

VITAL and HEALTH STATISTICS
DATA EVALUATION AND METHODS RESEARCH

Comparison of Timed and Untimed Presentation of the Goodenough-Harris Test of Intellectual Maturity

A methodological study of the effects of timed versus un-
timed administration of the human figure drawing technique
used in measuring intellectual maturity of adolescents.

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Public Health Service
Health Services and Mental Health Administration



Public Health Service Publication No. 1000 Series 2-No. 35

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C., 20402 - Price 30 cents

NATIONAL CENTER FOR HEALTH STATISTICS

THEODORE D. WOOLSEY, *Director*

PHILIP S. LAWRENCE, Sc.D., *Associate Director*

OSWALD K. SAGEN, Ph.D., *Assistant Director for Health Statistics Development*

WALT R. SIMMONS, M.A., *Assistant Director for Research and Scientific Development*

ALICE M. WATERHOUSE, M.D., *Medical Consultant*

JAMES E. KELLY, D.D.S., *Dental Advisor*

EDWARD E. MINTY, *Executive Officer*

MARGERY R. CUNNINGHAM, *Information Officer*

DIVISION OF HEALTH EXAMINATION STATISTICS

ARTHUR J. McDOWELL, *Director*

PAUL T. BRUYERE, M.D., *Deputy Director*

JAMES T. BAIRD, JR., *Chief, Analysis and Reports Branch*

HENRY W. MILLER, *Chief, Operations and Quality Control Branch*

PETER V. V. HAMILL, M.D., *Medical Advisor*

JAMES E. KELLY, D.D.S., *Dental Advisor*

HAROLD J. DUPUY, Ph.D., *Psychological Advisor*

Public Health Service Publication No. 1000-Series 2-No. 35

Library of Congress Catalog Card Number 79-600269

FOREWORD

The Health Examination Survey (HES) is one of the major continuing programs of the National Center for Health Statistics, an agency authorized by Congress to provide statistical information on the amount, distribution, and effects of illness and disability in the United States. The collection, analysis, and publication of data obtainable *only* through direct examination of people is the particular task of the HES. Examination programs for national samples of segments of our population began in November 1959 with a survey of adults between the ages of 18 and 79 (designated Cycle I). Mobile examination centers with their teams of specialists began traveling throughout the United States, setting up in diverse locations to examine individuals selected in the national probability sample. The basic pattern of operation has continued through successive surveys and has included examinations of a sample of children 6-11 years of age (Cycle II) and the current program focused on adolescents 12-17 years of age (Cycle III).

While the initial effort in the adult examination program was devoted primarily to obtaining information on several prevalent chronic diseases, when attention was directed toward younger age groups, the concern logically shifted to factors related to growth and development. At this point it became obvious that social and personal adjustment in the context of school and home is an integral part of healthy growth. Health problems of the developmental years are primarily those of retarded and disrupted growth, and the nature of personality development, as evidenced in acquisition of communication skills, general mental abilities, and interpersonal relationships, must be considered in assessment procedures. The decision of the survey planners to include some measurements of social and personality

function followed naturally, and psychologists became new members of the examining team.

Because time and physical limitations must inevitably be imposed on a comprehensive health survey, no one health factor—whether dental, physiological, physical, or psychological—can be evaluated as thoroughly as it would be in a typical clinical or research setting. As a case in point, sound, widely accepted, brief tests of the psychological factors found to be important to the goals of the survey did not exist. To cover the necessarily broad area, it was decided that the battery should be composed of either the briefest tests available or abbreviated and specially administered versions of widely used psychological instruments. The resulting battery, used in the children's survey and continued into the adolescent's survey, reflects the more frequent decision to use parts of longer tests and special administration procedures. Incumbent on the user of abbreviated tests is the need to conduct methodological studies to determine relationships between the new form and the original established instrument or other criterion measures or both. In the case of psychological data, the National Center for Health Statistics has attempted to fulfill this obligation primarily through contracts with several scientists. The study reported here is the result of one such contract.

There are some limited, but obvious and direct, benefits to be gained from this study; first, the Center is better prepared to report accurate data on the intellectual maturity of adolescents derived from the Goodenough-Harris test and, second, the readers of our reports are better prepared to fully understand these data. It is also our hope that those who are generally interested in the study and use of the human

figure drawing technique will benefit from the further exploration presented here. Those who are especially interested in the development and use of brief, survey-type instruments should be encouraged by the results of the study.

The Center has been fortunate in having Dr. Dale Harris as principal investigator on this study. In addition to the assistance he provided under this contract, Dr. Harris has served as a consultant to our program since the beginning of the children's survey. His longtime and intimate association with children's human figure drawings

has been an invaluable contribution to the survey. Our association with Dr. Harris, his colleagues, and his staff at The Pennsylvania State University has made not only the final product a worthwhile contribution, but also the entire undertaking a pleasurable experience.

Glenn D. Pinder
Assistant Psychological Advisor
Division of Health Examination
Statistics

CONTENTS

	Page
Foreword -----	iii
Introduction -----	1
Individual Versus Group Administration-----	2
Method -----	2
Results-----	2
Timed Versus Untimed Administration-----	3
Method -----	3
Results-----	8
Summary-----	12
References -----	13
Appendix I. Test Instructions-----	14
Appendix II. Test Forms-----	15

IN THIS REPORT the effects of modifications of the standard instructions for the administration of the Goodenough-Harris Drawing Test to adolescents are explored. Specifically, the question was what effect an imposed limitation on the time given to complete a human figure drawing would have on the total Goodenough-Harris score. Such a time limit was found to be necessary in order to use this test as a measure of intellectual maturity in a national survey of adolescents.

In a preliminary study, it was determined that a subject tested in a group situation achieved essentially the same score as when tested individually. This fact permitted the following study to proceed with collection of data by group administration of the test.

Two drawings, a person and the self, were obtained from a large group of 12-, 14-, and 16-year-old students. Some of the subjects were tested under the usual untimed conditions and some, with instructions imposing a 5-minute time limit. From this pool, 200 students were selected in each of the three age groups, half tested under timed and half under untimed conditions. The sample was controlled for sex, age, and race and stratified by father's occupation.

