children, the sudden gains which appeared during nursery school and were subsequently maintained might have appeared as a slow cumulative development in the home environment. This would assume a temporary nursery school effect, sudden and gradually wearing off, balanced by a more permanent home effect, gradually making itself manifest.

(3) It may be that children from the genetic background represented by the University nursery school groups tend to do better on the type of intelligence tested at the older ages than on the type tested in the first few years. In this case, then, the permanent change would be explained as due to a change in the character of the test with the older children to a test upon which they were, by genetic constitution, more extreme deviates than they were upon the early tests.

It must be admitted that we do not have much evidence to back up any of these suggestions. However, it does not seem that they are excluded as possibilities.

Peterson's study (79), which compares a small group of nursery school children with another small group of non-nursery-school children as they enter and go through kindergarten, may be cited as suggesting that the intellectual increment for nursery school children at kindergarten age is not great. Only a small and diminishing superiority was found for the nursery school group when they were

compared with non-nursery-school children of comparable socio-economic status.

A rather vitriolic critique of Wellman's conclusions has been published by Simpson (89). Specific criticisms are directed primarily at the first of Wellman's publications (104).

Lamson (46) endeavors to determine to what extent intelligence quotients are increased by children who participate in what is described as "a vital curriculum—one that considers children's interests and capacities, that requires activity and self-direction on the part of children, and that makes progress at the child's own optimum rate possible." No evidence of gain is found in this school.

Studies by Hawthorne (34), by Lowry (59), and by Scruggs (86) investigate the effect of special training in reading upon intelligence test performance. Hawthorne found that pupils from 5 to 12 years of age, who were of average intelligence, but retarded in reading, and who had improved in reading during remedial instruction at twice the normal rate, showed no corresponding improvement on a group intelligence test. Lowry, on the other hand, found that 50 children who were given three months of intensive reading drill, which produced an average gain of 1.36 grades in five reading tests and 2.72 grades in two speed-of-reading tests, showed a gain from a pre-test on Form A to a post-test on Form B of the Otis Intermediate Examination of 11.76 points in IQ. There was no control group to take care of practice effects or possible inequality of standardization of the two tests.

Scruggs (86) also finds gains in intelligence from special reading instruction. A group of fifth-grade negro children who "followed an especially prepared course in reading involving intensive and extensive work . . . exercises in vocabulary development, following directions, fact getting, recognition of central thought, organization and summarization, rate and speed of reading and in manipulation" gained more than a control group on a variety of verbal and non-verbal intelligence tests. There is reported to be some residual gain a year and two years later.

Durrell (20) finds that children whose reading ability is better than would be expected from their Binet IQ do better on paper-and-pencil intelligence tests than on the Binet.

CONDITIONS OF TESTING

Finally, we shall make brief mention of a number of studies which indicate the effect of certain conditions of testing upon the constancy of the resulting IQ's.

Madsen (61) and Jordan (42) report on the reliability of Binet tests given by student examiners, reporting test-retest correlations ranging from .65 to .84.

Mayer (65) and Rust (82) study the effect of negativism upon the scores of young children.

Adkins (1), Dave (16), Lämmermann (45), and Snedden (97) report evidence of practice effects where the same test is repeated or another test is given after a brief interval.

Benton (7), Ferguson (23), and Maller and Zubin (62) report studies in which special incentives were introduced into the testing situation. In no case did the special incentives result in significant increases in score, as compared to a control group.

Cattell (11) gives evidence of the variation in standard between different trained examiners.

Canady (9) reports a small-scale investigation of the effect of race of examiner upon test performance, comparing the performances of white and negro children tested by white and negro examiners.

Lodge (56) reports evidence of seasonal fluctuation in IQ. Analyzing results for a group of children living in their own home environment, tested at six-month intervals, he finds that tests given between November 1 and April 30 tend to be higher than those given during the other half of the year. It is suggested that some of the apparent preschool effect may be due to such a seasonal variation.

