DATA FROM THE HEALTH EXAMINATION SURVEY

Series 11
Number 107

Intellectual Development of Children as Measured by the Wechsler Intelligence Scale

United States

Information from the distribution of raw scores for the Vocabulary and Block Design subtests as well as scaled and standard scores or deviation IQ's derived from them by age and sex for noninstitutionalized children 6 through 11 years of age in the United States, obtained by administering this short form of the WISC to a representative sample of this population.

DHEW Publication No. (HSM) 72-1004

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Public Health Service
Health Services and Mental Health Administration
National Center for Health Statistics
Rockville, Md. August 1971

Series 11 reports present findings from the National Health Examination Survey, which obtains data through direct examination, tests, and measurements of samples of the U.S. population. Reports 1 through 37 relate to the adult program; additional reports concerning this program are forthcoming and will be numbered consecutively. The present report is one of a number of reports of findings from the children and youth programs, Cycles II and III of the Health Examination Survey. These reports, emanating from the same survey mechanism, are being published in Series 11 but are numbered consecutively beginning with 101. It is hoped this will guide users to the data in which they are interested.

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402 - Price 50 cents

NATIONAL CENTER FOR HEALTH STATISTICS

DIVISION OF HEALTH EXAMINATION STATISTICS

COOPERATION OF THE BUREAU OF THE CENSUS

In accordance with specifications established by the National Health Survey, the Bureau of the Census, under a contractual agreement, participated in the design and selection of the sample, and carried out the first stage of the field interviewing and certain parts of the statistical processing.

Vital and Health Statistics-Series 11-No. 107

DHEW Publication No. (HSM) 72-1004

Library of Congress Catalog Card Number 78-608552

CONTENTS

SYMBOLS	
Data not available----------	---
Category not applicable----------	...
Quantity zero----------	-
Quantity more than 0 but less than 0.05----	0.0
Figure does not meet standards of reliability or precision----------	*

INTELLECTUAL DEVELOPMENT OF CHILDREN
AS MEASURED BY THE WECHSLER INTELLIGENCE SCALE FOR CHILDREN

Jean Roberts, *Division of Health Examination Statistics*

INTRODUCTION

This report presents information on the distribution of the levels of intellectual development of children 6-11 years of age in the noninstitutional population of the United States as estimated from the Wechsler Intelligence Scale for Children test data obtained from the two subtests used in the Health Examination Survey of 1963-65—Vocabulary and Block Design. Consideration is limited to age and sex differentials only in the first of a series of reports on the test findings.

The Health Examination Survey is one of the major programs organized within the National Center for Health Statistics (NCHS) to carry out the National Health Survey, authorized in 1956 by the 84th Congress as a continuing Public Health Service activity.

Three different survey programs are employed to accomplish the objectives of the National Health Survey.[1] One of these is the Health Interview Survey, in which persons are asked to give information related to their health or that of other household members. A second program, surveys of health resources, obtains health data as well as health resources and utilization information through surveys of hospitals, nursing homes, and other resident institutions and the entire range of personnel in the health occupations. The third major program used for the National Health Survey is the Health Examination Survey.

In the Health Examination Survey, data are collected by direct physical examinations, tests, and measurements performed on the sample population studied. This is the most accurate way to obtain definite diagnostic data on the prevalence of certain medically defined illnesses. It is the only way to secure information on unrecognized and undiagnosed conditions as well as on a variety of physical, physiological, and psychological measurements within the population. In addition, it makes possible the study of relationships among the various examination findings and between these findings and certain demographic and socioeconomic factors.

The Health Examination Survey is carried out as a series of separate programs referred to as "cycles." Each cycle is concerned with some specific segment of the total United States population, usually a particular age group, and with certain specified aspects of the health of that subpopulation. The first cycle obtained data on the prevalence of certain chronic diseases and on the distribution of various measurements and other characteristics in a defined adult population, as described previously.[2,3]

For the second program or cycle of the Health Examination Survey (HES), on which this report is based, a probability sample of the Nation's noninstitutionalized children 6-11 years of age was selected and examined. The examination focused particularly on health factors related to growth and development but also screened for heart disease, congenital abnormalities, ear-nose-throat diseases, and neuromusculo-skeletal abnormalities. It included an examination by a pediatrician; examination by a dentist; tests administered by a psychologist;

and a variety of tests, procedures, and measurements given by technicians. A comprehensive description of the survey plan, sample design, content of the examination, and operation of the survey has been presented in another report.[4]

This second program of the Survey was started in July 1963 and field collection operations completed in December 1965. Out of the 7,417 children selected for the sample, 7,119 (96 percent) were examined. This national sample is representative of the roughly 24 million noninstitutionalized children in the United States 6-11 years of age.

A standardized single-visit examination was given to each child by the examining team in the specially designed mobile units used for the Survey. Prior to the examination, information was obtained from the parent of the child; this included demographic and socioeconomic data on the household members as well as a medical history and behavioral and related data on the child to be examined. Ancillary data for the child including grade placement, teacher's rating of his behavior and adjustment, and health problems known to the teacher were requested from the school. Birth certificates for verification of the child's age and information related to the child at birth were also obtained.

THE PSYCHOLOGICAL TEST BATTERY

After consultation with five child psychologists from some of the leading universities and the National Institute of Mental Health, a 60-minute test battery to assess the mental aspects of growth and development was assembled to be individually administered as part of the standard examination. The battery consisted of measures of or closely related to intelligence, as well as other tests designed to assess other developmental factors.

For direct estimates of intelligence, the battery included the Vocabulary and Block Design subtests of the Wechsler Intelligence Scale for Children (WISC) and a form of the Draw-a-Person Test. For the assessment of personality factors, five cards of the Thematic Apperception Test (TAT) were included. Estimates of school achievement in basic skills of arithmetic computation and reading were obtained using two subtests of the Wide Range Achievement Test (WRAT). These latter tests were also included to make possible the assessment of relationships among measures of school achievement, intellectual status, and social and emotional adjustment.

A comprehensive evaluation of the psychological procedures selected for the second cycle of the Health Examination Survey was included in the methodological study contract report by a recognized authority in the field of psychology—Dr. S. B. Sells of the Institute of Behavioral Research, Texas Christian University. This study included a literature review of previous research and evaluation known to be available on each of the battery components, recommendations concerning the types of inferences which could appropriately be made from the results to be obtained from the battery, and recommendations with respect to additional research which was deemed necessary to make logical use of the data collected. The results are published in the Center's methodological series.[5]

THE WECHSLER INTELLIGENCE SCALE

The Wechsler Intelligence Scale for Children, which was published in 1949, extended the well-known Wechsler intelligence scale for adolescents and adults into the childhood ranges of 5-15 years.[3] During the two decades since its publication the WISC has been the subject of extensive investigation and has achieved wide school and clinic use where individual measures of intelligence are desired.

The concept of intelligence represented in the WISC is that of an aggregate or global capacity of the individual to act purposefully, to think rationally, and to deal effectively with his environment. While intelligence is composed of elements or abilities—such as verbal, abstract, numerical, or general—which though not entirely independent, are quantitatively differentiable, it is also generally considered to be part of personality itself. The theory underlying the WISC is that intelligence cannot be separated from the rest of the personality, and a deliberate attempt has been made to take into account other

factors such as persistence, drive, or energy level, which contribute to the total effective intelligence of the individual. This effort is reflected both in the composition of the WISC and in the impartial weights assigned to the separate subtests. No attempt has been made in the WISC to bring together a series of tests that measure "primary abilities" or to order them into a hierachy of relative importance.

The WISC consists of 12 tests, two of which are considered supplementary or alternative tests in the adult scale; the tests are divided into two major subgroups identified as Verbal and Performance. Most of the verbal tests correlate better with each other than with tests of the performance group; the reverse is also true.

The Wechsler intelligence tests use the *deviation intelligence quotient* (IQ). This measure supplants the *mental age* concept and evaluates the performance of each individual on the basis of the distribution of scores of a representative sample of his own chronological age. In the standardization of the WISC, Wechsler kept the standard deviation of intelligence quotients constant from year to year, with the result that a child's obtained deviation IQ does not vary unless his actual test performance as compared with his peers varies.

Raw scores for each subtest are converted to *scaled scores* which have a mean of 10 and standard deviation of 3 for each age level. The sum of five scaled scores for the Verbal Series constitutes the Verbal Scale score (VS), and similarly the Performance Scale score (PS) is the sum of the scaled scores for five of the tests in the Performance Series. The Full Scale score (FS) is the sum of the Verbal Scale and the Performance Scale. Deviation intelligence quotients have been derived by a similar conversion process for VS, PS, and FS. The (deviation) IQ scales at each age have a mean of 100 and a standard deviation of 15.

Short Forms of the WISC

Time limitations in the survey examination did not permit administering the full WISC in addition to the other tests and measurements which needed to be obtained in an adequate appraisal of growth and development of children. Indeed only two parts of the test could be adequately administered in the time available. Consideration, hence, needs to be given to the effect that this will have on the national estimates of the intelligence of children presented here.

Several investigators have assessed the efficiency and accuracy of various combinations of two or more subtests of the WISC in estimating intelligence as measured by the Full Scale.[7-12] Of these, only one study, that by Silverstein,[12] dealt with a fairly large sample of predominantly normal children. He determined the correlations with the Full Scale of all possible short forms of 2, 3, 4, and 5 subtests using the WISC standardization data for 200 children at each of three age levels—7½, 10½, and 13½ years The correlations increase, as expected, as the length of the short form increases. Among the 10 most valid short forms of each length at each age level, the correlations range from 0.81 for the poorest predictor from the dyad sets at age 7½ years to 0.97 for the best pentad combination at age 10½. There is a tendency for the correlations at age 10½ to be higher than those at 13½, which in turn tend to be higher than those at 7½ years. Using the Wherry-Doolittle method which entails differential weighting of subtest scores rather than the simple summation of scaled scores did not result in appreciably higher validities. The best dyad predictor at age 10½ years (but not at the other two age levels) was found to be the Vocabulary-Block Design combination ($r_{FS \cdot V+BD} = 0.91$) used in this survey.

The other studies found are based on smaller atypical groups of children who were either mentally retarded, physically disabled, or referred to child guidance clinics or social agencies. One of these, the study by Schwartz and Levitt[11] among 179 retarded children also determined the efficiency of all possible 2 through 6 subtest combinations in predicting Full Scale intelligence test results. Here the correlations range from 0.79 for the best dyad to 0.95 for the best hexad. The Vocabulary-Block Design combination yielded a correlation of 0.74.

In the other studies the correlation of the two-set combination used in this survey with the Full Scale WISC was found by Finley[8] to be

0.68 among some 309 mentally retarded children; by Greenmun[9] to be 0.92 among 632 referrals to Child Guidance Clinic at Texas Christian University; by Simpson and Bridges[13] to be 0.87 among 120 children referred to social agencies and child guidance clinics; and by Wight and Sandry[14] to be 0.91 among 83 children hospitalized for a physical disability.

Hite[15] has confirmed Wechsler's data[6] indicating that Vocabulary and Block Design are the most reliable subtests of the WISC battery, with the highest split-half reliability coefficients and the lowest standard errors of measurement. His study included a stratified sample of 200 children aged 5½, 6½, and 7½ years in addition to the WISC standardization groups. Hagen[16] and Cohen[17] in the United States and Gault[18] in Australia have reported that both of these subtests are highly loaded on the so-called "general" factor of intelligence obtained in factor analysis of the WISC over the entire age range of 5 to 15 years.

Hence the two subtests selected for use in this survey appear in general to be the best available or at least as good as any two subtests for this age range although they are somewhat better for the older than the younger children. The validity of this short form of the WISC as a basis for the estimation of Full Scale intelligence scores is clearly of a fairly high order.

FIELD ADMINISTRATION

The Vocabulary and Block Design subtests of the WISC were the second and third procedures, following the human figure drawing test, in the 60-minute individual testing session allotted for each child examined. The testing was done in a small, adequately lighted climate-controlled, and sound-treated room in the mobile examination center by psychologists who had obtained at least a master's degree and who had had previous experience in administering tests to children. There were two psychologists (usually a man and woman to whom the examinees were assigned approximately at random) on the examining team at all times. The examiners were selected, trained in field testing procedures, and supervised by the Psychological Advisor to the Health Examination Survey. In the initial training and ensuing supervision of the examiners, strong emphasis was placed on uniform methods of test administration, scoring, and recording of data. During the course of the children's survey, a total of 25 examiners participated in administering the test.

The standard WISC Record Form (copyright 1949 by the Psychological Corporation) was used by the psychologist for recording the child's responses and scores in the Vocabulary and Block Design subtests. Standard procedures of administration were followed as specified in the WISC Manual,[6] except that every child was started with the first word in the Vocabulary test.

Vocabulary Subtest

The Examiner started by saying: "I want to see how many words you know. Listen carefully and tell me what these words mean. 'Scooter' . . . .what is a 'scooter'?" He recorded the Subject's responses and then proceeded with the words in the order listed, repeating at each presentation, "What is a . . .?" or "What does . . . mean?"

With more intelligent and older subjects the formal question was usually omitted after the third word and just the word pronounced. The Examiner used for the words the preferred pronunciation as given in *Webster's New Collegiate Dictionary,* copyright 1961. All subjects were started with the first word listed, even children over 8 years of age.

When it was difficult for the Examiner to decide whether the subject did or did not know the meaning of a word, he would say: "Please explain a little more," or make some other neutral inquiry. This type of neutral inquiry was used when the response was so vague it could not readily be scored or when it indicated enough understanding of the word so that a better response might be evoked by such neutral inquiry.

For words which sometimes evoke a response based on the homonym—such as "cede" ("seed"), "ware" ("wear"), "pear" ("pair"), "fairy" ("ferry")—the Examiner would pronounce the word again and ask: "What else does

...mean?" He would not spell the word for the Subject. The scoring here was based on the response for the word intended regardless of the quality of the response to the homonym.

The Vocabulary subtest was discontinued after five consecutive failures. Each word was scored "2," "1," or "0" except words 1 through 5, which were scored only "2" or "0." A maximum of 80 points could be obtained in this subtest.

Block Design Subtest

Children under 8 years of age were started with Design A; those 8 years or older began with Design C and were given automatic full credit for Design A and B if they passed Design C.

Design A.—The Examiner took four blocks in hand and said to the Subject: "You see these blocks have different colors on their different sides. I am going to put them together to make something with them. Watch me."

The Examiner arranged the four blocks slowly as shown on Design A of the reference card for this subtest. He then gave four other blocks to the Subject and said: "Now make one just like mine. Tell me when you are finished." If the Subject failed, he was told: "Watch me again." The Examiner would demonstrate a second time using the blocks just employed by the Subject. After the demonstration with the Subject's blocks, these were mixed up but the Examiner's blocks were still together in the model and the Subject was asked "Now try it again and be sure to make it just like mine."