Comparison of the scores obtained on each of the two drawings under timed versus untimed conditions revealed that the 14- and 16-year-olds obtained significantly lower scores on the drawings of a person under the timed conditions. No significant differences were found for the drawings of a person of the 12-year-olds or for the drawings of the self of any age group. After closer observation of these data and a comparison with norms published for the Goodenough-Harris test, it was concluded that, although conditions of restricted time may attenuate results of the test, they apparently do not constrict scores sufficiently to invalidate the technique. With restandardization the use of the test under conditions of limited time is clearly feasible. There appears a distinct possibility of providing adequate norms for the early adolescent years from data collected in the national survey in which the timed administration was used.

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-----	*

COMPARISON OF TIMED AND UNTIMED PRESENTATION OF THE GOODENOUGH-HARRIS TEST OF INTELLECTUAL MATURITY

Dale B. Harris, Ph.D., *Department of Psychology, The Pennsylvania State University*

INTRODUCTION

The National Center for Health Statistics has for several years been conducting an intensive Health Examination Survey. Cycle III, as the current survey of the adolescent years 12-17 is known, focuses on many areas of health including aspects of both physical and mental growth and development. Previous surveys known as Cycle I and Cycle II¹ explored, respectively, the health of American adults and children. In the course of 3 hours and 30 minutes the youths receive the attention of physicians, dentists, technicians, and psychologists. Of this time, a total of 70 minutes is devoted to the psychological examination.

In the brief span of 70 minutes psychologists administer a carefully selected battery of tests designed to study intellectual development, educational achievement, and certain other aspects of social and emotional growth. The battery was chosen after consultation with a group of child psychologists from five universities and was subjected to pilot testing. The battery for Cycle III is virtually the same as that used in Cycle II with the addition of a test of literacy and a behavioral questionnaire. The basic test battery used in Cycles II and III is given in table 1.

Use of the basic test battery provides comparable information from one survey to the next and permits analysis of the data by age through the entire range—6-17 years. Such comparison, of course, presupposes standardized administration and scoring of the various test protocols.

There are available objective procedures for the Wechsler subscales, the Thematic Apperception Test, and the Wide Range Achievement Test; however, drawing protocols may be administered and scored in many ways depending on the age of the subject and the researcher's inclinations.

For the Health Examination Survey the standardized method of administering and scoring drawing tests outlined in Harris' *Children's Drawings as Measures of Intellectual Maturity*² was adopted with certain modifications. In the Goodenough-Harris procedure, subjects are asked to make three drawings in sequence—a man, a

Table 1. Test battery used by the National Center for Health Statistics in Cycles II and III of the Health Examination Survey

Title
1. Vocabulary subtest from the Wechsler Intelligence Scale for Children
2. Block design subtest from the Wechsler Intelligence Scale for Children
3. Human figure drawing
4. Selected cards from the Thematic Apperception Test
5. Wide Range Achievement Test (1963 revisions of the arithmetic and reading sections)

woman, and the self. In Cycle II, subjects were asked to draw a person and, in Cycle III, a person and the self. Scoring followed the Goodenough-Harris standards exactly, using the appropriate key for the sex of the person portrayed. When sex could not be determined from the figure, the scoring key for the male figure was used.

Although the same objective procedures were required for continuity in Cycle III, certain difficulties arose in the Cycle III pilot tests. In the Goodenough-Harris directions no time limit is given and, as a result, the youths took a surprisingly long time to draw the required "picture of a person" and "picture of the self." This meant that the standard test procedure was not adaptable to the brief 70-minute examination period. However, repeated observations were made, and it was noted that a "valid test" could be obtained from most adolescents in a specific period. That period of time proved to be about 5 minutes.

The Goodenough-Harris directions were changed from "Take your time and work very carefully" to "You will have five minutes to draw a picture. . . ." (See appendix I for complete instructions.) The changed instructions were employed throughout the Cycle III survey so that the objective scoring method could be applied to maintain continuity of data collected from one age group to another. However, the question remained: Can a valid Goodenough-Harris figure drawing test be administered under a timed condition?

The purpose of this research under contract #PH43-67-759 from the National Center for Health Statistics was to explore that question. The statistical hypothesis is that there is no difference between the scores of protocols administered under timed and untimed conditions.^a

To deal accurately yet economically with the central hypothesis of this report, it was first necessary to see whether the Goodenough-Harris test could be presented to individuals and to groups with similar results. If, indeed, the

^aA related hypothesis (although not discussed in this project) is that, if a difference does exist between timed and untimed conditions, a statistical relationship would obtain between results of the two conditions such that tables of equivalent scores could be constructed for the two conditions.

test could be administered to classroom groupings it would expedite answering the main question. The specific hypothesis in this comparison is that there is no difference between scores of protocols administered under group and under individual conditions. The report which follows, then is divided into two parts, the first addressed to the question of equivalence of group and individual administrations and the second to the main question of the equivalence of timed and untimed administrations.

INDIVIDUAL VERSUS GROUP ADMINISTRATION

Method

To assess the effect of group presentation of the Goodenough-Harris test, 102 eighth-grade schoolchildren from the Bald Eagle Area School District in Central Pennsylvania were tested using the standard Goodenough-Harris instructions.^b The children were ages 13-14 with a median age of 13 years, 8 months.

All children received both group and individual test administrations using a counterbalanced design in which 54 students were tested, first, in an individual testing session and, second, in a classroom situation (group I), and 48 students were first tested in classroom groups and second, individually (group II). The period between test administrations was approximately 4 weeks.

All test protocols were scored independently by two carefully trained scorers and differences reconciled through conference. The data were subjected to appropriate analysis to assess the effects of group versus individual presentation and to test the possible interaction of method and order of administration.

Results

The mean difference (\bar{D}) between group and individual scores was calculated for group I (individual-group) and for group II (group-individual). Each \bar{D} was tested by a *t*-test at the

^bThe field work for this portion of the research was carried out by Mrs. Mary Whaley.