BIBLIOGRAPHY

1. ADKINS, D. C. The effects of practice on intelligence test scores. *J. educ. Psychol.*, 1937, 28, 222-231.
2. ARTHUR, G. The prediction value of the Kuhlmann-Binet scale for inmates of a State school for the feeble-minded. *J. appl. Psychol.*, 1933, 17, 188-194.
3. ARTHUR, G. The predictive value of the Kuhlmann-Binet scale for a partially Americanized school population. *J. appl. Psychol.*, 1937, 21, 359-364.
4. ASHER, E. J. The inadequacy of current intelligence tests for testing Kentucky mountain children. *J. genet. Psychol.*, 1935, 46, 480-486.
5. BALKEN, E. R., & MAURER, S. Variations in psychological measurements associated with increased vitamin B complex feeding in young children. *J. exp. Psychol.*, 1934, 17, 85-92.
6. BAYLEY, N. Mental growth during the first three years. *Genet. Psychol. Monogr.*, 1933, 14, 1-92.
7. BENTON, A. L. Influence of incentives upon intelligence test scores of school children. *J. genet. Psychol.*, 1936, 49, 494-497.
8. BROWN, R. R. The time interval between test and re-test in its relation to the constancy of the intelligence quotient. *J. educ. Psychol.*, 1933, 24, 81-96.

9. CANADY, H. G. The effect of "rapport" on the I.Q.: a new approach to the problem of racial psychology. *J. Negro Educ.*, 1936, 5, 209-219.
10. CATTELL, P. Do the Stanford-Binet IQ's of superior boys and girls tend to decrease or increase with age? *J. educ. Res.*, 1933, 26, 668-673.
11. CATTELL, P. Stanford-Binet IQ variations. *Sch. & Soc.*, 1937, 45, 615-618.
12. COFFEY, H. S., & WELLMAN, B. L. The role of cultural status in intelligence changes of preschool children. *J. exp. Educ.*, 1936, 5, 191-202.
13. CRISSEY, O. L. Mental development as related to institutional residence and educational achievement. *Univ. Ia Stud. Child Welf.*, 1937, 13, No. 1.
14. CRISSEY, O. L. The mental development of children of the same IQ in differing institutional environments. *Child Developm.*, 1937, 8, 217-220.
15. CUNNINGHAM, B. V. Infant IQ ratings evaluated after an interval of seven years. *J. exp. Educ.*, 1934, 3, 84-87.
16. DAVE, K. J. Effects of practice upon intelligence tests. *Brit. J. educ. Psychol.*, 1938, 8, 313.
17. DAVIDSON, C. M. Permanency of mental alertness measures. *J. appl. Psychol.*, 1938, 22, 307-310.
18. DAWSON, S. Environmental influence on mentality. *Brit. J. Psychol.*, 1936, 27, 129-134.
19. DRISCOLL, G. P. The developmental status of the preschool child as a prognosis of future development. *Child Developm. Monogr.*, 1933, No. 13.
20. DURRELL, D. D. Influence of reading ability on intelligence measures. *J. educ. Psychol.*, 1933, 24, 412-416.
21. ENGEL, A. M. Constancy of I.Q.; experiments with mental defectives. *Nation's Sch.*, 1937, 19, 19-21.
22. ENGLISH, H. B., & KILLIAN, C. D. The constancy of the IQ at different age levels. *J. consult. Psychol.*, 1939, 3, 30-32.
23. FERGUSON, H. H. Incentives and an intelligence test. *Aust. J. Psychol. Phil.*, 1937, 15, 39-53.
24. FORAN, T. G. The constancy of the intelligence quotient; a review. *Cath. Univ. Amer. educ. Res. Bull.*, 1926, 1, No. 10.
25. FORAN, T. G. A supplementary review of the constancy of the intelligence quotient. *Cath. Univ. Amer. educ. Res. Bull.*, 1929, 4, No. 9.
26. FRITZ, M. F. The effect of diet on intelligence and learning. *Psychol. Bull.*, 1935, 32, 355-363.
27. FURFEY, P. H., & MUEHLENBEIN, J. The validity of infant intelligence tests. *J. genet. Psychol.*, 1932, 40, 219-223.
28. GARRISON, S. C. Retests on adults at an interval of ten years. *Sch. & Soc.*, 1930, 32, 326-328.
29. GESELL, A., AMATRUDA, C. S., CASTNER, B. M., & THOMPSON, H. Biographies of child development. New York: Hoeber, 1939.
30. GILDEA, H., & MACOUBREY, C. Factors affecting the constancy of the intelligence quotients of problem children. *Smith Coll. Stud. soc. Work*, 1933, 3, 229-247.
31. GOODENOUGH, F. L. The Kuhlmann-Binet tests for children of school age: a critical study and evaluation. Minneapolis: Univ. Minnesota Press, 1928.
32. HALES, W. M. Results of a group intelligence re-test with a reformatory group. *J. juv. Res.*, 1937, 21, 181-187.