A score of 2 points was given for passing on the first trial and 1 point for passing on the second trial for each of Designs A, B, and C. Failure was scored on an item either for faulty design or for failure to complete the design in the allotted time as specified on the references card. Each trial was timed separately.

Design B.—The Examiner brushed up all the blocks before continuing, then assembled Design B behind a screen and presented the model to the subject in a completed form saying: "Now make one like this. Do it yourself."

If the Subject failed on the first trial of Design B, his blocks were brushed up by the Examiner and he was told: "Watch me do it." The examiner left the model intact and duplicated it with the Subject's blocks. After the second demonstration, the Subject's blocks were brushed up and he was asked: "Now you try it."

The subtest was discontinued if the Subject failed on both trials of Design B.

Design C.—If the Subject succeeded on either the first or second trial of Design B, the Examiner removed the blocks that served as a model and put the reference card C in their place. He then said: "This time we are going to put the blocks together by making them look like this picture. Watch me first." The Examiner constructed the design slowly and when finished said: "You see, the top of these blocks looks the same as this picture." After brushing up the demonstration, he asked the Subject: "Now you look at the picture and make one just like it with these blocks. Go ahead."

If the Subject failed to complete the design accurately or in the allotted time, the blocks were brushed up and a second demonstration was given. Then the blocks were brushed up again and the Subject was asked: "Go ahead. See if you can get it this time."

For Subjects under 8 years of age, if both trials for Design C were failed, this subtest was discontinued. For those 8 years or older if both trials of Design C were failed, Designs A and B were administered as for the younger subjects. The subtest was then discontinued and they were given as a score whatever points they earned on Designs A and B.

Designs 1-7.—For those passing Design C, the Examiner placed the reference card Design 1 before the Subject and said: "Now you make one like this." The Subject was provided with four blocks. The test was continued in a similar manner with the succeeding designs. When Design 5 was reached, the Examiner took the other five blocks out and said: "Now make one like this, using nine blocks."

In this part of the subtest—Designs 1-7—no second trials were given and the test was discontinued after two consecutive failures. Credit of 4 points was given for each Design in this section completed correctly. Bonus points were given, as specified in the WISC Manual,[6] for completing the design in less than the allot-

ted time. No credit was given for partially correct or incomplete performance. A maximum of 55 points could be obtained on this subtest.

Quality Control

The maintenance of standard administration procedures and uniform methods of recording are essential in large data collecting operations such as the Health Examination Survey. In addition to the initial training of examiners in the survey procedures which for the psychological portions included the memorization of all test instructions, several ongoing procedures were devised to assure the continuing quality of the data. The field psychologists exchanged all test forms daily and checked for any apparent errors in administration or mistakes in recording.

Each field psychologist tape-recorded one entire testing session each week. The tapes were sent to the supervisor who reviewed them and made notes of errors and suggestions regarding testing procedures. These notes were sent to the examiners. In addition, the Psychological Advisor or Supervising Field Psychologist made periodic visits to the field for direct observation and supervision of the work. The test forms were also spot-checked when they arrived at headquarters.

The WISC Vocabulary subtests administered on one day of each week were rescored by the field psychologist who had not given the particular tests. Points of disagreement were discussed by the two field psychologists and if agreement could not be reached, were referred to the Psychological Advisor for decision.

FINDINGS

Verbal—Vocabulary Raw Scores

On the measure of the "verbal" aspects of intellectual development used from the WISC—the Vocabulary subtest—the mean raw scores attained by the 24 million noninstitutionalized children 6-11 years of age in the United States, as estimated from the Health Examination Survey findings of 1963-65, increased steadily with age from 16.4 points at 6 years to 34.9 points at 11 years out of a total possible score of 80 points (table 1). If the scores are considered at 4-month age intervals, there is also a fairly smooth progressive increase in means from 15.0 points in the first 4 months of age 6 years to 36.2 in the last 4 months of age 11 years (table 2 and figure 1).

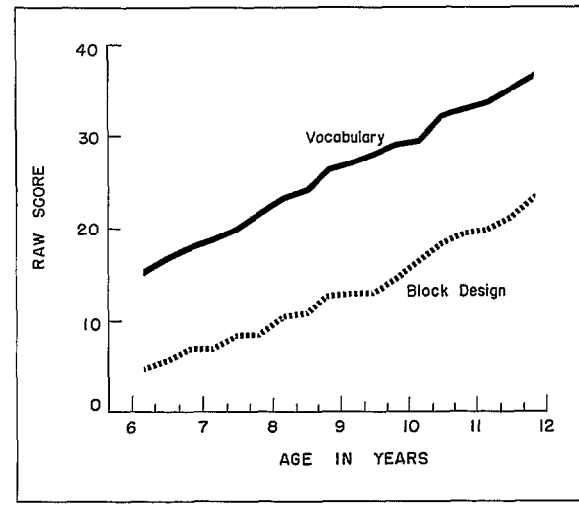

Figure 1. Average raw scores on the Vocabulary and Block Design subtests of the WISC at 4-month age intervals: United States, 1963-65.

The variation in raw scores obtained among these children, as indicated by the standard deviations of the raw score distributions, also showed a consistent increase with age but at a slightly decelerating rate (tables 1 and 2 and figure 2). Since the highest score attained was among the 11-year-olds and that fell 16 points short of the maximum possible, the Vocabulary subtest apparently tested the full range of ability that could be measured by it throughout the age range 6-11 years. Because of the slowing rate of increase with age in variability but not mean score, the relative variability, as roughly measured by the coefficient of variation (ratio of the standard deviation to the mean), shows a gradual downward trend with increasing age (table 2 and figure 3).

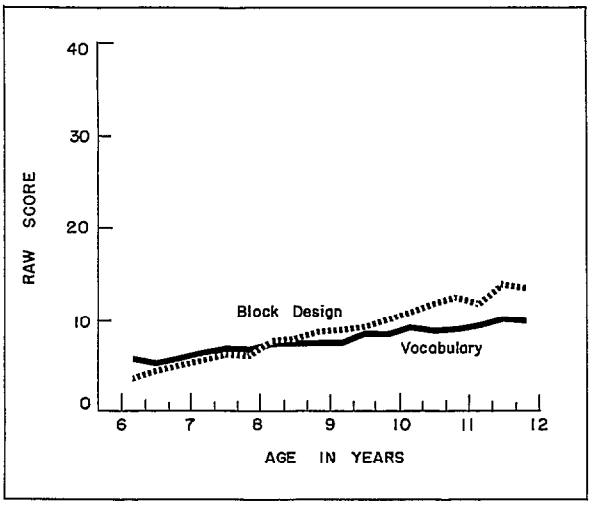

Figure 2. Standard deviations of raw scores on the Vocabulary and Block Design subtests of the WISC at 4-month age intervals: United States, 1963-65.

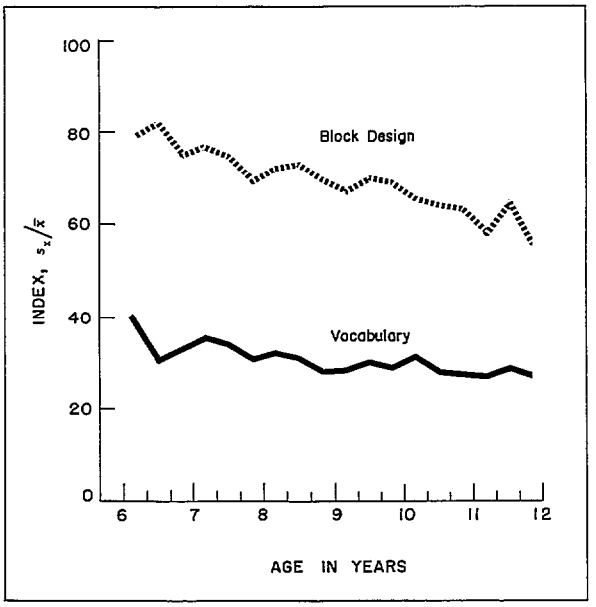

Figure 3. Coefficient of variation, $s_x/\bar{x}$, on the Vocabulary and Block Design subtests of the WISC at 4-month age intervals: United States, 1963-65.

Boys on the average consistently scored higher than girls throughout the age range tested (table 1 and figure 4). The mean differences in scores were statistically significant at the 5-percent level for children 6-10 years of age on this subtest.

Information by grade in school at the time of testing shows a consistent pattern of increase in mean score with grade level from those in kindergarten to those in seventh grade. Younger children in the respective grades tended to obtain higher scaled scores and older children lower scaled scores on the average than those

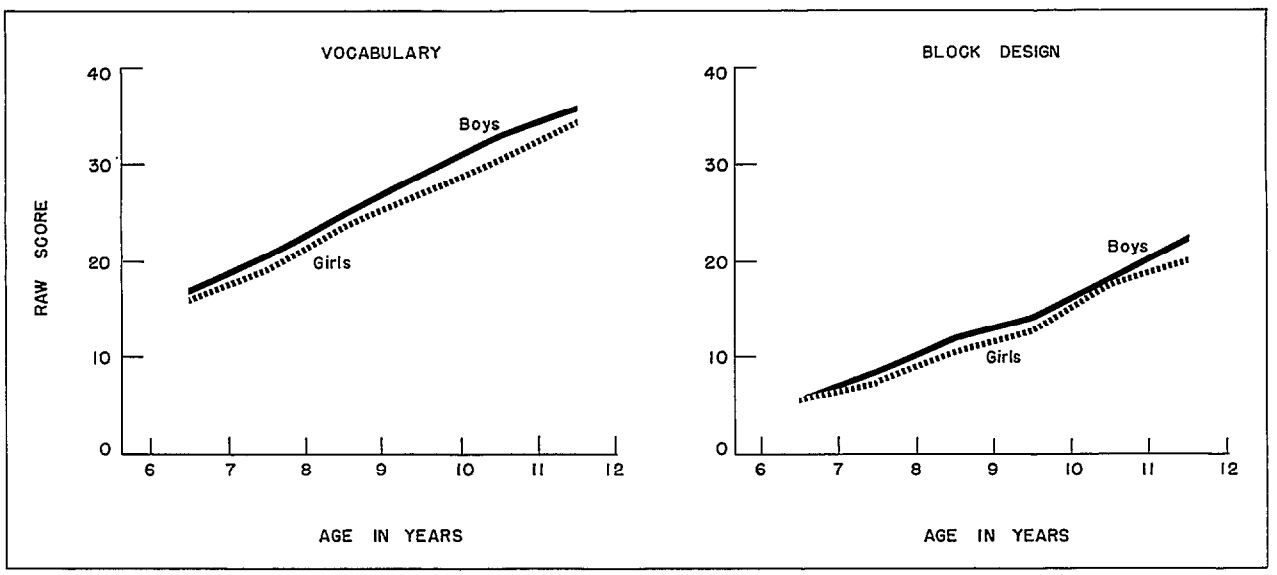

Figure 4. Average raw scores on the Vocabulary and Block Design subtests of the WISC, by sex: United States, 1963-65.

in the normal grade placement for their age (table 4).

As evident in the age trends, boys showed higher mean scores than girls at each grade level as well as in the special ungraded classes for mentally retarded and physically handicapped children. The mean differences were statistically significant at the 5-percent level from the first through the seventh grade. It should be kept in mind that the kindergarten group and the seventh graders included here will be primarily the youngest of the 6-year-olds and the oldest of the 11-year-olds, respectively. Deviations from the general pattern, indicated above, at the extremes of the grade levels and at the extremes of the age range within grade probably reflect only sampling errors with the small number of children involved at these extremes.

Performance—Block Design Raw Scores

Results with the measure of the "performance" aspects of intellectual development used from the WISC—the Block Design subtest—show a similar pattern to from with the Vocabulary subtest with few exceptions. Mean raw scores on the Block Design increased steadily with age, a trend paralleling that for Vocabulary, but remained 10 to 15 points below since the upper limit possible on this performance subtest is only 55 points. However, the raw scores themselves on the two subtests cannot be considered comparable. Here, the mean raw scores range from 5.7 points at 6 years to 21.2 points at 11 years (table 3). At 4-month age intervals the means start at 4.8 points for those in the first 4 months of age 6 years and increase steadily to 22.9 for children in the last 4 months of age 11 years (table 2 and figure 1). The upward trend while consistent is somewhat less smooth than that for the Vocabulary subtest and even shows a slight (but not statistically significant) slow down between the fourth and the eighth month of age 9 years.

The variation in these scores also increased with age but at a slightly faster rate than for the Vocabulary subtest (tables 2 and 3 and figure 2). Since the absolute amount of variability was of about the same order of magnitude for both subtests, while the mean scores differed substantially, the relative variability as measured roughly by the coefficient of variation was about twice as large on the Block Design

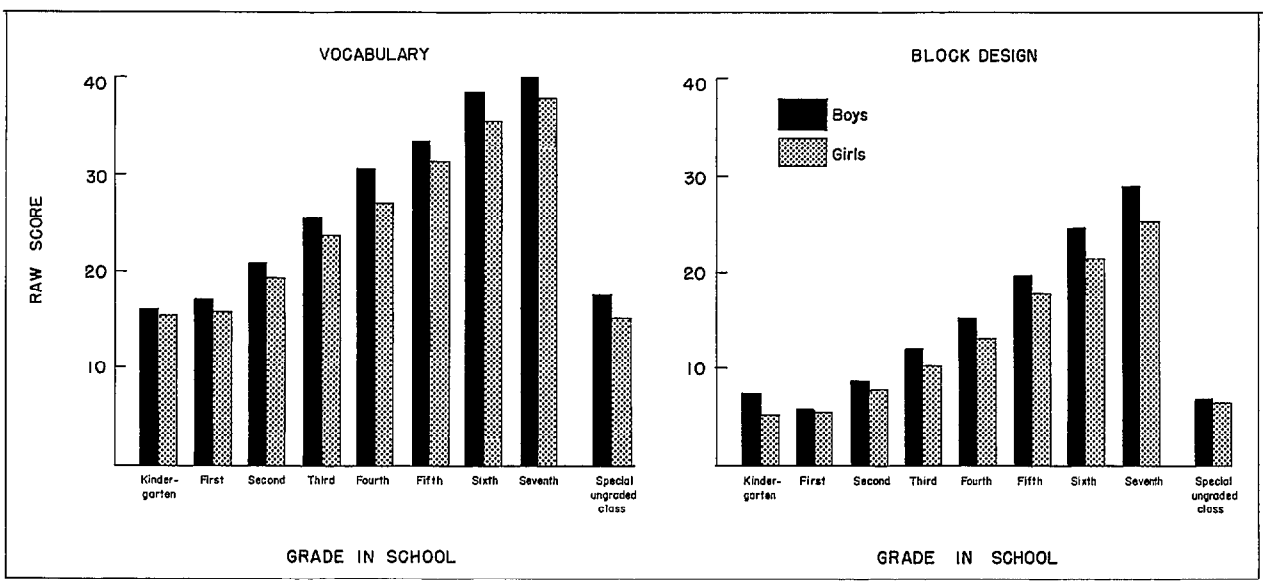

Figure 5. Average raw scores on the Vocabulary and Block Design subtests of the WISC, by grade in school and sex: United States, 1963-65.

as on the Vocabulary subtest (table 2 and figure 3). This coefficient for the Block Design subtest showed more irregular fluctuation but a sharper downward trend with age than the "verbal" test part used. The latter pattern appears to suggest that the full range of this aspect of intelligence (performance) was apparently not as adequately sampled as on the Vocabulary subtest, particularly among older children since at both 10 and 11 years a few came within one point of the maximum score attainable on this subtest.