Table 2. Statistical summary of individual-group hypothesis

Test administration	<i>N</i>	\bar{X}	S.D.	Mean difference	<i>t</i> for effect of method of administration	<i>t</i> for interaction of method and order of administration	
<u>Group I (individual-group)</u>							
Individual----- Group-----	} 54	40.76 39.44	7.47 9.72	} 1.31	1.79	} 0.41	
<u>Group II (group-individual)</u>							
Group----- Individual-----	} 48	35.75 36.56	7.99 7.68	} 0.81	0.86		

NOTE: *N*= number of subjects; \bar{X} = mean raw score; S.D.= standard deviation.

.05 level to see if it differed significantly from zero (table 2). The \bar{D} for group I (\bar{D}_I) was 1.31, and \bar{D} for group II (\bar{D}_{II}) was 0.81. The *t* obtained for \bar{D}_I was 1.79, and *t* obtained for \bar{D}_{II} was 0.86. Since $t_{.05}$ for 40 degrees of freedom is 2.021 and for 60 degrees of freedom is 2.000, the null hypothesis was retained. It may therefore be concluded that there is no significant difference between raw scores of tests administered under group conditions and those given under individual conditions.

To lend greater credence to the above conclusion it is necessary to show that these results were not dependent on some interaction between method of administration and order of presentation. (For a complete discussion of methods of analysis for this design see reference three.) To demonstrate, \bar{D}_{II} was subtracted from \bar{D}_I and was tested by *t* ratio at the .05 level to see if it differed significantly from zero. $\bar{D}_I - \bar{D}_{II}$ was 0.50 and yielded a *t* ratio of 0.41 which is not significant at the .05 level for 100 degrees of freedom. Therefore it was possible to retain the hypothesis that there was no interaction between type of administration and order of presentation. Thus, there is additional support for the conclusion that no essential difference exists between group and individual administration of the Goodenough-Harris Drawing Test.

TIMED VERSUS UNTIMED ADMINISTRATION

Because there is no statistical basis for assuming a fundamental difference between test results obtained under group as opposed to individual testing conditions, it was possible to conduct the main research with groups rather than on an individual basis at a very substantial savings in cost. A project was planned to examine the comparability of results obtained under timed and untimed conditions using group administrations.^c

Method

To assess the effect of timed presentation of the Goodenough-Harris Drawing Test, a total of 1,775 adolescents were examined (table 3). This group of students was drawn from 11 communities, 10 of these located throughout Pennsylvania and one in Ohio. The communities and their respective school systems were selected with a view toward covering all segments of the population—large and small towns, rural areas, and

^cThe field work for this portion of the research was carried out by Mrs. Barbara Buchanan.

Table 3. Distribution of 1,775 test subjects, by school, grade,¹ and sex

School	7th grade		9th grade		11th grade	
	Boys	Girls	Boys	Girls	Boys	Girls
	Number					
Bellefonte Area School District-----	26	26	13	12	22	38
Bristol Borough School District-----	22	29	25	29	21	12
Dalton, Ohio, School District-----	23	23	23	31	18	16
Harrisburg City School District-----	26	29	16	29	22	17
Hollidaysburg Area School District-----	21	36	35	31	19	19
Huntingdon Area School District-----	24	24	12	32	43	17
Lemoyne Area School District-----	28	30	30	27	15	39
Marple-Newtown Area Schools-----	33	36	28	39	26	21
Mt. Lebanon School District-----	39	38	21	33	19	24
Penns Hills Area School District-----	43	37	32	29	17	45
Tyrone Area School District-----	33	31	24	29	38	20

¹There were 677 students in the 7th grade, 580 in the 9th grade, and 528 in the 11th grade.

large cities and their suburbs. Table 4 describes the communities.

In each school system, two classes of students at three grade levels—seven, nine, and 11—were tested. The three grades chosen were composed predominately of students at ages 12, 14, and 16. One class at each grade level was given the timed test; the other, the untimed test.

The students tested under the timed condition were given the exact instructions used by the psychologists of the National Center for Health Statistics in Cycle III. Standard introductory instructions included the statement: "You will have five minutes to draw a person (or self)." (See appendix I for complete statement.) After 3 minutes the researcher said, "You have about two minutes." At 5 minutes the researcher said, "Are you almost finished?" The 5-minute statement was only given in the Cycle III Survey if the subject was not quite finished with his drawing. In the present study the 5-minute statement was given to all timed groups because all classes tested had subjects who were nearly finished. In Cycle III, when the 5-minute statement was given, a maximum of 2 extra minutes was allowed. Thus, the maximum time allowed to the individuals tested in Cycle III and the groups tested in the present research was 7 minutes. It is relevant to note that the majority of students in each

timed class were able to finish their drawings in the allotted 5-minute period.

In the untimed groups the instructions were given without the statement of time limitation. The researcher stated in addition to the standard Cycle III instructions, "When you have completed your drawing of a person, please turn the page and read the directions at the top. If there are any questions about the second drawing, you may ask me." This procedure is recommended by Harris in his directions for older children (see reference two, page 241). The researcher also announced to the class at the appropriate time, "The period is half over." The statement about the class period was made for two reasons. First, most groups were tested during a school period of 40 to 45 minutes. Where class periods were shorter, additional time was requested of the school administration so that all untimed groups had 40 to 45 minutes for two drawings. Second, in pilot work it was found that many adolescents would finish their first drawing and then return to it with a flurry of erasing and redrawing rather than going on to the second drawing. By stating that half the period was over, almost all of the students would turn over the test form and begin the drawing of the self. As each student completed the two drawings the researcher moved about the room collecting papers. Those who

finished early were requested to turn to their own assigned work.

Under both the timed and untimed conditions the researcher moved about the periphery of the room and through alternate rows of desks. Her manner was interested but not evaluative, and she made an attempt to be neutral as specified in the Cycle III instructions and yet to be a motivating stimulus as suggested by Harris in his directions for older children (see reference two, page 241).

In older groups, above fifth or sixth grade, it may be necessary to offer strong encour-

agement to some children, who will say they can't do the task.

. . .

In this case, it is well to have two examiners who can walk about the room speaking to individuals who seem reluctant to attempt the task.