33. HALLOWELL, D. K. Stability of mental test ratings for preschool children. *J. genet. Psychol.*, 1932, 40, 406-421.
34. HAWTHORNE, J. W. The effect of improvement in reading ability on intelligence test scores. *J. educ. Psychol.*, 1935, 26, 41-51.
35. HERRING, A. An experimental study of the reliability of the Bühler baby test. *J. exp. Educ.*, 1937, 6, 147-160.
36. HINTON, R. T., JR. The effect of a constant and superior environment upon the I.Q.'s of one hundred and seventeen children. (Abstract.) *Psychol. Bull.*, 1938, 35, 681.
37. HIRSCH, N. D. M. Experimental study upon three hundred school children over a six-year period. *Genet. Psychol. Monogr.*, 1930, 6, 487-549.
38. HOAKLEY, Z. P. The variability of intelligence quotients. *Proc. Amer. Ass. Stud. Feeblemind.*, 1932, 37, 119-146.
39. HOLLINGWORTH, L. S., & KAUNITZ, R. M. The centile status of gifted children at maturity. *J. genet. Psychol.*, 1934, 45, 106-120.
40. HONZIK, M. P. The constancy of mental test performance during the preschool period. *J. genet. Psychol.*, 1938, 52, 285-302.
41. HUBBARD, R. M. A study of the reliability and validity of the Bühler infant scale. *J. genet. Psychol.*, 1935, 47, 361-384.
42. JORDAN, J. S. Reliability of Stanford-Binet intelligence quotients derived by student examiners. *J. educ. Res.*, 1932, 26, 295-301.
43. KAWIN, E. Children of preschool age. Chicago: Univ. Chicago Press, 1934.
44. KAWIN, E., & HOEFER, C. A comparative study of a nursery-school versus a non-nursery-school group. Chicago: Univ. Chicago Press, 1931.
45. LÄMMERMANN, H. Die Konstanz die Uebbarkeit von Denkleistungen. *Z. angew. Psychol.*, 1934, 46, 3-87.
46. LAMSON, E. E. To what extent are intelligence quotients increased by children who participate in a rich, vital school curriculum? *J. educ. Psychol.*, 1938, 29, 67-70.
47. LAUDERBACH, J. C., & HAUSE, E. On the reliability and validity of derived scores yielded by the McCall multi-mental scale. *J. appl. Psychol.*, 1932, 16, 322-323.
48. LEAHY, A. M. Nature-nurture and intelligence. *Genet. Psychol. Monogr.*, 1935, 17, 236-308.
49. LINCOLN, E. A. Preliminary report on the Stanford Binet IQ changes of superior children. *J. exp. Educ.*, 1933, 1, 287-292.
50. LINCOLN, E. A. The Stanford Binet IQ changes of superior children. *Sch. & Soc.*, 1935, 41, 519-520.
51. LINCOLN, E. A. A study of changes in the intelligence quotients of superior children. *J. educ. Res.*, 1935, 29, 272-275.
52. LINCOLN, E. A. Stanford-Binet IQ changes in the Harvard growth study. *J. appl. Psychol.*, 1936, 20, 236-242.
53. LINCOLN, E. A., & WADLEIGH, V. L. Change and constancy in group test I.Q.'s. *Sch. & Soc.*, 1932, 35, 606-608.
54. LITHAUER, D. B., & KLINEBERG, O. A study of the variation in IQ of a group of dependent children in institution and foster home. *J. genet. Psychol.*, 1933, 42, 236-242.
55. LIVESAY, T. M. Does test intelligence increase at the college level? *J. educ. Psychol.*, 1939, 30, 63-68.