Throughout the age range tested, boys on the average scored higher than girls on the Block Design subtest also (tables 1 and 3 and figure 4). Here the mean differences were statistically significant at the 5-percent level for children 7-9 and 11 years of age.

On the Block Design as on the Vocabulary subtest, there was an increase in mean score attained with successive grade levels but only from first grade on for both boys and girls (figure 5). The mean score for boys, but not for girls, on this "performance" subtest was slightly lower in the first grade than in the kindergarten group; however, this difference was small enough to be due to chance alone in a sample of the size and design used in the survey. The mean differences in scores between boys and girls were large enough to be statistically significant at the 5-percent level from the second through the sixth grades. As for the Vocabulary subtest, average Block Design scaled scores were higher for the younger children in each grade and lower for the older ones than they were for the children in the normal grade placement for their age (table 5).

Comparison With the WISC Standardization Group

The WISC,[6] published in 1949, was standardized on a sample of 100 boys and 100 girls at each year of age from 5 through 15. Included within this group were 55 institutionalized, feebleminded children within that chronological age range who had been rated by staff psychologists as having IQ's between 50 and 70. With the exception of the known feebleminded group, most of whom were tested within 2 months of their midyear, each child was tested within 1½ months of his midyear. Only white children were included in this standardization group.[6]

Within each year of age and for the total group, the sample was selected so that it would be representative of the proportions in the total United States in 1940 (1) in the four regions of the country—New England and Middle Atlantic, North Central, South Atlantic and South Central, and Mountain and Pacific States, (2) in the total urban and rural population, and (3) with father's occupation distributed similarly to all employed white males. The Midwest sample was made short of cases and the Western group was made slightly larger than the U.S. proportions in 1940 to allow for the wartime and postwar population shifts during the 1940's.

Standardization testing took place in 85 communities. The samples chosen were apparently all from school populations with the exception of the 55 children in institutions for the feebleminded.

Thus for the age group of concern in the present study the standardization group included 1,200 children—600 boys and 600 girls—at the ages of 6-11 years.

The present study was based on findings from 7,119 examined children out of a total sample of 7,417 drawn to represent the 24 million noninstitutionalized children 6-11 years of age in the United States. The sample design used for the Survey was a multistage, stratified probability sample of loose clusters of persons in land-based segments, as described in appendix I and in a previous publication.[19] The sampling frame used here insured the representativeness of the sample with respect to degree of urbanization, region, and the rate of population change from 1950 to 1960, the latter being used as an indirect indicator of the economic condition of the area. Further analysis of the demographic and socioeconomic characteristics of the sample and the examined group indicated that both groups are also closely representative of the population from which they were drawn with respect to age, sex, race, family income, and education of father. Data used in this report for each sample child are inflated in the estimation process to characterize the larger universe of which the sample child is repre-

sentative and to include an adjustment for the small nonresponse group. This made the final sample estimates of the population agree exactly with independent controls from the Bureau of the Census for the U.S. noninstitutionalized population of August 1, 1964 (approximate midsurvey point) by color and sex for each single year in the age range 6 through 11. Thus findings from the present study are based on approximately 1,100 children at each single year of age 6 through 11, or 5½ times as many as in the original standardization group.

Comparisons of the means and standard deviations of raw scores obtained on the Vocabulary and Block Design subtests from the

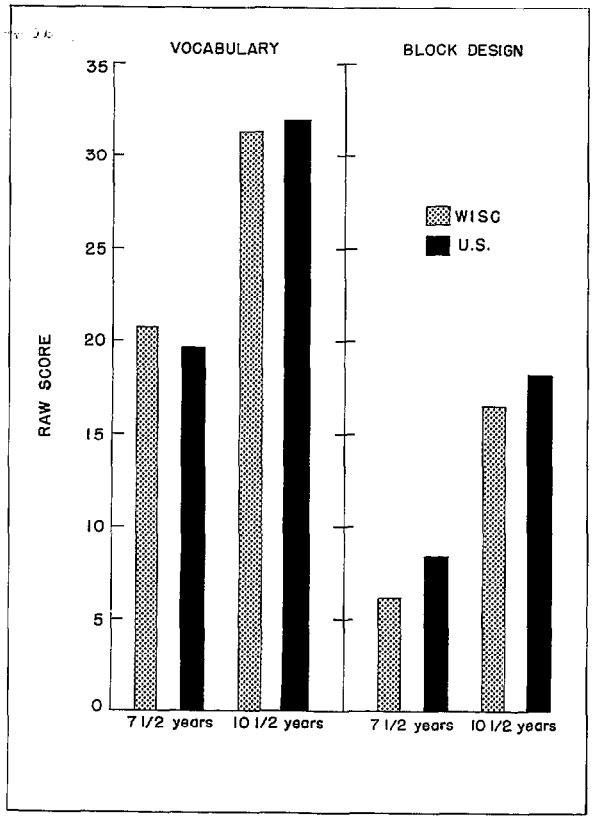

Figure 6. Average raw scores on the Vocabulary and Block Design subtests of the WISC for the WISC standardization group and the U.S. estimates from the Health Examination Survey at $7\frac{1}{2}$ and $10\frac{1}{2}$ years of age.

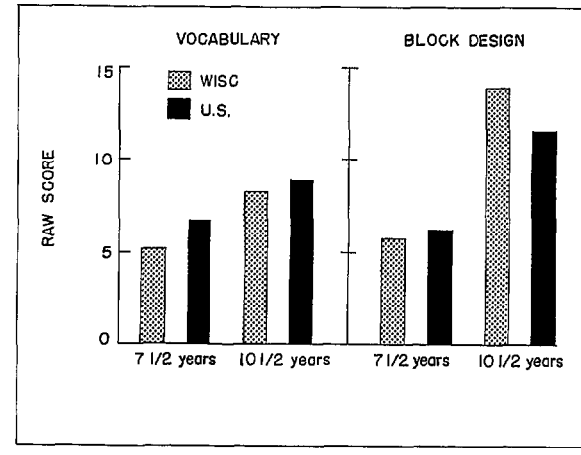

Figure 7. Standard deviation of raw scores on the Vocabulary and Block Design subtests of the WISC for the WISC standardization group and the U.S. estimates from the Health Examination Survey at $7\frac{1}{2}$ and $10\frac{1}{2}$ years of age.

WISC standardization sample and the U.S. estimates from the present study at ages 7½ and 10½ years are shown in figures 6 and 7. It will be noted that in comparison with the WISC standardization group the mean score on the Vocabulary subtest in the present study was lower and there was greater variability in scores (both significant at the 5-percent level) for the 7½-year-olds. At age 10½ years the mean and variability were both slightly, but not significantly, greater in the present study than in the WISC sample. For the Block Design subtest the means from the present study at these two ages exceeded those in the original standardization group; however, only at 7½ years was the difference large enough to have been statistically significant at the 5-percent level. The scores for the standardization group at 7½ years were slightly less variable than those of comparable age in the present study but at age 10½ years were substantially more variable.

As shown in figures 8 and 9, when comparison is made with the original standardization group throughout the age range included in the present study, the mean raw scores on the Vocabulary subtest in the present study were lower for the younger and older children up to the last quarter of age 8 and from the last

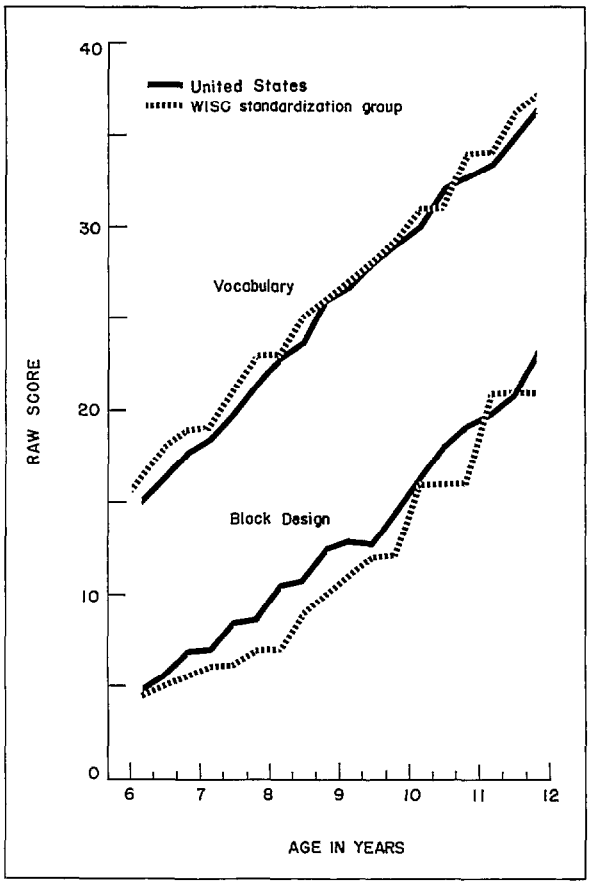

Figure 8. Average raw scores on the Vocabulary and Block Design subtests of the WISC for the WISC standardization group and the U.S. estimates from the Health Examination Survey at 4-month intervals.

quarter of age 10 on. Block Design means in the present study remained higher than those in the original standardization group up to age 11. From about 8 years on, greater variability was found on the Vocabulary subtest in the present study and less on the Block Design subtest than for the standardization group.

Boys tended to score higher than girls, on the average, on this test in both the original standardization study, where the entire battery of 12 subtests was used, and in the present study, which was limited to the two subtests. The differential in the present study is even more consistent and pronounced than in the original study. Seashore[31] found that, on the average, boys in the original standardization group did

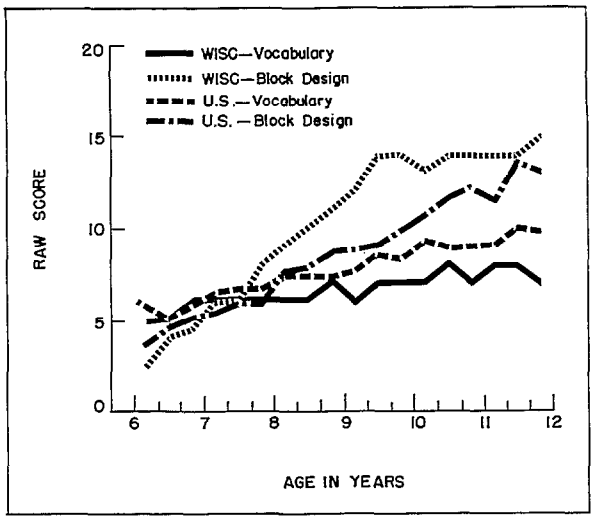

Figure 9. Standard deviations of raw scores on the Vocabulary and Block Design subtests of the WISC for the WISC standardization group and the U.S. estimates from the Health Examination Survey at 4-month age intervals.

slightly better than girls, the superiority being primarily in the older ages and the differences being small (3-4 standard score points at ages 8, 10, and 11 on the total verbal battery and at 8 and 10 years on the total performance battery, differences which would be statistically significant with the size of the sample used). In the present study, as previously noted, mean scores for boys on both subtests were higher than for girls throughout the age range tested, the differences being statistically significant at ages 6-10 years on Vocabulary and 7-9 and 11 years on the Block Design.

Scaled Scores

The WISC measures for the two aspects of intellectual development used here—the Vocabulary from the Verbal group and the Block Design from the Performance—are of different length,and are not directly comparable because of content, as has been indicated. The maximum score possible on the first of these is 80 points; on the second, 55. To compensate for the difference in test length and to make it possible to compare relative standing

within each subtest, raw scores for children within 4-month age intervals have been converted to scaled scores by setting the mean of each at 10 and the standard deviation at 3. The method of scaling is similar to that used for the 1949 WISC norms, but the means and standard deviations used were derived from the present study.

The correlations between scaled scores (as derived from the distributions in this national study) on the Vocabulary and Block Design subtests are in fairly close agreement with those from the original standardization group. At age 7 the correlation from the present study (+0.38 ± 0.027) was slightly higher than from Wechsler's group (+0.33), while at age 10 it was slightly lower (+0.50 ± 0.022 compared with +0.54 from the original standardization group).

The mean in terms of these scaled scores for single years of age by sex and grade on each subtest as determined from the present study are shown in tables 4 and 5. As on the raw score data, the mean scaled scores for boys slightly exceed those for girls on both subtests. In general the pattern by age within grade is similar for boys and girls. Children in the normal grade placement for their age tend to have about average scores, approximately 10 scaled score points, for those in their age group. The younger children in each grade have somewhat higher and the older children somewhat lower scaled scores than the average for their age group, the pattern and extent of the differences in scaled score units being similar for both subtests. The mean scaled scores for those in the special ungraded classes is, as expected, lower.

The basis for converting raw scores to scaled scores on the two test parts is shown in tables 6 and 7 for each of the 4-month age intervals for ages 6-11 years.

Standard Scores

The short form of the WISC consisting of the Vocabulary and Block Design subtests has been used here to obtain an estimate of the distribution of intelligence among noninstitutionalized children 6-11 years of age in the United States.

In making this estimate the scaled scores for the two subtests for each child have been added and the resultant combined distribution converted to standard scores or deviation intelligence quotients by a scale transformation setting the mean at 100 and the standard deviation at 15. These standard scores are of the order of magnitude of the older form of intelligence quotient which was the ratio of the mental age to the chronological age. The deviation intelligence quotient, however, makes possible the comparison of the performance of an individual with others of his own age, a measure which would not change on retest for a child unless his actual test performance as compared with his peers changes. The deviation IQ thus had the same relative meaning in each age group, unlike the older form of IQ.

The means of these standard scores by age, sex, and grade in school are shown in table 8. Here also, with the two subtest results combined, boys scored consistently higher than girls, the mean differences being statistically significant at the 5-percent level at each year of age with the exception of 6 and 10 years.

The pattern by grade is similar to that shown by the two sets of scaled scores. Those at the normal grade placement for their age test just slightly above normal from first grade on, while the younger children in each class have higher and the older children lower deviation IQ's on standard scores.

