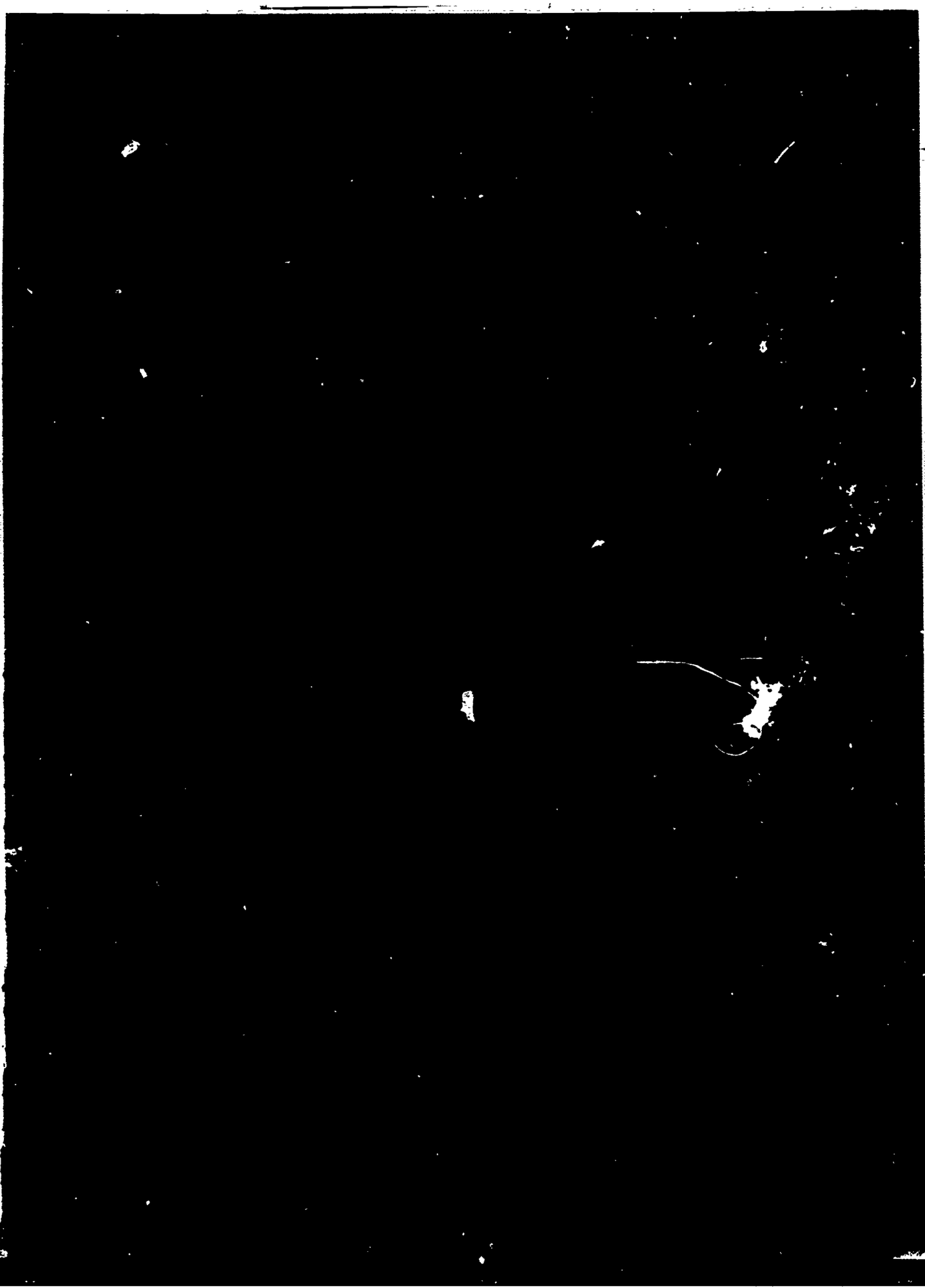
DOCUMENT RESUME

ED 080 600 TM 003 117

TITLE	Maine Assessment of Educational Progress, Report 5: Methodology; Citizenship and Writing, 1972.
INSTITUTION	Maine State Dept. of Educational and Cultural Services, Augusta.; Research Consortium for Educational Assessment.
REPORT NO	R-5
PUB DATE	72
NOTE	60p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Citizenship; Data Collection; *Educational Research; *Educational Status Comparison; *Measurement Instruments; Questionnaires; *Research Methodology; Secondary Grades; State Programs; Technical Reports; *Test Construction; Writing Skills
IDENTIFIERS	*Maine Assessment of Educational Progress

ABSTRACT

The methodology used in the Maine Assessment of Educational Progress in Citizenship and Writing is described in this report. The chapters of the report are: I. Introduction--background and purpose of the project, sampling considerations, Exercise Package development, administration and scoring, and data analysis plan; II. Sample Design--general considerations, sample frame construction, stratification of the sampling frame, sample allocation, school sample selection, pupil sample selection; III. The Exercise Package--development, citizenship exercises, writing exercises, student questionnaire, package assembly; IV. Supplementary Information--principal's questionnaire, department records, summary of data collected; and V. Administration and Scoring--time frame of Maine Assessment Project, summary of field operations, and exercise scoring. (For related documents, see TM 003 115, 116.) (DB)