56. LODGE, T. Variation in Stanford-Binet IQ's of pre-school children according to the months in which examinations were given. *J. Psychol.*, 1938, **6**, 385-395.
57. LORGE, I. Retests after ten years. *J. educ. Psychol.*, 1934, **25**, 136-141.
58. LORGE, I., & HOLLINGWORTH, L. S. Adult status of highly intelligent children. *J. genet. Psychol.*, 1936, **49**, 215-226.
59. LOWRY, E. Increasing the I.Q. *Sch. & Soc.*, 1932, **35**, 179-180.
60. LUDEMAN, W. W., & McANALLY, J. R. Intelligence of colony people. *J. educ. Psychol.*, 1930, **21**, 612-615.
61. MADSEN, I. N. Reliability and validity of Stanford-Binet tests when administered by student examiners. *J. educ. Res.*, 1934, **28**, 265-270.
62. MALLER, J. B., & ZUBIN, J. The effect of motivation upon intelligence test scores. *J. genet. Psychol.*, 1932, **41**, 136-151.
63. MASTERS, H. V., & UPSHALL, C. C. Study of the gains made by normal-school students in intelligence test scores. *J. educ. Res.*, 1934, **27**, 446-452.
64. MATEER, F. Pituitary acceleration of IQ's. (Abstract.) *Psychol. Bull.*, 1934, **31**, 686.
65. MAYER, B. A. Negativistic reactions of pre-school children on the New Revision of the Stanford-Binet. *J. genet. Psychol.*, 1935, **46**, 311-334.
66. McCONNELL, T. R. Change in scores on the psychological examination of the American Council on Education from freshman to senior year. *J. educ. Psychol.*, 1934, **25**, 66-69.
67. McNEMAR, Q. On the inadequacy of the reliability coefficient when used with I.Q. data. (Abstract.) *Psychol. Bull.*, 1936, **33**, 779-780.
68. MILLER, E. A. Emotional factors in intellectual retardation. *J. ment. Sci.*, 1933, **79**, 614-625.
69. MILLER, W. S. Variation of IQ's obtained from group tests. *J. educ. Psychol.*, 1933, **24**, 468-474.
70. MOWRER, W. M. C. Intelligence scales for preschool children. *Child Develpm.*, 1933, **4**, 318-322.
71. NELSON, V. L., & RICHARDS, T. W. Studies in mental development: I. Performance on Gesell items at six months and its predictive value for performance on mental tests at two and three years. *J. genet. Psychol.*, 1938, **52**, 303-325.
72. NELSON, V. L., & RICHARDS, T. W. Studies in mental development: III. Performance of twelve-months-old children on the Gesell schedule, and its predictive value for mental status at two and three years. *J. genet. Psychol.*, 1939, **54**, 181-191.
73. NEMZEK, C. L. Is the I.Q. constant? *Peabody J. Educ.*, 1931, **9**, 123-124.
74. NEMZEK, C. L. The constancy of the IQ's of gifted children. *J. educ. Psychol.*, 1932, **23**, 607-610.
75. NEMZEK, C. L. The constancy of the I.Q. *Psychol. Bull.*, 1933, **30**, 143-167.
76. O'NEILL, H. O. Variations in the intelligence quotients of 105 children. *Child Develpm.*, 1937, **8**, 357-363.
77. PARKER, H. T. Fluctuations in the intelligence quotients of subnormal children. *J. juv. Res.*, 1934, **18**, 163-168.
78. PARKER, H. T. The development of intelligence in subnormal children. *Aust. Coun. educ. Res. Ser.*, 1934, No. 27.