The basis for converting the combined scaled scores into standard scores with each 4-month age interval is contained in table 9.

Percentile Distributions

Another frequently used index of a child's relative standing in his age group, here with respect to intellectual development, is the percentile ranking of scores obtained for children of that age. The percentile rank equivalents of the raw scores for each of the subtests for boys and girls in the United States as determined in this survey for each year of age are shown in table 10. These percentile ranks give the relative standing of the score for a child in a theoretical group of 100, or the score below which the indicated percentage of children were found to fall.

It is generally assumed that intellectual development is normally distributed in the population. While the target in this study was limited to noninstitutionalized children and hence would not be expected to include as large a proportion of the mentally retarded as exist in the total child population, it is of interest to test the hypothesis of normality on the distribution of scores obtained in this study since this is the first time that data on these two subtests have been available on such a large and highly representative sample of the child population in the United States. An approximate chi-square test of the goodness of fit of the obtained distributions to the theoretical normal distribution shows the former to be essentially normal at each year of age on both subtests for both boys and girls as well as for the combined group. The goodness of fit for the Vocabulary subtest scores was much closer than that for the Block Design subtest, but even the latter did not differ more than would be expected by chance alone.

The smoothness and uniformity of the percentile distributions at each year of age for the total group from the Vocabulary subtest in contrast to the irregularity for the Block Design may be seen in figure 10. Figure 11 shows the

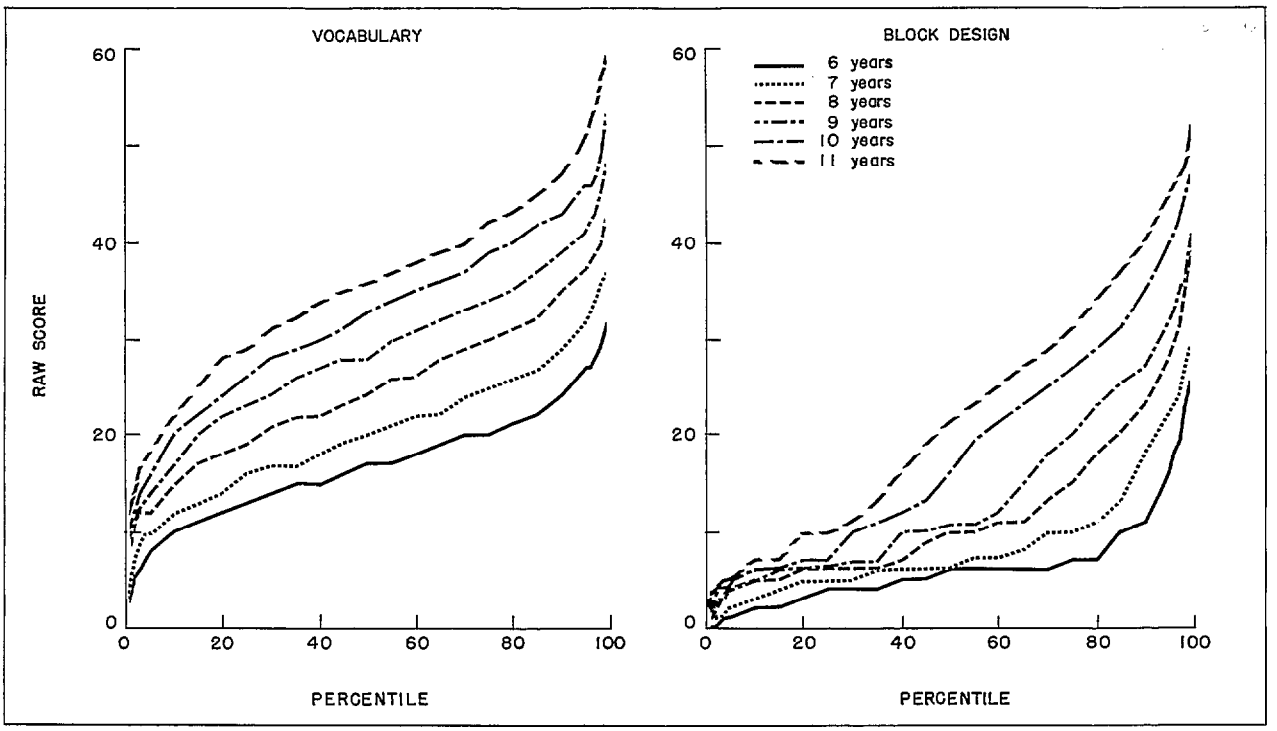

Figure 10. Percentile equivalents of raw scores on the Vocabulary and Block Design subtests of the WISC, by age: United States, 1963-65.

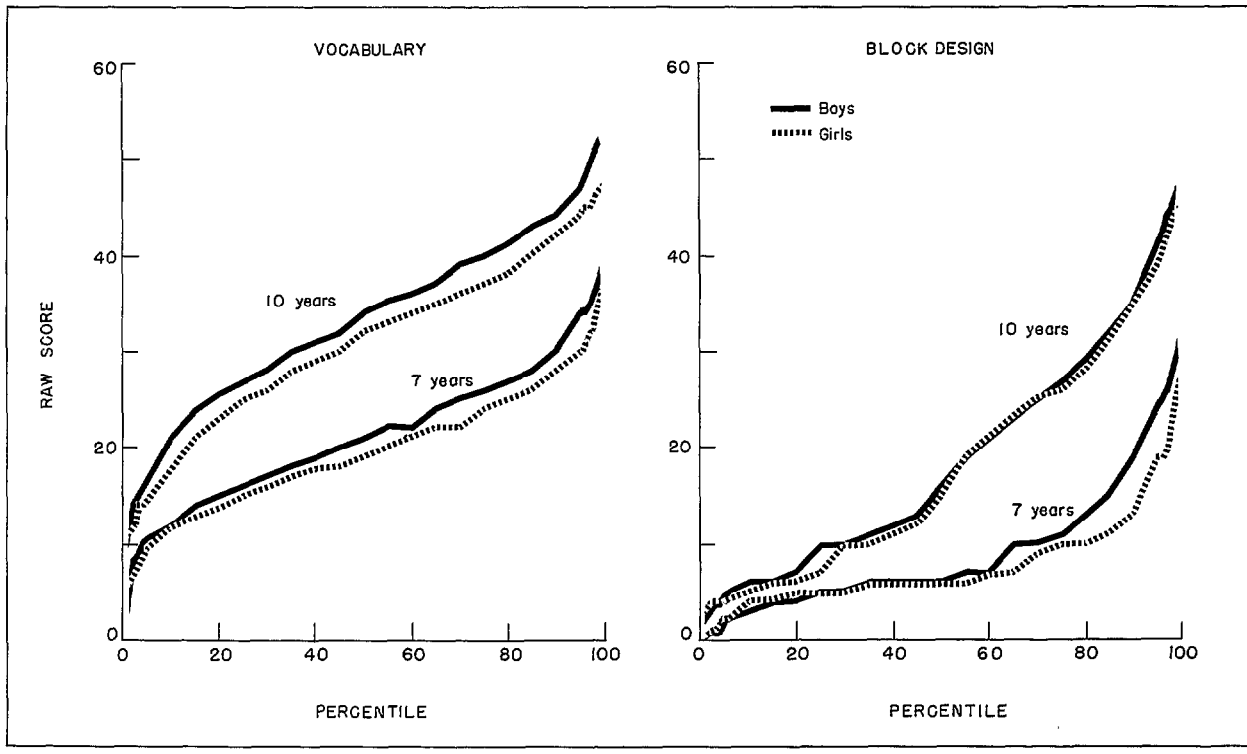

Figure 11. Percentile equivalents of raw scores on the Vocabulary and Block Design subtests of the WISC at 7 and 10 years of age, by sex: United States, 1963-65.

percentile distributions on each subtest for boys and girls separately at ages 7 and 10 years. Scores obtained by boys on the Vocabulary subtest consistently exceed those for girls throughout the distribution, while the pattern is less distinct on the Block Design.

DISCUSSION

As early as the start of this century, Alfred Binet and others showed that performance on the tests of judgment and reasoning varied systematically with chronological age in children. Binet was seeking to measure general intelligence but in an educational context so as to identify children who would require special educational facilities. The first such tests were considered successful because they differentiated between children known to do well and those known to be slow in school. Subsequent tests including the Wechsler Intelligence Scale for Children were validated against a Binet-type test and thus continued the influence of the school-learning context upon the measurement of intelligence.

The WISC is one of the more widely used and generally accepted individual measures of intelligence. The reliabilities (split-half) for the standardization groups on the Full Scale test were found to be 0.92, 0.95, and 0.94 at 7½, 10½, and 13½ years.[6] Validation studies using the Stanford-Binet as the external criteria show correlations between WISC Full Scale IQ's and Stanford-Binet IQ's to be generally high (in the mid 80's), with some even as high as the respective reliabilities of these tests. However, previous research has shown that WISC IQ's (deviation IQ's) tend to be substantially lower for the very young and for the gifted as well as for other samples across the normal range.[20-30]

Available research, cited previously, on the various possible short forms of the WISC indicate that while a longer test would give a better estimate of intelligence as measured by

the Full Scale WISC, the best or one of the better dyad predictors from it is the combination of the Vocabulary and the Block Design subtest.[6,8,9,11-13,15] The most extensive study based on the WISC standardization data indicated that these two subtests are the most closely correlated with the Full Scale at age 10½ years but only one of the better combinations at 7½ and 13½ years.[12] Wechsler[6] found these two subtests to be the most reliable of the WISC subtests and his results were later confirmed by Hite[15] for children aged 5-7 years.

The present study makes available for the first time Vocabulary and Block Design subtest findings for a highly representative national sample of the entire noninstitutional population of children 6-11 years of age with proportionate representation from all races. The sample is much larger—5½ times as great—as that used for the original test standardization and hence should provide an even more stable base for standardization of the two subtests than Wechsler's original data.

The national estimates of the Full Scale Standard Scores or Deviation Intelligence Quotients based on the findings from the two subtests in this study are, of course, less reliable than had the full 10 subtests been used. However, the combined scores do provide a rough measure of the distribution of intellectual development levels of noninstitutionalized children in this country, within the limitations described here.

It is of interest, though this study can provide no definitive explanation, that boys tend to outscore girls, even more consistently on these two subtests in the present study than they did on the entire set of subtests in the original standardization study. Seashore et al[31] in their original paper on the standardization of the Wechsler Intelligence Scale for Children had concluded, regarding these sex differentials, that the safest assumption was that boys and girls are the same in mental ability but that either the test items chosen turned out to be slightly biased in favor of the boys or that the sampling of boys was somehow chosen with a slight bias or both. They concluded this since Terman and Merrill[32] in their 1937 revision of the Stanford-Binet examination found the same situation and also could find no definitive answer from their data.

However, findings from the present study are based both on a far larger sample—more than 5 times as great at each year of age—and one that is more closely representative of the total child population of the United States from which it was drawn than the original group. Hence the factor of bias in selection of children can be eliminated as a consideration in the present findings. Although Seashore's data did not show test results separately for the two subtests used here, it is quite likely that this general differential between boys' and girls' performances would have been found for them in the original study because of the high correlation between these two subtests and the total Verbal and Performance sets in this test.

SUMMARY

This report presents national estimates of the intellectual development levels of noninstitutionalized children 6-11 years of age in the United States as determined by scores on the Vocabulary and Block Design subtest of the WISC obtained in the Health Examination Survey of 1963-65. In the survey a probability sample of 7,417 children was selected to represent the 24 million noninstitutionalized children of this age in the United States. The total of 7,119, or 96 percent of the sample examined, were found to be a microcosm of children of this age in the United States with respect to age, sex, race, region, and other available demographic and socioeconomic characteristics.

The findings are shown here by age, sex, and grade in school in the form of scaled scores, raw scores, and percentile distribution of raw scores for each subtest and as standard scores or deviation IQ estimates of the Full Scale IQ based on this dyad short form of the WISC.

Comparisons for the two subtests are made with findings in Wechsler's standardization group, which was about one-fifth the size of the group of examinees in this study. In contrast to the present study, Wechsler's group was restricted to white children but did include some institutionalized children. Mean scores on the

Block Design subtest were higher in the present study than in Wechsler's at both 7½ and 10½ years, the difference being statistically significant only for the younger group. On the Vocabulary subtest the present study means were significantly lower at 7½ years but slightly higher at 10½ years than in the original standardization group. Variability in scores was greater in the present study at both ages on the Vocabulary subtest, the difference here being significant only for those 7½ years of age. Block Design scores were slightly more variable in the present study for the younger group but significantly less for the 10½-year-olds.

Boys were found to outscore girls, on the average even more consistently on the two subtests used in the survey than they did on the entire test in Wechsler's original standardization study. Since the present survey findings are based on a substantially larger, more representative sample of noninstitutionalized American children than the original study, it is apparent that there is a sex differential in performance on this test, or at least on the two subtests used here, that cannot be attributed to a factor of bias in the selection of the sample of children examined.

The degree of reliability of the two subtest combination of the WISC as an estimator of the intelligence of children based on previous research is discussed.

REFERENCES

[1] National Center for Health Statistics: Origin, program, and operation of the U.S. National Health Survey. *Vital and Health Statistics.* PHS Pub. No. 1000-Series 1-No. 1. Public Health Service. Washington. U.S. Government Printing Office, Aug. 1963.

[2] National Center for Health Statistics: Plan and initial program of the Health Examination Survey. *Vital and Health Statistics.* PHS Pub. No. 1000-Series 1-No. 4. Public Health Service. Washington. U.S. Government Printing Office, July 1965.

[3] National Center for Health Statistics: Cycle I of the Health Examination Survey, sample and response. *Vital and Health Statistics.* PHS Pub. No. 1000-Series 11-No. 1. Public Health Service. Washington. U.S. Government Printing Office, Apr. 1964.

[4] National Center for Health Statistics: Plan, operation, and response results of a program of children's examinations. *Vital and Health Statistics.* PHS Pub. No. 1000-Series 1-No. 5. Public Health Service. Washington. U.S. Government Printing Office, Oct. 1967.

[5] National Center for Health Statistics: Evaluation of psychological measures used in the Health Examination Survey of children ages 6-11. *Vital and Health Statistics.* PHS Pub. No. 1000-Series 2-No. 15. Public Health Service. Washington. U.S. Government Printing Office, Mar. 1966.

[6] Wechsler, D.: *Wechsler Intelligence Scale for Children. Manual.* New York. The Psychological Corporation, 1949.

[7] Carleton, F.O., and Stacey, C.C.: Evaluation of selected short forms of the Wechsler Intelligence Scale for Children. *J. Clin. Psychol.* 10: 258-261, 1954.

[8] Finley, C.J., and Thompson, J.: An abbreviated Wechsler Intelligence Scale of Children for use with the educable mentally retarded. *Am. J. Ment. Deficiency* 63: 473-480, 1958.