MAINE ASSESSMENT OF EDUCATIONAL PROGRESS

REPORT 5

M E T H O D O L O G Y

Citizenship and Writing

1972

Carroll R. McGary, Commissioner
Department of Educational and Cultural Services
Augusta, Maine

Conducted in cooperation with Research Consortium
for Educational Assessment consisting of
Research Triangle Institute, Research Triangle Park, N.C.
Measurement Research Center, Iowa City, Iowa
American Institute for Research, Palo Alto, California

The preparation of this report was financially assisted by the State of Maine, Department of Educational and Cultural Services, Appropriation Account Number 4635; through a Federal grant from the U.S. Department of Health, Education and Welfare.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

ACKNOWLEDGEMENTS

The Maine Assessment of Educational Progress (MAEP) for 1972 was a joint effort by the Maine Department of Educational and Cultural Services and the Research Consortium for Educational Assessment. The Consortium consists of the Research Triangle Institute (RTI), the American Institutes for Research (AIR), and the Measurement Research Center (MRC) division of the Westinghouse Learning Corporation.

MAEP involved and required the efforts and cooperation of many people. The following personnel at RTI in the Center for Educational Research and Evaluation comprised the project staff: Dr. D. L. Bayless, Project Director; Dr. J. P. Bailey, Jr., Assistant Project Director; Dr. G. H. Dunteman and Ms. H. Lewis.

The Statistics Research Division at RTI was particularly helpful in a multitude of ways in survey operations and data processing. Personnel involved were Dr. B. V. Shah, Mr. J. C. Wright, Mr. D. W. Jackson, Ms. N. Monroe, Ms. N. Hunt, and Ms. L. Bressler.

Dr. V. Campbell and Dr. C. Finley at AIR contributed heavily to the instrument development and analyses efforts of the project, and Dr. L. Diana, Mr. J. O'Neil, and Ms. B. Dixon at MRC performed substantial work in the data handling and scoring efforts.

Maine Department personnel who contributed substantially to the success and completion of MAEP included: Dr. C. R. McGary, Commissioner of Education; Mr. A. Natale, Mr. A. Gordon, Dr. H. Maxcy, Mr. E. Hinckley, Ms. M. Reynolds, Dr. J. Pecoraro and Ms. S. Ezzy.

Finally, no report can be written successfully without detailed editorial and secretarial assistance. Particular thanks goes to Ms. F. Heald and Ms. L. Froyen in the Center for Educational Research and Evaluation at RTI, and Ms. A. Hanley of the Maine Department.

TABLES AND FIGURES

I. INTRODUCTION

Background and Purpose of the Project

In cooperation with the Department of Educational and Cultural Services of the State of Maine, the Research Consortium for Educational Assessment designed and implemented in 1972 the Maine Assessment of Educational Progress. The Consortium consists of three institutions which have been deeply involved in assessment for years: Research Triangle Institute (RTI) of North Carolina, Measurement Research Center (MRC) of Iowa City, Iowa, and American Institutes for Research (AIR) in Palo Alto, California.

The purpose of the Maine Assessment of Educational Progress (MAEP) in 1972 was to complete the first step or phase of a ten-year comprehensive needs assessment program of the in-school students of the state. The overall model is designed to provide specific information about knowledge, skills, understandings, and attitudes in subject matter areas. The current phase investigated the areas of Citizenship and Writing using the National Assessment of Educational Progress (NAEP) model; and subsequent yearly assessments will assess eight other areas.

National Assessment (NAEP) is a census-like study to measure important specific outcomes of education. The beginnings of NAEP date from about 1964 when the Exploratory Committee on Assessing the Progress of Education (ECAPE) was established with Carnegie Corporation funds to investigate the possibilities of a national assessment. ECAPE developed a plan and instrumentation for its work, and this in turn resulted in the overall design now followed by NAEP. Actual administration of National Assessment exercises in the schools began in the spring of 1969. NAEP's governing organization is the Education Commission of the States (ECS) composed of representatives of the states and territories, and NAEP is legally responsible to the public. Major funding is provided by the U. S. Office of Education.

The major goals of NAEP are twofold:

1. To make available the first census-like data on the educational attainments of young Americans; and

2. To measure any growth or decline which takes place in selected aspects of educational attainments of young Americans in certain subject areas.

Attainment of these goals should enable data to be gathered which will help answer the question, "How much good is the expenditure of so much money doing in terms of what Americans know and can do?" However, it is first necessary to determine what the educational system is trying to achieve. ECAPE, followed by CAPE (Committee on Assessing the Progress of Education) and NAEP, have long been involved in determining these objectives and developing questions and tasks (called exercises) to assess how well these objectives are being achieved. Exercises have been administered to thousands of people in four different age groups (9, 13, 17 and young adults) selected through random sampling procedures throughout the country. Citizenship and Writing, developed quite early in NAEP's history, are but two of ten areas for which objectives and exercises are being developed.