Before giving the instructions the researcher asked the subjects to specify the following information: Name (not used after correct birthdate and occupation of father were established), sex, grade, date of test, date of birth, father's occu-

Table 4. Description of communities where testing was done

Community	Classification	Location	Brief description
Bellefonte, Pa.	Small town	Mountainous region of Central Pennsylvania	An old, small community with some heavy and light industry, as well as farming. Employment is heavily operative.
Bristol, Pa.	Suburban	North of Philadelphia	Heavy industrial area with high percentage of ethnic groups. Employment is mostly craftsmen, operatives, sales, and service.
Dalton, Ohio	Village	Rich, northern Ohio farmlands (Wayne Co.)	An old, small village in an agricultural area. Other occupations, primarily crafts, commute to nearby industrial cities.
Harrisburg, Pa.	City	East central Pennsylvania	The State capital with a broad spectrum of occupations at all levels. The schools sampled represented largely the operative and labor groups.
Hollidaysburg, Pa.	Suburban	West of Altoona	A neat, flourishing community with a high representation of skilled and operative workers.
Huntingdon, Pa.	Small town	Mountainous region of central Pennsylvania	An old, somewhat depressed community represented primarily by working classes—sales, personal services, operatives, and labor.
Lemoyne, Pa.	Suburban	East of Harrisburg	Three communities served by Lemoyne Area School District encompass a cross section of occupations. Employment is largely managerial, clerical, and sales.
Camp Hill, Pa.	Suburban	East of Harrisburg	
New Cumberland, Pa.	Suburban	East of Harrisburg	
Marple, Pa.	Suburban	Southwest of Philadelphia	Two communities served by Marple-Newtown School District. They are primarily, well-to-do suburbs with great representation of professional, managerial, and clerical occupations.
Newtown, Pa.	Suburban	Southwest of Philadelphia	
Mt. Lebanon, Pa.	Suburban	Southwest of Pittsburgh	An old and wealthy suburb with a high representation of professional people.
Penns Hills, Pa.	Suburban	East of Pittsburgh	Developing community with a high proportion of sales and skilled crafts personnel.
Tyrone, Pa.	Small town	Central Pennsylvania	Primarily a mill town. With the major occupations centered around a paper mill, most workers are operatives.

pation, and school. With regard to the father's occupation it was necessary for the researcher to give examples of occupations and to say specifically, "not *where* he works, but *what* he does." Then the instructions were given, and routinely the students asked the following questions: Should it be front or profile? Should it be a man or woman? Does it have to be someone real? Do I have to draw what I'm wearing today? In response to such questions the researcher made a neutral response as specified in Cycle III directions such as "Use your own judgment" or "Make it anyway you wish." The only restriction was on the use of photographs for copying the self. In this case the researcher would suggest, "No one should be copying himself from a picture."

Other interactions between researcher and students followed those specified by Harris in the above quote. Some adolescents said they simply could not draw. In that case the researcher would simply ask them to do their best to try, adding that the test was not one of drawing skill as they would find out in an explanation to be given after the test. Only two of the 1,775 students tested absolutely refused to take part when the test was introduced to the class. They were permitted to work quietly in the rear of the classroom.

Routinely the 16-year-olds were most interested in what the test was for and why it was being given. There was less interest among the 14-year-olds and less still among the 12-year-olds. In response to any inquiries about the purpose of the test, the researcher assured the students that all their questions would be answered when they were finished. As mentioned above, the promise of a discussion session was used as motivational material for the more reluctant students. After administration, a short and humorous presentation about the testing was given to each class, and an opportunity for questioning was provided. The students generally showed a keen interest in the entire testing process.

From the group of subjects tested according to the foregoing procedures, 200 students were chosen at each of three age levels, 12, 14, and 16 years, so controlled that the mean for each age sample fell close to the half-year interval (i.e., 12 years, 6 months; 14 years, 6 months; 16 years, 6 months). At each grade level 100 students had taken the test under the timed condition;

the other 100 students, under the untimed condition. In each group of 100, half of the subjects were male and the other half, female; the racial distribution was 90 percent white and 10 percent nonwhite, to correspond with proportions reported in the 1960 U.S. Census report for the Nation. Each age sample of 100 students was also carefully controlled by father's occupation, to represent the national distribution of occupations as reported in the 1960 census.

The rationale for the procedure of proportional selection from the several occupational strata is that socioeconomic status is a major variable, influencing ability test results more than geography. Such a quota sample would be more comparable to a truly random national sample than would a sample of convenience.

The argument that a man's occupation is a reasonable index for social stratification has been covered in many sociological works. For example, Slocum⁴ says:

In contemporary America, the type of work a man does, together with the reputation he establishes at work, must be regarded among the principal determinants of social rank.

Warner⁵ found correlation of 0.88 between occupation and judged social class. It is, of course, generally recognized in psychological and sociological research that ability test data are substantially correlated with social status, and this fact is generally taken into account in development of norms for tests. The norms of the Goodenough-Harris Drawing Test were established according to a stratified occupational level sampling procedure.

In the 1960 U.S. Census report on occupation by industry,⁶ all occupations are divided into nine categories. These categories, developed by Alba Edwards, represent the entire work force in the United States. (For a more complete discussion of Edwards' work see references four and five.) The categories provide a basis for grouping occupations and setting up representative sampling, although they do not necessarily constitute a rank order of occupations by skill or ability.