[9] Greenmun, R.: *Abbreviated Forms of the Wechsler Intelligence Scale for Children.* Unpublished master's thesis. Texas Christian University, Aug. 1965.

[10] Osborne, R.T., and Allen, J.: Validity of short forms of the WISC for mental retardates. *Psychol. Rep.* 11: 167-170, 1962.

[11] Schwartz, L., and Levitt, E.: Short forms of the Wechsler Intelligence Scale for Children in the educable non-institutionalized mentally retarded. *J. Educ. Psychol.* 51:187-190, 1960.

[12] Silverstein, A.B.: Validity of WISC short forms at three age levels. *J. Consult. Psychol.* 31(6): 635-636, 1967.

[13] Simpson, W.H., and Bridges, Jr., C.C.: A short form of the Wechsler Intelligence Scale for Children. *J. Clin. Psychol.* 15: 424, 1959.

[14] Wight, B., and Sandry, M.: A short form of the Wechsler Intelligence Scale for Children. *J. Clin. Psychol.* 18: 166, 1962.

[15]Hite, L.: *Analysis of Reliability and Validity of the Wechsler Intelligence Scale for Children.* Unpublished doctoral dissertation. Western Reserve University, 1953.

[16]Hagen, E.P.: *A Factor Analysis of the Wechsler Intelligence Scale for Children.* Unpublished doctoral dissertation. Columbia University, 1952.

[17]Cohen, J.: The factorial structure of the WISC at ages 7-6, 10-6, and 13-6. *J. Consult. Psychol.* 23:285-299, 1959.

[18]Gault, U.: Factorial patterns on the Wechsler Intelligence Scales. *Aust. J. Psychol.* 6:85-90, 1954.

[19]National Center for Health Statistics: Sample design and estimation procedures for a National Health Examination Survey of children. *Vital and Health Statistics.* PHS. Pub. No. 1000-Series 2. Public Health Service. In publication.

[20]Beeman G.: A comparative study of the WISC and Stanford-Binet with a group of more able and gifted 7-11 year old students. *California J. Educ. Res.* 11:77, 1960.

[21]Lucito, L., and Gallagher, J.: Intellectual patterns of highly gifted children on the WISC. *Peabody J. Educ.* 38: 131-136, 1960.

[22]Muhr, J.P.: *Validity of the Wechsler Intelligence Scale for Children at the five and six year level.* Unpublished master's thesis. University of Detroit, 1952.

[23]Pastovic, J.J., and Guthrie, G.M.: Some evidence on the validity of the WISC. *J. Consult. Psychol.* 15:385-386, 1951.

[24]Cohen, B.D., and Collier, M.J.: A note on WISC and other tests of children six to eight years old. *J. Consult. Psychol.* 16:226-227, 1952.

[25]Schachter, F.F., and Apgar, V.: Comparison of preschool Stanford-Binet and school-age WISC IQ's. *J. Educ. Psychol.* 49: 320-323, 1958.

[26]Triggs, F.O., and Cartee, J.K.: Pre-school pupil performance on the Stanford-Binet and the Wechsler Intelligence Scale for Children. *J. Clin. Psychol.* 9:27-29, 1953.

[27]Rottersman, L.: *A Comparison of the IQ Scores on the New Revised Stanford-Binet, Form L, the Wechsler Intelligence Scale for Children, and the Goodenough "Draw A Man" Test at the Six Year Age Level.* Unpublished master's thesis. University of Nebraska, 1950.

[28]Jones, S.: The Wechsler Intelligence Scale for Children applied to a sample of London primary school children. *Br. J. Educ. Psychol.* 32(2):119-133, 1962.

[29]Wagner, W.K.: *A Comparison of Stanford-Binet Mental Ages and Scaled Scores on the Wechsler Intelligence Scale for Children for Fifty Bowling Green Pupils.* Unpublished master's thesis. Bowling Green State University, 1951.

[30]Frandsen, A.N., and Higginson, J.B.: The Stanford Binet and the Wechsler Intelligence Scale for Children. *J. Consult. Psychol.* 15:236-238, 1951.

[31]Seashore, H., Wesman, A., and Doppelt, J.: The standardization of the Wechsler Intelligence Scale for Children. *J. Consult. Psychol.* 14:99-110, 1950.

[32]Terman, L.M., and Merrill, M.A.: *Measuring Intelligence.* Boston. Houghton Mifflin Co., 1937.

———— ○○○ ————

LIST OF DETAILED TABLES

Table 1. Mean and standard deviation (SD) of raw scores on the Vocabulary subtest of the WISC by age and sex and mean scores by grade, age, and sex for noninstitutionalized children: United States, 1963-65

Age and sex	Total		Grade in school								
	Mean	SD	Kindergarten	First	Second	Third	Fourth	Fifth	Sixth	Seventh	Special ungraded class
			Mean raw score								
Boys and girls											
6-11 years--	25.6	...	15.8	16.4	20.2	24.6	28.8	32.5	37.0	38.9	16.8
6 years-----	16.4	5.57	15.8	16.4	18.3	...	...	...	...	...	14.6
7 years-----	19.8	6.64	*	16.8	20.4	23.2	...	...	...	...	15.6
8 years-----	24.2	7.34	...	14.2	21.5	25.0	27.9	...	...	...	13.2
9 years-----	27.9	8.15	...	15.4	16.9	25.1	29.4	31.3	...	...	14.4
10 years----	31.6	9.00	...	*	13.9	20.6	29.0	33.2	37.1	*	16.1
11 years----	34.9	9.65	...	...	43.0	23.7	23.1	31.5	37.0	38.9	24.2
Boys											
6-11 years--	26.4	...	16.1	17.1	20.9	25.4	30.5	33.6	38.6	40.2	17.6
6 years-----	16.9	5.65	16.0	17.0	18.8	...	...	...	...	...	13.1
7 years-----	20.6	7.09	*	17.7	21.3	24.7	...	...	...	...	15.6
8 years-----	24.9	7.83	...	13.6	22.1	25.9	28.9	...	...	...	14.8
9 years-----	28.9	8.18	...	17.9	16.8	25.1	31.1	32.5	...	...	15.7
10 years----	32.7	9.25	...	*	15.2	23.1	30.9	34.6	38.7	...	15.8
11 years----	35.6	10.04	...	...	16.9	22.0	25.4	32.2	38.6	40.2	24.8
Girls											
6-11 years--	24.8	...	15.5	15.7	19.4	23.7	27.0	31.4	35.6	38.0	15.2
6 years-----	15.8	5.59	15.5	15.8	17.9	...	...	...	...	...	16.5
7 years-----	19.0	6.30	...	15.6	19.4	21.8	...	...	...	...	14.4
8 years-----	23.5	7.12	...	14.9	20.7	24.1	26.8	...	...	...	11.0
9 years-----	26.8	8.06	...	12.4	16.6	25.0	27.7	29.9	...	...	12.0
10 years----	30.4	8.72	...	...	9.4	16.7	26.5	32.0	35.6	*	16.0
11 years----	34.1	9.34	...	...	11.0	25.0	19.2	30.2	35.6	38.0	21.6

Table 2. Mean, standard deviation, and coefficient of variation of raw scores on the Vocabulary and Block Design subtests of the WISC for noninstitutionalized children by 4-month age intervals: United States, 1963-65

Age interval	Vocabulary subtest			Block Design subtest		
	Mean	Standard deviation	Coefficient of variation	Mean	Standard deviation	Coefficient of variation
	Raw score			Raw score		
6 years:						
0-3 months	15.0	5.7	.38	4.8	3.8	.79
4-7 months	16.3	5.1	.31	5.6	4.6	.82
8-11 months	17.7	5.8	.33	6.8	5.1	.75
7 years:						
0-3 months	18.4	6.5	.35	7.0	5.4	.77
4-7 months	19.6	6.7	.34	8.3	6.2	.75
8-11 months	21.4	6.7	.31	8.5	5.9	.69
8 years:						
0-3 months	22.9	7.4	.32	10.5	7.6	.72
4-7 months	23.7	7.3	.31	10.7	7.8	.73
8-11 months	26.0	7.3	.28	12.4	8.6	.69
9 years:						
0-3 months	26.7	7.6	.28	12.9	8.7	.67
4-7 months	28.0	8.5	.30	12.8	9.0	.70
8-11 months	29.0	8.3	.29	14.3	9.9	.69
10 years:						
0-3 months	30.3	9.3	.31	16.4	10.6	.65
4-7 months	31.9	8.8	.28	18.1	11.6	.64
8-11 months	32.7	8.9	.27	19.3	12.2	.63
11 years:						
0-3 months	33.3	9.1	.27	19.6	11.4	.58
4-7 months	34.8	10.0	.29	20.8	13.6	.65
8-11 months	36.2	9.8	.27	22.9	13.1	.57

Table 3. Mean and standard deviation of raw scores on the Block Design subtest of the WISC by age and sex and mean scores by grade, age, and sex for noninstitutionalized children: United States, 1963-65

Age and sex	Total		Grade in school								
	Mean	SD	Kinder-garten	First	Second	Third	Fourth	Fifth	Sixth	Seventh	Special ungraded class
			Mean raw score								
Boys and girls											
6-11 years--	12.7	...	6.2	5.7	8.2	11.1	14.1	18.8	22.8	26.6	6.8
6 years-----	5.7	4.51	6.3	5.5	7.8	...	...	...	...	...	4.5
7 years-----	8.0	5.86	*	6.6	8.2	9.8	...	...	...	...	4.3
8 years-----	11.2	8.04	...	4.3	9.2	11.6	14.8	...	...	...	3.7
9 years-----	13.3	9.23	...	4.6	5.6	10.6	14.1	17.6	...	...	5.9
10 years----	17.9	11.48	...	*	5.3	10.6	14.4	19.6	22.1	*	7.5
11 years----	21.2	12.76	...	...	11.0	11.6	11.3	17.2	23.0	26.6	11.2
Boys											
6-11 years--	13.3	...	7.3	5.9	8.7	12.0	15.2	19.7	24.4	28.8	7.0
6 years-----	5.8	5.04	7.4	5.5	7.6	...	...	...	...	...	3.7
7 years-----	8.5	6.68	*	7.4	8.9	10.5	...	...	...	...	4.1
8 years-----	12.0	8.81	...	4.3	9.4	12.8	16.0	...	...	...	2.6
9 years-----	14.0	9.56	...	3.4	6.6	10.4	15.4	18.9	...	...	4.7
10 years----	18.2	11.60	...	*	5.3	12.5	14.9	20.4	22.4	...	7.7
11 years----	22.3	13.00	...	...	5.0	12.8	12.6	18.5	24.8	28.8	12.5
Girls											
6-11 years--	12.1	...	5.2	5.5	7.7	10.2	13.1	17.8	21.3	25.1	6.4
6 years-----	5.7	4.10	5.2	5.5	7.8	...	...	...	...	...	5.7
7 years-----	7.3	4.88	...	5.6	7.5	9.2	...	...	...	...	5.1
8 years-----	10.3	7.14	...	4.1	8.9	10.3	13.6	...	...	...	4.6
9 years-----	12.6	8.85	...	4.9	4.0	10.8	13.0	16.1	...	...	7.3
10 years----	17.5	11.42	...	...	5.4	7.8	13.8	18.8	21.6	*	6.8
11 years----	20.1	12.59	...	...	2.0	8.3	9.0	15.1	21.3	25.1	7.2

Table 4. Mean scaled scores on the Vocabulary subtest of the WISC for noninstitutionalized children, by grade, age, and sex: United States, 1963-65

Age and sex	Grade in school									
	Total	Kindergarten	First	Second	Third	Fourth	Fifth	Sixth	Seventh	Special ungraded class
	Mean scaled scores[1]									
Boys and girls										
6 years	10	10	10	11	...	...	...	...	...	9
7 years	10	*	9	10	12	...	...	...	...	8
8 years	10	...	6	9	11	12	...	...	...	6
9 years	10	...	5	6	9	11	...	...	...	5
10 years	10	...	*	4	6	9	10	12	*	5
11 years	10	...	...	12	7	6	9	11	11	7
Boys										
6 years	10	10	10	12	...	...	...	...	...	8
7 years	10	*	9	11	12	...	...	...	...	8
8 years	11	...	6	9	11	12	...	...	...	6
9 years	10	...	6	6	9	11	12	...	...	6
10 years	10	...	*	4	7	10	11	12	...	5
11 years	10	...	...	5	6	7	9	11	12	7
Girls										
6 years	10	10	10	11	...	...	...	...	...	10
7 years	10	...	8	10	11	...	...	...	...	7
8 years	10	...	6	9	10	11	...	...	...	5
9 years	10	...	4	6	9	10	11	...	...	4
10 years	10	...	...	2	5	8	10	11	*	5
11 years	10	...	...	3	7	5	9	10	11	6

[1]Scaled scores with mean set at 10 and standard deviation of 3.

Table 5. Mean scaled scores on the Block Design subtest of the WISC for noninstitutionalized children, by grade, age, and sex: United States, 1963-65

Age and sex	Grade in school									
	Total	Kinder-garten	First	Second	Third	Fourth	Fifth	Sixth	Seventh	Special ungraded class
Boys and girls	Mean scaled scores[1]									
6 years	10	10	10	12	...	...	...	...	...	9
7 years	10	*	9	10	11	...	...	...	...	8
8 years	10	...	7	9	10	12	...	...	...	7
9 years	10	...	7	8	10	12	...	...	...	8
10 years	10	...	*	7	8	9	10	11	*	7
11 years	10	...	...	8	8	8	9	10	11	8
Boys										
6 years	10	11	10	12	...	...	...	...	...	9
7 years	10	*	9	10	11	...	...	...	...	8
8 years	11	...	7	9	11	12	...	...	...	7
9 years	10	...	7	8	9	11	12	...	...	7
10 years	10	...	*	7	9	9	10	11	...	7
11 years	10	...	...	6	8	8	9	11	12	8
Girls										
6 years	10	10	10	12	...	...	...	...	...	10
7 years	9	...	9	10	10	...	...	...	...	8
8 years	10	...	7	9	10	11	...	...	...	8
9 years	10	...	7	7	9	10	11	...	...	8
10 years	10	...	...	7	7	9	10	11	*	7
11 years	10	...	...	5	7	7	9	10	11	7

[1]Scaled scores with mean set at 10 and standard deviation of 3.