79. PETERSON, T. J. A preliminary study of the effects of previous nursery school attendance upon five-year-old children entering kindergarten. *Univ. Ia Stud. Child. Welf.*, 1938, 14, 197-248.
80. PULL, L. E. The effect of improvement in nutrition on the mental capacity of young children. *Child Developm.*, 1938, 9, 123-126.
81. RICHEY, A. The effects of diseased tonsils and adenoids on intelligence quotients of 204 children. *J. juv. Res.*, 1934, 18, 1-4.
82. RUST, M. M. The effect of resistance on intelligence test scores of young children. *Child Developm. Monogr.*, 1931, No. 6.
83. SCHELL, M. Infection by intestinal protozoa in relation to the intelligence of siblings. *Child Developm.*, 1933, 4, 253-258.
84. SCHOTT, E. L. Variability of mental ratings in retests of neuropsychiatric cases. *Amer. J. Psychiat.*, 1930, 10, 213-227.
85. SCHOTT, E. L. IQ changes in foster home children. *J. appl. Psychol.*, 1937, 21, 107-112.
86. SCRUGGS, S. D. Effect of improvement in reading upon the intelligence of negro children. *Bull. Educ., Univ. Kans.*, May, 1935. (Special issue.)
87. SEAGOE, M. V. An evaluation of certain intelligence tests. *J. appl. Psychol.*, 1934, 18, 432-436.
88. SHERMAN, M., & KEY, C. B. The intelligence of isolated mountain children. *Child Developm.*, 1932, 3, 279-290.
89. SIMPSON, B. R. The wandering IQ: is it time for it to settle down? *J. Psychol.*, 1939, 7, 351-367.
90. SKEELS, H. M. The relation of the foster home environment to the mental development of children placed in infancy. *Child Developm.*, 1936, 7, 1-5.
91. SKEELS, H. M. Mental development of children in foster homes. *J. genet. Psychol.*, 1936, 49, 91-106.
92. SKEELS, H. M. Mental development of children in foster homes. *J. consult. Psychol.*, 1938, 2, 33-43.
93. SKEELS, H. M., & FILLMORE, E. A. The mental development of children from under-privileged homes. *J. genet. Psychol.*, 1937, 50, 427-439.
94. SKEELS, H. M., UPDEGRAFF, R., WELLMAN, B. L., & WILLIAMS, H. M. A study of environmental stimulation: an orphanage preschool project. *Univ. Ia Stud. Child Welf.*, 1938, 15, No. 4.
95. SKODAK, M. The mental development of adopted children whose true mothers are feeble-minded. *Child Developm.*, 1938, 9, 303-308.
96. SKODAK, M. Children in foster homes. *Univ. Ia Stud. Child Welf.*, 1939, 16, No. 1.
97. SNEDDEN, D. Practice effect. *J. educ. Res.*, 1931, 24, 376-380.
98. STUTSMAN, R. Factors to be considered in measuring the reliability of a mental test, with special reference to the Merrill-Palmer scale. *J. educ. Psychol.*, 1934, 25, 630-633.
99. SYMMES, E. F. An infant testing service as an integral part of a child guidance clinic. *Amer. J. Orthopsychiat.*, 1933, 3, 409-430.
100. Terman, L. M., & MERRILL, M. A. Measuring intelligence: a guide to the administration of the new revised Stanford-Binet tests of intelligence. New York: Houghton Mifflin, 1937.
101. THORNDIKE, R. L. The effect of interval between test and retest on the constancy of the IQ. *J. educ. Psychol.*, 1933, 24, 543-549.

102. TRAXLER, A. E. Reliability, constancy and validity of the Otis IQ. *J. appl. Psychol.*, 1934, **18**, 241-251.
103. UPDEGRAFF, R. The determination of a reliable intelligence quotient for the young child. *J. genet. Psychol.*, 1932, **41**, 152-166.
104. WELLMAN, B. L. Some new bases for interpretation of the I.Q. *J. genet. Psychol.*, 1932, **41**, 116-126.
105. WELLMAN, B. L. The effect of pre-school attendance upon the I.Q. *J. exp. Educ.*, 1933, **1**, 48-69.
106. WELLMAN, B. L. Growth in intelligence under differing school environments. *J. exp. Educ.*, 1934, **3**, 59-83.
107. WELLMAN, B. L. Mental growth from pre-school to college. *J. exp. Educ.*, 1937, **6**, 127-138.
108. WELLMAN, B. L. Intelligence of pre-school children as measured by the Merrill-Palmer scale of performance tests. *Univ. Ia Stud. Child Welf.*, 1938, **15**, No. 4.
109. WELLS, J., & ARTHUR, G. Effect of foster-home placement on the intelligence ratings of children of feeble-minded parents. *Ment. Hyg., N. Y.*, 1939, **23**, 277-285.
110. WHEELER, L. R. The intelligence of East Tennessee mountain children. *J. educ. Psychol.*, 1932, **23**, 351-370.
111. WOLCOTT, W. Changes in Thorndike intelligence test scores at the end of the college course. *Sch. & Soc.*, 1933, **37**, 627-632.
112. WOODALL, C. S. Analysis of I.Q. variability. *Proc. Amer. Ass. Stud. Feeblemind.*, 1931, **36**, 247-262.