In the present study, Negro children were distributed throughout the samples according to proportions reported in the 1960 Census. The samples thus drawn are reported in detail in tables 5 and 6. In these tables, the "ideal pro-

Table 5. Distribution¹ of sample taking drawing test under timed conditions, by age, sex, and occupation of father

Occupation of father	Ideal proportion	12 years			14 years			16 years		
		Total	Boys	Girls	Total	Boys	Girls	Total	Boys	Girls
Total----	100.0(8.0)	98(7)	49(4)	49(3)	100(10)	47(5)	53(5)	101(10)	49(2)	52(8)
Professional---	10.0(0.5)	10	5	5	11 (1)	5	6(1)	10	5	5
Farm-----	4.5(0.5)	5	2	3	5	2	3	5	3	2
Managerial----	10.0 (-)	11(1)	6(1)	5	10	5	5	10	5	5
Clerk, sales---	13.0(0.5)	14(1)	7(1)	7	14 (1)	6	8(1)	14 (1)	7	7(1)
Craftsmen, foremen-----	19.0(1.0)	20(1)	10	9(1)	20 (1)	10(1)	10	20 (1)	10	10(1)
Operatives-----	18.0(2.0)	20(2)	10(1)	10(1)	20 (2)	11(2)	9	20 (2)	10(1)	10(1)
Service-----	5.0(1.0)	6(1)	2	4(1)	6 (1)	2	4(1)	6 (1)	3	3(1)
Labor-----	7.0(2.5)	8(1)	5(1)	3	10 (3)	4(1)	6(2)	10 (3)	4(1)	6(2)
Unemployed, or not reported--	3.5 (-)	4	2	2	5 (1)	2(1)	2	6 (2)	2	4(2)

¹Figures in parentheses represent the number of Negro children contained in totals for both the ideal proportion and realized samples.

Table 6. Distribution¹ of sample taking drawing test under untimed conditions, by age, sex, and occupation of father

Occupation of father	Ideal proportion	12 years			14 years			16 years		
		Total	Boys	Girls	Total	Boys	Girls	Total	Boys	Girls
Total----	100.0(8.0)	100(10)	51(5)	49(5)	101(10)	49(2)	52(8)	98(7)	50(4)	48(3)
Professional---	10.0(0.5)	11 (1)	6(1)	5	11 (1)	5	6(1)	10	5	5
Farm-----	4.5(0.5)	4	2	2	5	3	2	5	3	2
Managerial----	10.0 (-)	10	5	5	10	5	5	10	5	5
Clerk, sales---	13.0(0.5)	14 (1)	7	7(1)	14 (1)	8(1)	6	14(1)	7(1)	7
Craftsmen, foremen-----	19.0(1.0)	19	10	9	20 (1)	9	11(1)	20(1)	11(1)	9
Operatives-----	18.0(2.0)	20 (2)	11(2)	9	20 (2)	9	11(2)	20(2)	10(1)	10(1)
Service-----	5.0(1.0)	6 (1)	4(1)	2	7 (2)	3	4(2)	6(1)	3(1)	3
Labor-----	7.0(2.5)	10 (3)	4(1)	6(2)	10 (3)	5(1)	5(2)	9(2)	4	5(2)
Unemployed, or not reported--	3.5 (-)	6 (2)	2	4(2)	4	2	2	4	2	2

¹Figures in parentheses represent the number of Negro children contained in totals for both the ideal proportion and realized samples.

portion" is that reported in the Census. The proportion actually achieved for this study matches the expected or target figure quite closely; proportions requiring fractions of a person, arbitrarily distributed between boys and girls and between Negro and white students under the limitations of the available subjects may readily be noted in these tables.

Tables 7, 8, and 9 show how the communities from which the subjects were drawn contribute to the occupational strata of the sample. Note that the actual sample selection is generally in keeping with the community description from table 4. An exception is notable in the community of Huntingdon, Pennsylvania, where a relatively small, working class community houses Juniata College, a fact which causes a number of professional occupations to appear in this predominantly blue-collar community.

The drawing protocols of all students in the sample were scored independently by two well-trained scorers and differences reconciled through conference, and the data were subjected to appropriate analysis to assess the effects of timed versus untimed administration.

Results

The mean raw scores were calculated for tests done under both conditions, timed and untimed, for the three age groups. Tables 10 and 11 report relevant statistics. Note that for the drawing of a person, the mean of untimed scores is higher than the mean of timed scores for each age group. However, by comparison, in the drawing of the self the mean of the untimed score is higher only for 14- and 16-year-old students. Twelve-year-old students yield essentially the same scores on the drawing of the self regardless of condition.

A t for independent samples was calculated for each age group for the mean scores on drawings of a person and the self to test the difference between the timed and untimed conditions. Note in table 10 that for the drawing of the person the untimed score was significantly higher

in two cases, i.e., for 14- and 16-year-olds ($t = 3.38$ and 3.42 , respectively). Because $t_{.05}$ for 200 degrees of freedom is 1.96, the null hypothesis was rejected for ages 14 and 16, but retained for age 12 ($t = 1.51$). In testing treatment differences with the self drawing (table 11) it was found that although the mean of untimed scores was higher for 14- and 16-year-olds, it was not significantly higher than the mean of timed scores ($t = 0.90$ and 1.72 , respectively). No significant difference at the .05 level was found for 12-year-olds ($t = 0.08$).

Up to this point, only the data obtained in this study have been considered. It is possible, and desirable, to compare these results with the published norms of the Goodenough-Harris Drawing Test. These comparisons, however, require certain explanations and justification.

The Goodenough-Harris method of evaluating drawings of a man, a woman, and the self was developed to provide a reliable and valid estimate of intellectual level up to and including the early years of adolescence. The test serves this function reasonably well, and it is this function only which is considered in the present report. However, the Health Examination Survey's instructions requested the drawing of a *person*, and while most subjects elected to draw a person of his or her own sex, a small group in each sex drew human figures of opposite sex.^d These two small subgroups, i.e., boys drawing a woman and girls drawing a man, necessitate two additional comparisons for each age group with the appropriate Goodenough-Harris norms. Both of these subgroups will have less stable mean scores and standard deviations than the much larger groups of students drawing figures of the same sex.

^dClinical psychologists utilize this fact (that some subjects draw opposite-sex figures), in individual cases, to draw certain psychodynamic inferences concerning the personality of the youthful artist. There is, according to Harris,² however, no evidence that any systematic relationship exists between personality characteristics of subjects and their intellectual ability as estimated from the same drawings.