Table 6. Scaled score equivalents for raw scores on the Vocabulary subtests of the WISC for non-institutionalized children, by 4-month age intervals: United States, 1963-65

Raw score	6 years			7 years		
	0-3 months	4-7 months	8-11 months	0-3 months	4-7 months	8-11 months
	Scaled score[1]					
00	02	00	01	02	01	00
01	03	01	01	02	02	01
02	03	02	02	02	02	01
03	04	02	02	03	03	02
04	04	03	03	03	03	02
05	05	03	03	04	03	03
06	05	04	04	04	04	03
07	06	05	04	05	04	04
08	06	05	05	05	05	04
09	07	06	06	06	05	04
10	07	06	06	06	06	05
11	08	07	07	07	06	05
12	08	08	07	07	07	06
13	09	08	08	08	07	06
14	09	09	08	08	07	07
15	10	09	09	08	08	07
16	11	10	09	09	08	08
17	11	10	10	09	09	08
18	12	11	10	10	09	08
19	12	12	11	10	10	09
20	13	12	11	11	10	09
21	13	13	12	11	11	10
22	14	13	12	12	11	10
23	14	13	13	12	12	11
24	15	15	13	13	12	11
25	15	15	14	13	12	12
26	16	16	14	13	13	12
27	16	16	15	14	13	12
28	17	17	15	14	14	13
29	17	17	16	15	14	13
30	18	18	16	15	15	14
31	18	19	17	16	15	14
32	19	19	17	16	16	15
33	19	20	18	17	16	15
34	20	20	18	17	16	16
35	20	21	19	18	16	16
36	21	22	19	18	17	16
37	21	22	20	19	18	17
38	22	23	20	19	18	17
39	23	23	21	19	19	18

[1] Mean of raw score distribution set at 10 and standard deviation at 3.

Table 6. Scaled score equivalents for raw scores on the Vocabulary subtests of the WISC for non-institutionalized children, by 4-month age intervals: United States, 1963-65—Con.

8 years			9 years			10 years			11 years		
0-3 months	4-7 months	8-11 months	0-3 months	4-7 months	8-11 months	0-3 months	4-7 months	8-11 months	0-3 months	4-7 months	8-11 months
Scaled score[1]											
01	00	...	...	...	...	...	...	...	...	...	...
01	01	00	00	00	00	01	...	...	...	...	...
02	01	00	00	01	00	01	00	00	00	00	00
02	01	01	01	01	01	01	00	00	00	00	00
02	02	01	01	02	01	02	00	00	00	01	00
03	02	01	01	02	01	02	01	01	01	01	00
03	03	02	02	02	02	02	01	01	01	01	01
04	03	02	02	03	02	02	01	01	01	02	01
04	04	03	03	03	02	03	02	02	02	02	01
04	04	03	03	03	03	03	02	02	02	02	02
05	04	03	03	04	03	03	03	02	02	03	02
05	05	04	04	04	04	04	03	03	03	03	02
06	05	04	04	04	04	04	03	03	03	03	03
06	06	05	05	05	04	04	04	03	03	03	03
06	06	05	05	05	05	05	04	04	04	04	03
07	06	05	05	05	05	05	04	04	04	04	04
07	07	06	06	06	05	05	05	04	04	04	04
08	07	06	06	06	06	06	05	05	05	05	04
08	08	07	07	06	06	06	05	05	05	05	04
08	08	07	07	07	06	06	06	05	05	05	05
09	08	08	07	07	07	07	06	06	06	06	05
09	09	08	08	08	07	07	06	06	06	06	05
10	09	08	08	08	07	07	07	06	06	06	06
10	10	09	09	08	08	08	07	07	07	06	06
10	10	09	09	09	08	08	07	07	07	07	06
11	11	10	09	09	09	08	08	07	07	07	07
11	11	10	10	09	09	09	08	08	08	07	07
12	11	10	10	10	09	09	08	08	08	08	07
12	12	11	11	10	10	09	09	08	08	08	07
12	12	11	11	10	10	10	09	09	09	08	08
13	13	12	11	11	10	10	09	09	09	09	08
13	13	12	12	11	11	10	10	09	09	09	08
14	13	12	12	11	11	11	10	10	10	09	09
14	14	13	12	12	11	11	10	10	10	09	09
14	14	13	13	12	12	11	11	10	10	10	09
15	15	14	13	12	12	12	11	11	11	10	10
15	15	14	14	13	13	12	11	11	11	10	10
16	15	15	14	13	13	12	12	11	11	11	10
16	16	15	14	14	13	12	12	12	12	11	11
17	16	15	15	14	14	13	12	12	12	11	11

[1]Mean of raw score distribution set at 10 and standard deviation at 3.

Table 6. Scaled score equivalents for raw scores on the Vocabulary subtests of the WISC for non-institutionalized children, by 4-month age intervals: United States, 1963-65—Con.

Raw score	6 years			7 years		
	0-3 months	4-7 months	8-11 months	0-3 months	4-7 months	8-11 months
	Scaled score[1]					
40	23	24	22	20	19	18
41	24	24	22	20	20	19
42	24	25	23	21	20	19
43	25	26	23	21	21	20
44	25	26	24	22	21	20
45	26	27	24	22	21	20
46	26	27	25	23	22	21
47	27	28	25	23	22	21
48	27	29	26	24	23	22
49	28	29	26	24	23	22
50	28	30	27	25	24	23
51	29	30	27	25	24	23
52	29	31	28	25	25	24
53	30	31	28	26	25	24
54	30	32	29	26	25	25
55	...	...	29	27	26	25
56	...	...	30	27	26	25
57	...	...	30	28	27	26
58	...	...	31	28	27	26
59	...	...	31	29	28	27
60	...	...	32	29	28	27
61	...	...	32	30	29	28
62	...	...	33	30	29	28
63	...	...	33	30	30	29
64	...	...	...	...	30	29
65	...	...	...	...	30	29
66	...	...	...	...	31	30
67	...	...	...	...	31	30
68	...	...	...	...	...	...
69	...	...	...	...	...	...
70	...	...	...	...	...	...
71	...	...	...	...	...	...
72	...	...	...	...	...	...
73	...	...	...	...	...	...
74	...	...	...	...	...	...
75	...	...	...	...	...	...
76	...	...	...	...	...	...
77	...	...	...	...	...	...
78	...	...	...	...	...	...
79	...	...	...	...	...	...
80	...	...	...	...	...	...

[1] Mean of raw score distribution set at 10 and standard deviation at 3.

Table 6. Scaled score equivalents for raw scores on the Vocabulary subtests of the WISC for non-institutionalized children, by 4-month age intervals: United States, 1963-65—Con.

8 years			9 years			10 years			11 years		
0-3 months	4-7 months	8-11 months	0-3 months	4-7 months	8-11 months	0-3 months	4-7 months	8-11 months	0-3 months	4-7 months	8-11 months
Scaled score[1]											
17	17	16	15	14	14	13	13	12	12	12	11
17	17	16	16	15	14	13	13	13	13	12	11
18	18	17	16	15	15	14	13	13	13	12	12
18	18	17	16	15	15	14	14	13	13	12	12
19	18	17	17	16	15	14	14	14	14	13	12
19	19	18	17	16	16	15	14	14	14	13	13
19	19	18	18	16	16	15	15	14	14	13	13
20	20	19	18	17	16	15	15	15	14	14	13
20	20	19	18	17	17	16	15	15	15	14	14
21	20	19	19	17	17	16	16	15	15	14	14
21	21	20	19	18	18	16	16	16	15	15	14
21	21	20	20	18	18	17	17	16	16	15	15
22	22	21	20	18	18	17	17	17	16	15	15
22	22	21	20	19	19	17	17	17	16	15	15
23	22	22	21	19	19	18	18	17	17	16	15
23	23	22	21	20	19	18	18	18	17	16	16
23	23	22	22	20	20	18	18	18	17	16	16
24	24	23	22	20	20	19	19	18	18	17	16
24	24	23	22	21	20	19	19	19	18	17	17
25	25	24	23	21	21	19	19	19	18	17	17
25	25	24	23	21	21	20	20	19	19	18	17
25	25	24	24	22	22	20	20	20	19	18	18
26	26	25	24	22	22	20	20	20	19	18	18
26	26	25	24	22	22	21	21	20	20	18	18
27	27	26	25	23	23	21	21	21	20	19	18
27	27	26	25	23	23	21	21	21	20	19	19
27	27	26	26	23	23	22	22	21	21	19	19
28	28	27	26	24	24	22	22	22	21	20	19
28	28	27	26	24	24	22	22	22	21	20	20
29	29	28	27	24	24	22	23	22	22	20	20
29	29	28	27	25	25	23	23	23	22	21	20
29	29	29	27	25	25	23	23	23	22	21	21
30	30	29	28	26	25	23	24	23	23	21	21
30	30	29	28	26	26	24	24	24	23	21	21
...	...	30	29	26	26	24	24	24	23	22	22
...	...	30	29	27	27	24	25	24	24	22	22
...	...	31	29	27	27	25	25	25	24	22	22
...	...	31	30	27	27	25	25	25	24	23	22
...	...	31	30	28	28	25	26	25	25	23	23
...	...	...	...	28	28	26	26	26	25	23	23
...	...	...	...	28	28	26	26	26	25	24	23

[1]Mean of raw score distribution set at 10 and standard deviation at 3.

Table 7. Scaled score equivalents for raw scores on the Block Design subtest of the WISC for noninstitutionalized children, by 4-month intervals: United States, 1963-65

Raw score	6 years			7 years		
	0-3 months	4-7 months	8-11 months	0-3 months	4-7 months	8-11 months
	Scaled score[1]					
00	06	06	06	06	06	06
01	07	07	07	07	06	06
02	08	08	07	07	07	07
03	09	08	08	08	07	07
04	09	09	08	08	08	08
05	10	10	09	09	08	08
06	11	10	10	09	09	09
07	12	11	10	10	09	09
08	13	12	11	11	10	10
09	13	12	11	11	10	10
10	14	13	12	12	11	11
11	15	14	13	12	11	11
12	16	14	13	13	12	12
13	17	15	14	13	12	12
14	17	15	14	14	13	13
15	18	16	15	14	13	13
16	19	17	15	15	14	14
17	20	17	16	16	14	14
18	21	18	17	16	15	15
19	21	19	17	17	15	15
20	22	19	18	17	16	16
21	23	20	18	18	16	16
22	24	21	19	18	17	17
23	25	21	20	19	17	17
24	25	22	20	19	18	18
25	26	23	21	20	18	18
26	27	23	21	21	19	19
27	28	24	22	21	19	19
28	29	25	23	22	19	20
29	29	25	23	22	20	20
30	30	26	24	23	20	21
31	31	27	24	23	21	21
32	32	27	25	24	21	22
33	33	28	26	24	22	22
34	33	28	26	25	22	23
35	34	29	27	26	23	23
36	35	30	27	26	23	24
37	36	30	28	27	24	24
38	...	...	29	27	24	25
39	...	...	29	28	25	25
40	...	...	30	28	25	26
41	...	...	30	29	26	26
42	...	...	31	30	26	27
43	...	...	31	30	27	27
44	...	...	...	...	27	28
45	...	...	...	...	28	28
46	...	...	...	...	28	29
47	...	...	...	...	29	29
48	...	...	...	...	29	30
49	...	...	...	...	30	30
50	...	...	...	...	30	31
51	...	...	...	...	...	...
52	...	...	...	...	...	...
53	...	...	...	...	...	...
54	...	...	...	...	...	...
55	...	...	...	...	...	...

[1]Mean of raw score distribution set at 10 and standard deviation at 3.

Table 7. Scaled score equivalents for raw scores on the Block Design subtest of the WISC for noninstitutionalized children, by 4-month intervals: United States, 1963-65—Con.

8 years			9 years			10 years			11 years		
0-3 months	4-7 months	8-11 months	0-3 months	4-7 months	8-11 months	0-3 months	4-7 months	8-11 months	0-3 months	4-7 months	8-11 months
Scaled score[1]											
06	06	06	06	06	06	05	05	05	05	05	05
06	06	06	06	06	06	06	06	06	05	06	05
07	07	06	06	06	06	06	06	06	05	06	05
07	07	07	07	07	07	06	06	06	06	06	05
07	07	07	07	07	07	07	06	06	06	06	06
08	08	07	07	07	07	07	07	06	06	07	06
08	08	08	08	08	07	07	07	07	06	07	06
09	09	08	08	08	08	07	07	07	07	07	06
09	09	08	08	08	08	08	07	07	07	07	07
09	09	09	09	09	08	08	08	07	07	07	07
10	10	09	09	09	09	08	08	08	07	08	07
10	10	10	09	09	09	08	08	08	08	08	07
11	11	10	10	10	09	09	08	08	08	08	08
11	11	10	10	10	10	09	09	08	08	08	08
11	11	11	10	10	10	09	09	09	09	09	08
12	12	11	11	11	10	10	09	09	09	09	08
12	12	11	11	11	11	10	09	09	09	09	08
13	12	12	11	11	11	10	10	09	09	09	09
13	13	12	12	12	11	10	10	10	10	09	09
13	13	12	12	12	11	11	10	10	10	10	09
14	14	13	12	12	12	11	10	10	10	10	09
14	14	13	13	13	12	11	11	10	10	10	10
14	14	13	13	13	12	12	11	11	11	10	10
15	15	14	14	13	13	12	11	11	11	10	10
15	15	14	14	14	13	12	12	11	11	11	10
16	16	14	14	14	13	12	12	11	11	11	10
16	16	15	15	14	14	13	12	12	12	11	11
16	16	15	15	15	14	13	12	12	12	11	11
17	17	15	15	15	14	13	13	12	12	12	11
17	17	16	16	15	14	14	13	12	12	12	11
18	17	16	16	16	15	14	13	13	13	12	12
18	18	16	16	16	15	14	13	13	13	12	12
18	18	17	17	16	15	14	14	13	13	12	12
19	19	17	17	17	16	15	14	13	14	13	12
19	19	17	17	17	16	15	14	14	14	13	13
20	19	18	18	17	16	15	14	14	14	13	13
20	20	18	18	18	17	16	15	14	14	13	13
20	20	19	18	18	17	16	15	14	15	14	13
21	21	19	19	18	17	16	15	15	15	14	13
21	21	19	19	19	17	16	15	15	15	14	14
22	21	20	19	19	18	17	16	15	15	14	14
22	22	20	20	19	18	17	16	15	16	14	14
22	22	20	20	20	18	17	16	16	16	15	14
23	22	21	20	20	19	18	16	16	16	15	15
23	23	21	21	20	19	18	17	16	16	15	15
24	23	21	21	21	19	18	17	16	17	15	15
24	24	22	21	21	20	18	17	17	17	16	15
24	24	22	22	21	20	19	17	17	17	16	16
25	24	22	22	22	20	19	18	17	17	16	16
25	25	23	23	22	21	19	18	17	18	16	16
25	25	23	23	22	21	19	18	18	18	16	16
26	26	23	23	23	21	20	19	18	18	17	16
26	26	24	24	23	21	20	19	18	19	17	17
27	26	24	24	23	22	20	19	18	19	17	17
27	27	24	24	24	22	21	19	19	19	17	17
27	27	25	25	24	22	21	20	19	19	18	17

[1]Mean of raw score distribution set at 10 and standard deviation at 3.

Table 8. Mean standard scores on the Vocabulary and Block Design short form of the WISC for noninstitutionalized children, by grade, age, and sex: United States, 1963-65

Age and sex	Grade in school									
	Total	Kinder-garten	First	Second	Third	Fourth	Fifth	Sixth	Seventh	Special ungraded class
	Mean standard score[1]									
Boys and girls										
6-11 years--	100.0	103.3	98.4	99.4	99.7	100.3	101.4	103.6	105.0	82.1
6 years----------	100.0	103.6	99.9	103.4	...	...	...	...	...	96.5
7 years----------	100.1	*	95.9	101.1	104.9	...	...	...	...	89.1
8 years----------	99.9	...	82.9	95.9	101.3	105.0	...	...	...	80.0
9 years----------	100.1	...	81.5	82.5	95.6	102.5	105.8	...	...	80.3
10 years----------	100.0	...	*	75.6	86.3	96.0	102.7	106.4	*	79.0
11 years----------	100.0	...	...	75.5	85.7	84.5	95.4	103.1	105.1	84.3
Boys										
6-11 years--	101.5	106.3	99.5	100.6	101.1	102.7	103.0	106.0	107.6	83.1
6 years----------	101.1	106.5	100.9	103.5	...	...	...	...	...	91.0
7 years----------	102.0	*	98.1	103.2	107.5	...	...	...	...	87.7
8 years----------	101.5	...	81.7	96.7	103.5	107.7	...	...	...	78.6
9 years----------	101.8	...	79.4	82.6	95.1	105.4	108.4	...	...	79.7
10 years----------	101.3	...	*	75.6	89.5	98.0	104.5	108.0	...	78.2
11 years----------	101.4	...	...	74.4	83.4	86.8	96.9	105.6	107.6	85.0
Girls										
6-11 years--	98.4	100.2	97.1	98.2	98.2	97.9	99.8	101.3	103.3	80.2
6 years----------	98.9	100.2	98.8	103.2	...	...	...	...	...	98.8
7 years----------	98.2	...	92.8	99.1	102.0	...	...	...	...	85.9
8 years----------	98.2	...	82.6	94.5	99.2	102.4	...	...	...	76.6
9 years----------	98.3	...	75.9	80.5	96.0	99.7	102.7	...	...	77.7
10 years----------	98.6	...	...	71.0	79.8	93.0	101.0	104.5	*	77.3
11 years----------	98.5	...	...	65.0	82.4	79.2	93.0	100.7	103.3	78.4

[1]Mean of combined distribution of Vocabulary and Block Design scaled scores set at 100, standard deviation at 15.

Table 9. Standard score equivalents of sum of scaled scores from WISC short form—Vocabulary and Block Design subtests—for noninstitutionalized children, by 4-month age intervals: United States, 1963-65

Sum of two scaled scores	6 years			7 years		
	0-3 months	4-7 months	8-11 months	0-3 months	4-7 months	8-11 months
	Standard score					
00	39	39	39	38	39	40
01	42	42	42	41	42	43
02	45	45	45	44	45	46
03	48	48	48	48	48	49
04	51	51	51	51	51	52
05	54	54	54	54	54	55
06	57	57	57	57	57	58
07	60	60	60	60	60	61
08	63	63	63	63	63	64
09	66	66	66	66	66	67
10	69	69	69	69	69	70
11	72	72	72	72	72	73
12	75	76	75	75	75	76
13	78	79	78	78	78	79
14	81	82	81	81	81	82
15	85	85	84	84	84	85
16	88	88	87	87	87	88
17	91	91	90	90	90	91
18	94	94	93	93	93	94
19	97	97	96	96	96	97
20	100	100	99	99	99	100
21	103	103	102	102	102	103
22	106	106	106	106	105	106
23	109	109	109	109	108	109
24	112	112	112	112	112	112
25	115	115	115	115	115	115
26	118	118	118	118	118	118
27	121	121	121	121	121	121
28	124	124	124	124	124	124
29	127	127	127	127	127	127
30	130	130	130	130	130	130
31	133	134	133	133	133	133
32	136	137	136	136	136	136
33	139	140	139	139	139	139
34	142	143	142	142	142	142
35	145	146	145	145	145	145
36	148	149	148	148	148	148
37	151	152	151	151	151	151
38	154	155	154	154	154	154
39	158	158	157	157	157	157
40	161	161	160	160	160	160
41	164	164	163	164	163	163
42	167	167	166	167	166	166
43	170	170	169	170	169	169
44	173	173	172	173	172	172
45	176	176	175	176	175	175

Table 9. Standard score equivalents of sum of scaled scores from WISC short form—Vocabulary and Block Design subtests—for noninstitutionalized children, by 4-month age intervals: United States, 1963-65—Con.

8 years			9 years			10 years			11 years		
0-3 months	4-7 months	8-11 months	0-3 months	4-7 months	8-11 months	0-3 months	4-7 months	8-11 months	0-3 months	4-7 months	8-11 months
Standard score											
40	39	39	41	41	40	41	42	41	41	40	43
43	42	42	44	44	43	44	45	44	44	44	45
46	45	45	47	47	46	47	48	47	47	46	48
49	48	48	50	49	49	50	51	50	49	49	51
52	51	51	53	52	52	52	53	53	52	52	54
55	54	54	56	55	55	55	56	56	55	55	57
58	57	57	59	58	58	58	59	59	58	58	60
61	60	60	62	61	61	61	62	62	61	61	63
64	63	63	65	64	64	64	65	65	64	64	65
67	66	66	68	67	67	67	68	67	67	67	68
70	69	69	71	70	70	70	71	70	70	70	71
73	72	73	74	73	73	73	74	73	73	73	74
76	75	76	76	76	76	76	77	76	76	76	77
79	78	79	79	79	79	79	80	79	79	79	80
82	81	82	82	82	82	82	82	82	82	82	83
85	84	85	85	85	85	85	85	85	85	85	85
88	87	88	88	88	88	88	88	88	88	88	88
91	90	91	91	91	91	91	91	91	91	91	91
94	93	94	94	94	94	94	94	94	94	94	94
97	97	97	97	97	96	97	97	97	97	97	97
100	100	100	100	100	99	100	100	100	99	100	100
103	103	103	103	103	102	103	103	103	102	102	103
106	106	106	106	105	105	106	106	105	105	105	105
109	109	109	109	108	108	109	109	108	108	108	108
112	112	112	111	111	111	112	111	111	111	111	111
115	115	115	114	114	114	115	114	114	114	114	114
118	118	118	117	117	117	118	117	117	117	117	117
121	121	121	120	120	120	120	120	120	120	120	120
124	124	124	123	123	123	123	123	123	123	123	123
127	127	127	126	126	126	126	126	126	126	126	125
130	130	130	129	129	129	129	129	129	129	129	128
133	133	133	132	132	132	132	132	132	132	132	131
136	136	136	135	135	135	135	135	135	135	135	134
139	139	139	138	138	138	138	138	138	138	138	137
142	142	142	141	141	141	141	140	141	141	141	140
145	145	145	144	144	144	144	143	143	144	144	142
148	148	148	146	147	147	147	146	146	147	147	145
151	151	151	149	150	150	150	149	149	149	150	148
154	154	154	152	153	153	153	152	152	152	153	151
157	157	157	155	156	156	156	155	155	155	156	154
160	160	160	158	158	159	159	158	158	158	159	157
163	163	163	161	161	162	162	161	161	161	162	160
166	166	166	164	164	165	165	164	164	164	165	162
169	169	169	167	167	168	168	167	167	167	168	165
172	172	172	170	170	171	171	169	170	170	170	168
175	175	175	173	173	174	174	172	173	173	173	171

Table 10. Percentile equivalent of raw scores on the Vocabulary and Block Design subtests of the WISC for noninstitutionalized children, by age and sex: United States, 1963-65

Percentile and sex	Age in years													
	6-11	6	7	8	9	10	11	6-11	6	7	8	9	10	11
Boys and girls	Vocabulary raw score							Block Design raw score						
99	51	31	37	42	47	52	59	46	25	29	39	40	46	50
98	48	29	36	40	45	49	57	42	22	26	35	36	44	48
97	46	28	34	39	43	47	54	40	19	24	31	35	42	47
96	45	27	33	38	42	46	52	37	18	23	30	33	41	46
95	44	27	32	37	41	46	51	36	16	22	28	32	40	45
90	40	24	29	35	39	43	47	29	11	18	23	27	35	40
85	37	22	27	32	37	42	45	25	10	13	20	25	31	37
80	35	21	26	31	35	40	43	22	7	11	18	23	29	34
75	33	20	25	30	34	39	42	19	7	10	15	20	27	31
70	31	20	24	29	33	37	40	15	6	10	13	18	25	28
65	30	19	22	28	32	36	39	12	6	9	11	15	23	27
60	28	18	22	26	31	35	38	11	6	7	11	12	21	25
55	27	17	21	26	30	34	37	10	6	7	10	11	19	23
50	25	17	20	24	28	33	36	10	6	6	10	11	16	21
45	24	16	19	23	28	31	35	7	5	6	9	10	13	19
40	22	15	18	22	27	30	34	7	5	6	7	10	12	15
35	21	15	17	22	26	29	32	6	4	6	6	7	11	13
30	20	14	17	21	24	28	31	6	4	5	6	7	10	11
25	18	13	16	19	23	26	29	6	4	5	6	6	7	10
20	17	12	14	18	22	24	28	5	3	5	6	6	7	10
15	15	11	13	17	20	22	25	5	2	4	5	6	6	7
10	14	10	12	15	17	20	22	4	2	3	5	5	6	6
5	11	8	10	12	14	16	18	3	1	2	4	4	5	5
4	11	7	10	12	13	15	17	2	1	2	4	4	4	5
3	10	6	9	12	12	14	16	2	1	1	3	3	4	5
2	9	5	7	11	11	12	14	1	0	1	2	3	4	4
1	7	3	4	9	9	10	12	0	0	0	1	2	2	2
Boys														
99	53	31	38	44	48	52	60	46	27	31	42	42	47	51
98	49	30	36	41	46	51	58	43	24	28	38	38	45	49
97	47	29	35	40	44	49	57	41	21	26	37	36	44	48
96	46	28	34	39	43	48	55	38	19	25	32	34	42	46
95	45	27	34	39	42	47	53	37	18	24	31	33	41	46
90	41	25	30	36	39	44	48	30	11	19	24	28	35	42
85	38	23	28	34	38	43	46	26	10	15	22	26	32	38
80	36	22	27	32	37	41	44	23	8	13	18	24	29	35
75	34	21	26	31	35	40	42	20	7	11	16	21	27	32
70	32	20	25	30	34	39	41	18	6	10	14	19	25	30
65	31	20	24	28	32	37	40	14	6	10	12	16	23	28
60	29	19	22	27	32	36	39	12	6	7	11	14	21	26
55	27	18	22	26	31	35	38	11	6	7	11	12	19	24

Table 10. Percentile equivalent of raw scores on the Vocabulary and Block Design subtests of the WISC for noninstitutionalized children, by age and sex: United States, 1963-65—Con.

Percentile and sex	Age in years													
	6-11	6	7	8	9	10	11	6-11	6	7	8	9	10	11
Boys—Con.	Vocabulary raw score							Block Design raw score						
50	26	17	21	25	30	34	37	10	5	6	10	11	16	22
45	25	16	20	24	28	32	36	10	5	6	10	10	13	20
40	23	16	19	23	27	31	34	7	5	6	9	10	12	18
35	22	15	18	22	26	30	33	6	4	6	7	9	11	15
30	20	14	17	21	25	28	31	6	4	5	6	7	10	12
25	19	13	16	20	24	27	30	6	3	5	6	6	10	11
20	18	12	15	18	22	26	28	5	3	4	6	6	7	10
15	16	11	14	17	21	24	25	5	2	4	5	6	6	7
10	14	10	12	15	19	21	23	4	1	3	5	5	6	6
5	12	9	11	13	15	17	18	2	1	2	4	4	5	5
4	11	8	10	12	15	16	17	2	0	1	3	4	4	5
3	10	8	9	12	14	15	16	1	0	1	3	3	4	4
2	9	7	8	11	13	14	14	1	0	1	2	3	3	4
1	7	4	3	10	11	10	13	0	0	0	1	2	2	2
Girls														
99	49	32	36	40	45	47	58	44	22	26	34	37	45	49
98	46	28	34	38	43	46	53	41	19	23	30	36	43	48
97	45	27	32	38	42	45	51	38	18	20	28	34	42	46
96	44	26	31	37	41	45	50	36	18	19	27	32	40	45
95	42	25	30	36	40	44	49	35	15	19	26	31	39	43
90	39	23	28	33	37	42	46	28	10	13	22	26	35	39
85	36	22	26	31	35	40	44	24	9	11	19	24	31	36
80	34	21	25	30	34	38	42	21	7	10	16	21	28	33
75	32	20	24	29	33	37	41	18	7	10	13	19	26	30
70	30	19	22	28	31	36	40	14	6	9	11	16	25	27
65	29	18	22	27	30	35	38	11	6	7	11	13	23	25
60	27	18	21	26	30	34	37	11	6	7	10	11	21	23
55	26	17	20	25	29	33	36	10	6	6	10	11	19	21
50	24	16	19	24	28	32	35	9	6	6	9	10	15	19
45	23	16	18	23	27	30	34	7	5	6	7	10	12	15
40	22	15	18	22	26	29	33	6	5	6	7	7	11	13
35	20	14	17	21	25	28	32	6	5	6	6	7	10	11
30	19	14	16	20	23	26	30	6	4	5	6	6	10	11
25	18	13	15	19	22	25	29	6	4	5	6	6	7	10
20	16	12	14	18	20	23	27	5	4	5	6	6	6	9
15	15	11	13	16	18	21	25	5	3	4	5	6	6	7
10	13	10	12	14	16	18	22	4	2	4	5	5	5	6
5	11	7	10	12	13	15	18	3	1	2	4	4	4	5
4	10	6	9	12	12	14	17	3	1	2	4	4	4	5
3	10	5	8	11	12	14	16	2	1	1	3	3	4	5
2	8	4	7	11	11	12	14	2	0	1	3	2	4	4
1	6	2	6	9	8	11	12	1	0	0	1	1	3	3

APPENDIX I

TECHNICAL NOTES

The Survey Design

The sample design for the second cycle of the Health Examination Survey, similar to the one used for the first cycle, was that of a multistage, stratified probability sample of loose clusters of persons in land-based segments. Successive elements dealt with in the process of sampling are the primary sampling unit (PSU), census enumeration district (ED), segment, household, eligible child (EC), and the sample child (SC).