Table 7. Distribution of sample for age 12, by occupation of father and community

Community	Occupation of father								
	Pro- fes- sional	Farm	Mana- gerial	Clerk, sales	Crafts- men, foremen	Oper- atives	Serv- ice	Labor	Unem- ployed, or not reported
	Number								
Bellefonte, Pa-	-	-	-	-	1	5	-	2	-
Bristol, Pa----	-	-	1	-	9	2	3	1	2
Dalton, Ohio---	-	5	-	-	3	3	-	1	1
Harrisburg, Pa-	-	-	-	2	-	4	1	5	2
Hollidaysburg, Pa-----	-	-	-	-	1	3	1	4	1
Huntingdon, Pa-	3	3	1	-	5	7	4	1	1
Lemoyne, Pa----	1	1	1	5	5	6	1	1	-
Marple-Newtown, Pa-----	-	-	4	3	4	-	-	-	1
Mt. Lebanon, Pa-	13	-	10	9	1	-	-	-	1
Penns Hills, Pa-	4	-	4	9	11	2	2	1	1
Tyrone, Pa-----	-	-	-	-	-	7	-	2	-

Table 8. Distribution of sample for age 14, by occupation of father and community

Community	Occupation of father								
	Pro- fes- sional	Farm	Mana- gerial	Clerk, sales	Crafts- men, foremen	Oper- atives	Serv- ice	Labor	Unem- ployed, or not reported
	Number								
Bellefonte, Pa-	-	2	-	-	4	2	-	-	-
Bristol, Pa----	1	-	1	5	4	6	2	3	-
Dalton, Ohio---	1	6	-	1	2	6	-	1	-
Harrisburg, Pa-	1	-	-	1	1	3	2	6	1
Hollidaysburg, Pa-----	-	-	-	2	7	6	-	-	-
Huntingdon, Pa-	2	-	1	3	4	5	4	4	1
Lemoyne, Pa----	1	-	7	4	4	4	3	1	2
Marple-Newtown, Pa-----	7	-	3	4	2	1	1	-	4
Mt. Lebanon, Pa-	10	-	7	6	1	-	-	-	1
Penns Hills, Pa-	-	-	1	2	9	5	-	4	-
Tyrone, Pa-----	-	2	-	-	2	2	1	-	-

Table 9. Distribution of sample for age 16, by occupation of father and community

Community	Occupation of father									
	Pro- fes- sional	Farm	Mana- gerial	Clerk, sales	Crafts- men, foremen	Oper- atives	Serv- ice	Labor	Unem- ployed, or not reported	
	Number									
Bellefonte, Pa-	2	3	-	2	4	5	3	-	1	
Bristol, Pa----	-	-	1	2	6	4	1	-	1	
Dalton, Ohio----	-	4	-	-	1	4	-	-	-	
Harrisburg, Pa-	-	-	-	1	2	3	2	3	1	
Hollidaysburg, Pa-----	-	-	-	1	1	2	-	1	-	
Huntingdon, Pa-	5	1	1	6	5	3	1	4	3	
Lemoyne, Pa----	-	1	2	2	7	7	3	5	2	
Marple-Newtown, Pa-----	3	1	7	3	2	1	2	-	1	
Mt. Lebanon, Pa-	9	-	5	3	2	1	-	-	-	
Penns Hills, Pa-	1	-	4	7	11	6	-	2	2	
Tyrone, Pa-----	-	-	-	2	-	4	-	2	1	

Table 10. Comparison of timed and untimed results of the drawing of a "person"

Age and test condition	N	\bar{X}	Standard error of mean	t
<u>12 years</u>				
Timed-----	98	41.4	0.87	} 1.51
Untimed----	100	43.2	0.80	
<u>14 years</u>				
Timed-----	100	41.8	0.81	} 13.38
Untimed----	101	45.5	0.74	
<u>16 years</u>				
Timed-----	101	43.1	0.98	} 13.42
Untimed----	98	47.6	0.89	

¹p < .05

NOTE: N= number of subjects; \bar{X} = mean raw score.

To compare the results of person and self drawings obtained under timed and untimed conditions with the Goodenough-Harris norms for adolescents, the mean raw score and standard deviation were calculated for each type of drawing (person, either man or woman, and self) for each sex and age. Tables 12 and 13 present these comparisons.

Norms for the Goodenough-Harris test are available only through age 15. Therefore, for the sake of comparison the available norms for 15-year-olds have been compared with the results for 16-year-olds which were obtained in the present study. This practice seems justified as the test shows increasingly smaller mean age increments in the early adolescent years, and by age 15, increases from year to year are quite small.

Because the group tested under the untimed condition received instructions analogous to those of the Goodenough-Harris test (with the primary exception that the sex of the drawing was not

Table 11. Comparison of timed and untimed results of the drawing of the "self"

Age and test condition	N	\bar{X}	Standard error of mean	t
<u>12 years</u>				
Timed-----	98	40.7	0.87	} 0.08
Untimed----	100	40.6	0.90	
<u>14 years</u>				
Timed-----	100	42.1	0.80	} 0.90
Untimed----	101	43.2	0.82	
<u>16 years</u>				
Timed-----	101	41.9	0.93	} 1.72
Untimed----	98	44.5	1.21	

NOTE: N= number of subjects; \bar{X} = mean raw score.

specified), it was expected that scores would equal or nearly equal the norms. Indeed, the results of this study compare very favorably with national norms reported (tables 12 and 13). In only three of 18 possible comparisons does the discrepancy of mean scores exceed three score points, which is the standard error of measurement for the test at the ages concerned in this study. That the means of subjects tested in one geographic region compare so favorably with national norms may be taken as justification for the attention given to constructing the quota sample, in order to assure this result.