At the first stage, the nearly 2,000 PSU's into which the United States (including Hawaii and Alaska) had been divided and then grouped into 357 strata for use in the Current Population Survey and Health Interview Survey were further grouped into 40 super-strata for use in Cycle II of the Health Examination Survey. The average size of each Cycle II stratum was 4.5 million persons, and all strata fell between the limits of 3.5 and 5.5 million. Grouping into 40 strata was done in a way that maximized homogeneity of the PSU's included in each stratum, particularly with regard to the degree of urbanization, geographic proximity, and degree of industrialization. The 40 strata were classified into four broad geographic regions (each with 10 strata) of approximately equal population and cross-classified into four broad population density groups (each of 10 strata). Each of the 16 cells contained either two or three strata. A single stratum might include only one PSU (or only part of a PSU, for example, New York City which represented two strata) or several score PSU's.

To take account of the possible effect that the rate of population change between the 1950 and 1960 Census might have had on health, the 10 strata within each region were further classified into four classes ranging from those with no increase to those with the greatest relative increase. Each such class contained two or three strata.

One PSU was then selected from each of the 40 strata. A controlled selection technique was used in which the probability of selection of a particular PSU was proportional to its 1960 population. In the controlled selection an attempt was also made to maximize the spread of the PSU's among the States. While not every one of the 64 cells in the 4x4x4 grid contributes a PSU in the sample of 40 PSU's the controlled selection technique ensured the sample's matching the marginal distributions in all three dimensions and being closely representative of all cross-classifications.

Table I. Number of examinees, by age and sex: Health Examination Survey, 1963-65

Age	Both sexes	Boys	Girls
6-11 years	7,119	3,632	3,487
6 years	1,111	575	536
7 years	1,241	632	609
8 years	1,231	618	613
9 years	1,184	603	581
10 years	1,160	576	584
11 years	1,192	628	564

Generally, within a particular PSU, 20 ED's were selected with the probability of selection of a particular ED proportional to its population in the age group 5-9 years in the 1960 census, which by 1963 roughly approximated the population in the target age group for Cycle II. A similar method was used for selecting one segment (clusters of households) in each ED. Each of the resultant 20 segments was either a bounded area or a cluster of households (or addresses). All of the children in the age range properly resident at the address visited were EC's. Operational considerations made it necessary to reduce the number of prospective examinees at any one location to a maximum of 200. The EC's to be excluded for this reason from the SC group were determined by systematic subsampling.

The total sample included 7,417 children from 25 different States in the 6-11 year age group, with approximately 1,000 in each of the single years of age. The age-sex distribution for the 7,119 children in the total sample who were examined is shown in table I.

Reliability

Measurement processes employed in the survey were highly standardized and closely controlled. Of course, this does not mean that the correspondence between the real world and the survey results is exact. Data from the survey are imperfect for three major reasons: (1) results are subject to sampling error, (2) the actual conduct of a survey never agrees perfectly with the design, and (3) the measurement processes themselves are inexact even though standardized and controlled.

The first report on Cycle II[4] describes in detail the faithfulness with which the sample design was carried out. It notes that out of the 7,417 sample children the 7,119 who were examined—a response rate of 96 percent—gave evidence that they were a highly representative sample of children of this age in the noninstitutional population of the United States. The response levels for the various demographic subgroups—including those for age, sex, race, region, population density, parents' educational level, and family income—show no marked differentials. Hence it appears unlikely that nonresponse could bias the findings much in these respects.

Measures used to control the quality of the data from this survey in general have been cited previously;[4] those relating specifically to the Wechsler Intelligence Scale for Children are outlined in the section of this report on Field Administration.

Data recorded for each sample child are inflated in the estimation process to characterize the larger universe of which the sample child is representative. The weights used in this inflation process are a product

Table II. Number of test parts missing for examinees for the Vocabulary and Block Design subtests of the WISC: United States, 1963-65

Age and sex	Total missing parts	Both subtests	Vocabulary only	Block Design only
Boys and girls				
6-11 years	88	51	31	6
6 years	30	14	14	2
7 years	16	10	6	-
8 years	12	9	3	-
9 years	12	7	3	2
10 years	10	6	2	2
11 years	8	5	3	-
Boys				
6-11 years	32	19	9	4
6 years	10	5	3	2
7 years	8	3	5	-
8 years	2	2	-	-
9 years	6	4	-	2
10 years	4	3	1	-
11 years	2	2	-	-
Girls				
6-11 years	56	32	22	2
6 years	20	9	11	-
7 years	8	7	1	-
8 years	10	7	3	-
9 years	6	3	3	-
10 years	6	3	1	2
11 years	6	3	3	-

of the reciprocal of the probability of selecting the child, an adjustment for nonresponse cases, and a poststratified ratio adjustment which increases precision by bringing survey results into closer alignment with known U.S. population figures by color and sex within single years of age 6-11.

In the second cycle of the Health Examination Survey the sample was the result of three stages of selection—the single PSU from each stratum, the 20 segments from each sample PSU, and the sample children from the eligible children. The probability of selecting an individual child is the product of the probabilities of selection at each stage.

Since the strata are roughly equal in population size and a nearly equal number of sample children were examined in each of the sample PSU's the sample design is essentially self-weighting with respect to the target population; that is, each child 6-11 years had about the same probability of being drawn into the sample.

The adjustment upward for nonresponse is intended to minimize the impact of this factor on final estimates by imputing to nonrespondents the characteristics of "similar" respondents. Here "similar" respondents were judged to be examined children in a sample PSU having the same age (in years) and sex as children not examined in that sample PSU.

The poststratified ratio adjustment used in the second cycle achieved most of the gains in precision which would have been attained if the sample had been drawn from a population stratified by age, color, and sex and makes the final sample estimates of population agree exactly with independent controls prepared by the Bureau of the Census for the noninstitutional population of the United States as of August 1, 1964 (approximate mid-survey point), by color and sex for each single year of age 6 through 11. The weight of every responding sample child in each of the 24 age, color, and sex classes is adjusted upward or downward so that the weighted total within the class equals the independent population control.

In addition to children not examined at all, there were some whose examination was incomplete in one procedure or another. The extent of missing data for the two WISC subtests is shown in table II.

Table III. Sampling errors for average raw scores on the WISC Vocabulary and Block Design subtests by age, sex, and selected grade levels for noninstitutionalized children: United States, 1963-65

Age and sex	Vocabulary				Block Design			
	Total group	First grade	Fourth grade	Seventh grade	Total group	First grade	Fourth grade	Seventh grade
Boys and girls								
6-11 years	.42	.38	.52	.93	.29	.21	.40	1.37
Boys								
6-11 years	.47	.47	.61	1.09	.33	.33	.41	1.75
6 years	.32	.30	...	...	.27	.30	...	...
7 years	.52	1.19	...	...	.29	.83	...	...
8 years	.46	1.79	.94	...	.39	.75	1.01	...
9 years	.59	*	.61	...	.46	2.01	.54	...
10 years	.79	*	1.19	...	.63	*	.84	...
11 years	.54	...	1.17	1.09	.62	...	1.47	1.75
Girls								
6-11 years	.38	.35	.61	1.03	.31	.21	.60	1.42
6 years	.31	.29	...	...	.24	.24	...	...
7 years	.28	.91	...	...	.25	.57	...	...
8 years	.48	1.69	.64	...	.36	1.46	.99	...
9 years	.51	*	.48	...	.42	1.97	.58	...
10 years	.67	...	1.59	*	.55	...	1.57	*
11 years	.64	...	1.92	1.03	.82	...	1.39	1.43

Sampling and Measurement Error

In the present report, reference has been made to efforts to minimize bias and variability of measurement techniques.

The probability design of the survey makes possible the calculation of sampling errors. The sampling error is used here to determine how imprecise the survey test results may be because they come from a sample rather than from the measurement of all elements in the universe.

The estimation of sampling errors for a study of the type of the Health Examination Survey is difficult for at least three reasons: (1) measurement error and "pure" sampling error are confounded in the data—it is not easy to find a procedure which will either completely include both or treat one or the other separately, (2) the survey design and estimation procedure are complex and accordingly require computationally involved techniques for the calculation of variances, and (3) from the survey are coming thousands of statistics, many for subclasses of the population for which there are a small number of cases. Estimates of sampling error are obtained from the sample data and are themselves subject to sampling error which may be large when the number of cases in a cell is small or even occasionally when the number of cases is substantial.

Estimates of approximate sampling variability for selected statistics used in this report are presented in tables III and IV. The estimates have been prepared by a replication technique which yields overall variability through observation of variability among random subsamples of the total sample as described previously. This method reflects both "pure" sampling variance and a part of the measurement variance. A similar pseudo-replication technique was used to estimate the standard errors of the correlation coefficients shown in the Findings section on scaled scores.

Table IV. Sampling errors for average standard scores on the WISC short form—Vocabulary and Block Design subtests—by age, sex, and selected grade levels for noninstitutionalized children: United States, 1963-65

Age and sex	WISC short form			
	Total group	First grade	Fourth grade	Seventh grade
Boys and girls				
6-11 years	.68	.92	.82	1.45
Boys				
6-11 years	.78	1.18	.80	1.87
6 years	.89	.90	...	...
7 years	.98	2.56	...	...
8 years	.85	2.45	1.96	...
9 years	.97	*	.99	...
10 years	1.08	*	1.54	...
11 years	.80	...	1.95	1.87
Girls				
6-11 years	.64	.85	1.16	1.44
6 years	.80	.73	...	...
7 years	.62	1.85	...	...
8 years	.81	2.67	1.35	...
9 years	.84	*	.94	...
10 years	.96	...	2.58	*
11 years	.97	...	2.28	1.45

In accordance with usual practice, the interval estimate for any statistic may be considered the range within one standard error of the tabulated statistic with 68-percent confidence, or the range within two standard errors of the tabulated statistic with 95-percent confidence. The latter is used as the level of significance in this report.

An overestimate of the standard error of a difference $\underline{d}=\underline{x}-\underline{y}$ of two statistics x and y is given by the formula $S_d=(S_x^2 + S_y^2)1/2$ where S_x and S_y are the sampling errors, respectively, of x and y, shown in tables III and IV.

Small Categories

In some tables, magnitudes are shown for cells for which the sample size is so small that the sampling error may be several times as great as the statistic itself. Obviously in such instances the statistic has no meaning in itself except to indicate that the true quantity is small. Such numbers, if shown, have been included in the belief that they may help to convey an impression of the overall story of the table.

Scaled and Standard Scores

The following formula was used for computing the scaled scores shown in this report in tables 4-7:

$$S_c \; S_i = \frac{1}{s_{x_i}}(3)(X-\bar{X}_i) + 10$$

where s_{x_i} is the standard deviation of the raw scores in the i^{th} age interval, $\bar{X}_i$ is the arithmetic average or mean raw score in that age interval (both derived from the inflated sample or rational estimates) and X is the raw score for which the scaled score is being derived.

In tables 8 and 9 the following formula was used for computing standard scores or deviation intelligence quotients:

$$SS_i = \frac{1}{s_{x_i}}(15)(X-\bar{X}_i) + 100$$

where s_{x_i} here is the standard deviation of the distribution of scaled scores obtained on the two subtests combined (for each person) in the i^{th} age interval, $\bar{X}_i$ is the mean of that distribution of combined scaled scores and X the sum of the two scaled scores for which the standard score is being derived.

APPENDIX II

WISC TEST PARTS USED

Vocabulary

(Some items similar to words in the Vocabulary subtest)

1. Scooter	Score 2 or 0	5. Vocabulary
2. Pen		
3. Cape		
4. Note		
5. Insect		
	Score 2, 1 or 0	
6. Pillow		
7. Tack		
8. Elephant		
9. Fall		
10. Emerald		
40. Transverse		

Block Design

8. BLOCK DESIGN							
Design	Time	Pass-Fail	Score				
A. 45"	1		2				
	2		0	1			
B. 45"	1		2				
	2		0	1			
C. 45"	1		2				
	2		0	1			
1. 75"			0	21-75 4	16-20 5	11-15 6	1-10 7
7. 150"			0	91-150 4	66-90 5	56-65 6	1-55 7

VITAL AND HEALTH STATISTICS PUBLICATION SERIES

Formerly Public Health Service Publication No. 1000

Series 1. Programs and collection procedures.—Reports which describe the general programs of the National Center for Health Statistics and its offices and divisions, data collection methods used, definitions, and other material necessary for understanding the data.

Series 2. Data evaluation and methods research.—Studies of new statistical methodology including: experimental tests of new survey methods, studies of vital statistics collection methods, new analytical techniques, objective evaluations of reliability of collected data, contributions to statistical theory.

Series 3. Analytical studies.—Reports presenting analytical or interpretive studies based on vital and health statistics, carrying the analysis further than the expository types of reports in the other series.

Series 4. Documents and committee reports.—Final reports of major committees concerned with vital and health statistics, and documents such as recommended model vital registration laws and revised birth and death certificates.

Series 10. Data from the Health Interview Survey.—Statistics on illness, accidental injuries, disability, use of hospital, medical, dental, and other services, and other health-related topics, based on data collected in a continuing national household interview survey.

Series 11. Data from the Health Examination Survey.—Data from direct examination, testing, and measurement of national samples of the civilian, noninstitutional population provide the basis for two types of reports: (1) estimates of the medically defined prevalence of specific diseases in the United States and the distributions of the population with respect to physical, physiological, and psychological characteristics; and (2) analysis of relationships among the various measurements without reference to an explicit finite universe of persons.

Series 12. Data from the Institutional Population Surveys.—Statistics relating to the health characteristics of persons in institutions, and their medical, nursing, and personal care received, based on national samples of establishments providing these services and samples of the residents or patients.

Series 13. Data from the Hospital Discharge Survey.—Statistics relating to discharged patients in short-stay hospitals, based on a sample of patient records in a national sample of hospitals.

Series 14. Data on health resources: manpower and facilities.—Statistics on the numbers, geographic distribution, and characteristics of health resources including physicians, dentists, nurses, other health occupations, hospitals, nursing homes, and outpatient facilities.

Series 20. Data on mortality.—Various statistics on mortality other than as included in regular annual or monthly reports—special analyses by cause of death, age, and other demographic variables, also geographic and time series analyses.

Series 21. Data on natality, marriage, and divorce.—Various statistics on natality, marriage, and divorce other than as included in regular annual or monthly reports—special analyses by demographic variables, also geographic and time series analyses, studies of fertility.

Series 22. Data from the National Natality and Mortality Surveys.—Statistics on characteristics of births and deaths not available from the vital records, based on sample surveys stemming from these records, including such topics as mortality by socioeconomic class, hospital experience in the last year of life, medical care during pregnancy, health insurance coverage, etc.

For a list of titles of reports published in these series, write to:

Office of Information
National Center for Health Statistics
Public Health Service, HSMHA
Rockville, Md. 20852