The data from the group tested under the timed conditions in this study provide, as should be expected, mean scores less in keeping with published norms. In nine of 18 comparisons in tables 12 and 13 the discrepancy exceeds three points, the standard error of measurement for the test, and in each instance the mean value of

Table 12. Comparison of timed and untimed results with the Goodenough-Harris norms for boys

Age and drawing	Norms ¹		Untimed			Timed			t
	\bar{X}	S.D.	N	\bar{X}	S.D.	N	\bar{X}	S.D.	
<u>12 years</u>									
Man-----	40.3	11.01	47	41.0	7.87	42	38.8	7.69	1.31
Woman-----	39.8	9.61	4	46.2	7.41	7	39.0	6.19	1.65
Self-----	40.3	11.01	51	39.4	10.21	49	37.4	7.92	1.14
<u>14 years</u>									
Man-----	44.7	10.51	46	45.4	7.42	44	41.0	8.86	² 2.54
Woman-----	44.1	9.41	3	44.0	9.54	3	36.7	8.26	1.01
Self-----	44.7	10.51	49	42.6	8.58	47	40.2	8.51	1.39
<u>16 years</u>									
Man-----	45.1	10.60	40	47.0	8.98	43	44.8	8.35	1.11
Woman-----	44.4	9.31	10	43.5	9.32	6	36.0	13.57	1.20
Self-----	45.1	10.60	50	42.9	14.01	49	41.9	9.06	0.43

¹Norms are from Harris (1963), tables 11 and 13.

²p < .05

NOTE: \bar{X} = mean raw score; S.D.= standard deviation; N= number of subjects.

Table 13. Comparison of timed and untimed results with the Goodenough-Harris norms for girls

Age and drawing	Norms ¹		Untimed			Timed			t
	\bar{X}	S.D.	N	\bar{X}	S.D.	N	\bar{X}	S.D.	
<u>12 years</u>									
Man-----	43.0	10.32	8	40.9	7.08	7	40.1	5.08	0.23
Woman-----	45.8	9.58	41	46.0	7.72	42	44.7	9.41	0.68
Self-----	45.8	9.58	49	41.7	7.52	49	44.0	8.37	-1.40
<u>14 years</u>									
Man-----	45.1	9.57	13	43.9	9.23	8	42.5	7.72	0.38
Woman-----	48.2	8.97	39	46.3	6.94	45	42.8	7.55	² 2.21
Self-----	48.2	8.97	52	43.7	8.00	53	43.9	7.67	0.10
<u>16 years</u>									
Man-----	45.2	9.01	15	48.3	7.83	13	40.3	11.35	² 2.13
Woman-----	48.2	8.48	33	49.4	8.72	39	43.2	9.89	² 2.81
Self-----	48.2	8.48	48	46.1	9.39	52	41.8	9.67	² 2.25

¹Norms are from Harris (1963), tables 11 and 13.

²p < .05

NOTE: \bar{X} = mean raw score; S.D. = standard deviation; N = number of subjects.

the sample is below the norm with which it is appropriately compared. Thus the conclusion that the test, when given under conditions of sharply limited time, requires restandardization for proper interpretation, is emphasized.

One final observation may be made. The investigator, as the primary agent in the 1963 restandardization of the Goodenough Draw-a-Man Test, can only regard with satisfaction the generally favorable comparison of the results of this study with those of his original work. While conditions of restricted time (a condition often experienced in clinical work and research) may attenuate the results of this test, they apparently do not constrict them sufficiently to invalidate the technique. The mean scores are still high; in only two of 18 comparisons are the discrepancies larger than six points, twice the standard error of measurement. Hence, the use of the test under conditions of reduced time is clearly feasible, and the establishment of new norms, a distinct possibility. This restandardization for the adolescent years could, of course, be based on data

collected in Cycle III of the Health Examination Survey. Such new norms could also be confirmed statistically, through regression equations, by a study applying the test under the two conditions to the same group, appropriately stratified and tested in counterbalanced order.

SUMMARY

A group of 102 students were tested with the Goodenough-Harris test in a counterbalanced design to assess the effects of group and individual presentation of the test. The effect of the treatment was not significant at the .05 level, *t* for group I (individual-group) was 1.79 and *t* for group II (group-individual) was 0.86. In addition, there was no interaction effect between treatment and order of presentation (*t*_{.05} = 0.41).

Since there was no difference in method of presentation, samples of high-school students at ages 12, 14, and 16 were tested in classroom groups. Two hundred students were selected in each age group, half tested under timed conditions

and half under untimed conditions. A significant difference at the .05 level was found for the drawing of a person by 14- and 16-year-olds ($t = 3.38$ and 3.42 , respectively), with the untimed tests yielding higher mean scores. No significant

difference was obtained for the person drawn by 12-year-olds or for the drawing of the self at any age. A comparison of these data with published norms is reported. Observations and results are discussed.

REFERENCES

¹National 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.

²Harris, D. B.: *Children's Drawings as Measures of Intellectual Maturity*. New York. Harcourt, Brace, and World, Inc., 1963.

³Scott, W. A., and Wertheimer, M.: *Introduction to Psychological Research*. New York. John Wiley and Sons, Inc., 1962. pp. 258-262.

⁴Slocum, W. L.: *Occupational Careers*. Chicago. Aldine Publishing Co., 1966. pp. 76-87.

⁵Warner, W. L.: *Social Class in America*. Chicago. Science Research Associates, 1949. pp. 131-159.

⁶U.S. Bureau of the Census: *U.S. Census of Population, 1960, Occupation by Industry, United States Summary*. Subject Report PC(2)-7C. Washington. U.S. Government Printing Office, 1961.



APPENDIX I
TEST INSTRUCTIONS

I. *Original Goodenough-Harris Test Instructions*

(Based on doing three pictures—the man, the woman, and the self)

I want you to make a picture of (a man, a woman, yourself). Make the very best picture that you can; take your time and work very carefully. Be sure to make the whole (man, woman, self), not just the head and shoulders.

II. *Cycle III Goodenough-Harris Test Instructions*

(Based on doing two pictures—the person and the self)

I want you to draw a picture of (a person, yourself). Make the very best picture you can. Be sure to make the whole (person, self), not just the head and shoulders. You will have five minutes to draw (a person, yourself). Work very carefully.



APPENDIX II. TEST FORMS

HES Cycle III

**Draw a picture of a person. Make the very best picture you can.
Be sure to make the whole person, not just the head and shoulders.**

M F

- 1. _____ 41. _____
- 2. _____ 42. _____
- 3. _____ 43. _____
- 4. _____ 44. _____
- 5. _____ 45. _____

- 6. _____ 46. _____
- 7. _____ 47. _____
- 8. _____ 48. _____
- 9. _____ 49. _____
- 10. _____ 50. _____

- 11. _____ 51. _____
- 12. _____ 52. _____
- 13. _____ 53. _____
- 14. _____ 54. _____
- 15. _____ 55. _____

- 16. _____ 56. _____
- 17. _____ 57. _____
- 18. _____ 58. _____
- 19. _____ 59. _____
- 20. _____ 60. _____

- 21. _____ 61. _____
- 22. _____ 62. _____
- 23. _____ 63. _____
- 24. _____ 64. _____
- 25. _____ 65. _____

- 26. _____ 66. _____
- 27. _____ 67. _____
- 28. _____ 68. _____
- 29. _____ 69. _____
- 30. _____ 70. _____

- 31. _____ 71. _____
- 32. _____ 72. _____
- 33. _____ 73. _____
- 34. _____
- 35. _____

R.S. _____

36. _____ **Scoring Problems**

37. _____ Yes No

38. _____ **If Yes, List:**

39. _____

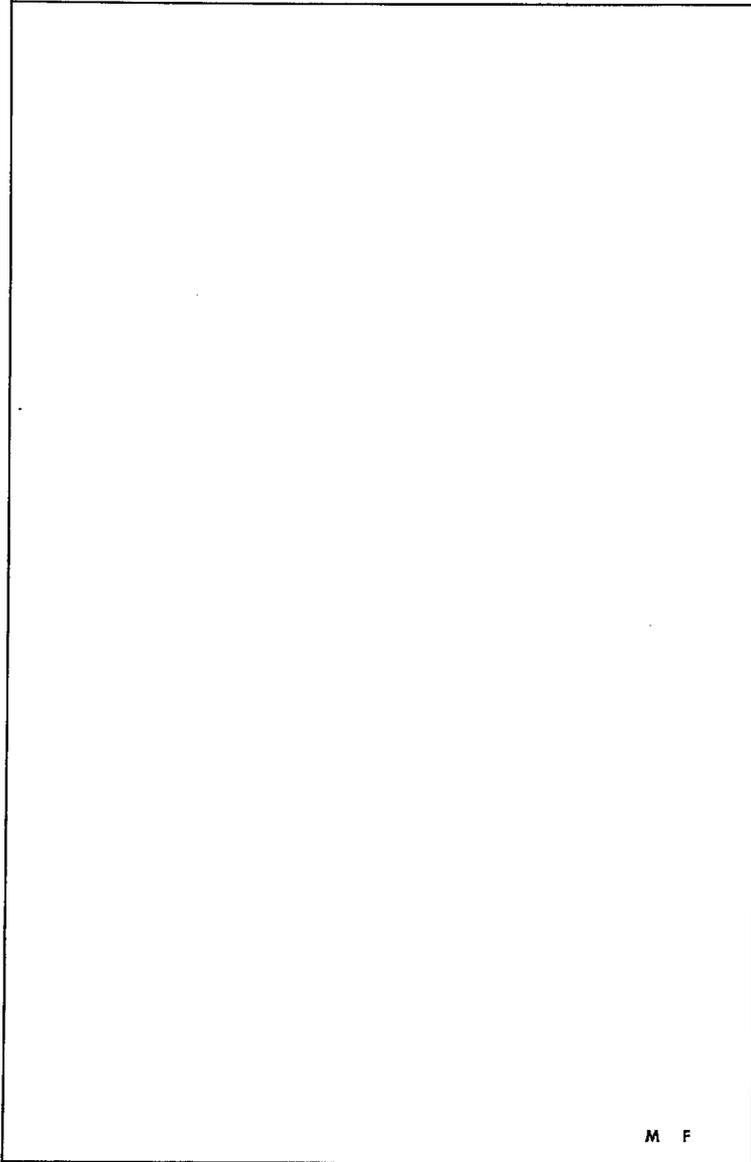
40. _____

Name _____ M F Scorer _____ Sample No. _____

Examiner _____ Test Quality Valid Invalid Not Administered

Copyright © 1968 by Harcourt, Brace & World, Inc., New York. All rights reserved. Printed in U.S.A.

Draw a picture of yourself. Make the very best picture you can.
 Be sure to make your whole self, not just your head and shoulders.

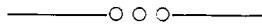


- 1. _____ 41. _____
- 2. _____ 42. _____
- 3. _____ 43. _____
- 4. _____ 44. _____
- 5. _____ 45. _____
- 6. _____ 46. _____
- 7. _____ 47. _____
- 8. _____ 48. _____
- 9. _____ 49. _____
- 10. _____ 50. _____
- 11. _____ 51. _____
- 12. _____ 52. _____
- 13. _____ 53. _____
- 14. _____ 54. _____
- 15. _____ 55. _____
- 16. _____ 56. _____
- 17. _____ 57. _____
- 18. _____ 58. _____
- 19. _____ 59. _____
- 20. _____ 60. _____
- 21. _____ 61. _____
- 22. _____ 62. _____
- 23. _____ 63. _____
- 24. _____ 64. _____
- 25. _____ 65. _____
- 26. _____ 66. _____
- 27. _____ 67. _____
- 28. _____ 68. _____
- 29. _____ 69. _____
- 30. _____ 70. _____
- 31. _____ 71. _____
- 32. _____
- 33. _____
- 34. _____
- 35. _____
- 36. _____ R.S. _____
- 37. _____ Scoring Problems
- 38. _____ Yes No
- 39. _____ If Yes, List: _____
- 40. _____

M F

Name _____ M F Scorer _____ Sample No. _____

Examiner _____ Test Quality Valid Invalid Not Administered



OUTLINE OF REPORT SERIES FOR VITAL AND HEALTH STATISTICS

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 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 on 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 manpower occupations, hospitals, nursing homes, and outpatient and other inpatient facilities.
- Series 20. Data on mortality.*—Various statistics on mortality other than as included in 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 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, medical experience in the last year of life, characteristics of pregnancy, etc.

For a list of titles of reports published in these series, write to: Office of Information
National Center for Health Statistics
U.S. Public Health Service
Washington, D.C. 20201