The Maine Department of Educational and Cultural Services, in an effort to gather statewide educational data, has adopted the basic elements of the NAEP model and adapted them to meet Maine needs. One would have difficulty coming up with direct evidence on the strengths and weaknesses of Maine schools in meeting the needs of our society as those needs are presently expressed in the objectives set by the schools themselves. To what extent are Maine students learning to read about and understand the scientific aspects of our society, or United States social structure, and thus developing into thoughtful citizens? Neither these questions nor many like them can be answered by information currently available in Maine.

These concerns lead to the question of Needs Assessment. The concept of educational needs assessment requires identifying learning outcomes which are desirable, and then determining a learner's status with respect to those outcomes. Put another way, one must first say what is important

for children to know as a result of their school experiences, and then systematically determine if they, in fact, do know what has been said is important for them to know.

After thorough investigation, careful thought, and with due consideration for the pressing urgency of the problem, the Educational and Cultural Services Department of Maine decided to embrace the NAEP model of objectives and exercises over a cyclical schedule of periodic assessments. Maine Assessment of Educational Progress (MAEP) will examine in Cycle I the ten subject matter areas of NAEP by scheduling two of them each year for five years. Citizenship and Writing were the first. The remainder are Science, Reading, Career and Occupational Development, Literature, Mathematics, Music, Art, and Social Studies. Cycle I will provide benchmark data. Cycle II will repeat Cycle I and will provide a measure of educational performance progress over time.

It was extremely difficult for the educational leaders in Maine to anticipate a set of educational issues uniquely relevant to Maine which they could expect the assessment program to answer directly. Some of the more general questions it was hoped data would shed some light on were:

1. Does the amount of money expended make a difference in student achievement in schools?
2. Does equal funding assure equal achievement?
3. Is student achievement related to teacher salaries?
4. Does community involvement and support affect student achievement?
5. Does the adequacy and amount of reading material in the home affect student achievement in school?
6. Are the needs of children being met by the school program?

These general questions, and many more like them, will underlie the critical examination of the achievement data produced by this project. Using survey techniques, public hearings, and other appropriate self-examination procedures both inside and outside the Maine Department, the criticality of

determined needs as revealed by the data will be assessed. What can be changed, improved, or implemented to meet the most critical needs will be determined. Areas in which more data are needed will form a basis for planning the type of data to be collected in subsequent years of assessment.

Sampling Considerations

The major concern in designing a sample for a Statewide Educational Assessment is to have the sample design compatible with the overall objectives of the assessment program. The cost effectiveness of selecting a sample of pupils to analyze the educational achievements in Citizenship and Writing of groups of 17-year-olds attending public and private schools in Maine, as contrasted with evaluating all 17-year-olds, is easily justified from a cost versus statistical precision viewpoint.

Design of a sample for educational assessment means consideration of the following factors of importance:

1. All schools and students of the target group must be available to be sampled.

2. The way in which the schools of the sample frame are to be grouped or stratified, as well as the way the outcome variables will be grouped, are both important.

3. The sample selection procedure to be used to select schools and pupils must be on a probability basis.

4. School and pupil sampling sizes at each stage of sampling must be considered for subsequent stages.

5. Weighting and estimating procedures to be used in analysis of the resultant data depend heavily on the sample design parameters used.

It is of extreme importance that the sample design be closely interwoven with the instrument development, data collection, and analysis phases of the assessment. The instrument development phase specifies which variables are to be analyzed. The data collection phase specifies how the data are to be collected, while the analysis phase specifies how the data are to be analyzed and reported.

The requirement of the Maine sample from a data analysis viewpoint was that it provide sample sizes of certain subgroups or subpopulations of Maine 17-year-olds in school so that Writing and Citizenship outcomes could be analyzed for these subgroups. In particular, the sample was spread across four geographical regions of the state (North, East, West, South) so that results could be reported by region. The sample was also expected to produce sample sizes to provide results of reliable statistical precision for subgroups defined by community variables, in the home variables, and other non-school and school variables.

The NAEP project assesses four age groups, each chosen to provide information at meaningful periods in educational life. Age 9 marks the end of most students' primary education. Age 13 is the end of most elementary education. Age 17 is usually close to the end of secondary education and is the last time many children are in school. The 10-year span of 26-35 for the young adult category provides a large enough population from which to sample adults who have finished their formal education.

The Maine Department of Educational and Cultural Services chose the 17-year-old population of in-school youngsters for its first assessment phase using the NAEP model, and a target of 2,000 students in the sample was set to represent the approximately 17,000 in the state. This was considered large enough to give statistical precision to the subgroup results to be reported. Detailed sampling methodology and techniques are given as Section II of this report.

Exercise Package Development

The exercise package for the MAEP project consisted of a 32-page booklet printed specially for the assessment project by the Measurement Research Center. Because of the basic design of this project, an overriding consideration in the development of the package was ensuring its compatibility with materials and procedures used in the national effort. Other considerations included the usual ones of exercise format, placement and

location, mode of administration, and procedures to ensure standardization among testing situations and conditions.

After the basic decision had been made to conduct the MAEP project in the subject matter areas of Citizenship and Writing, the available released exercises from NAEP were carefully examined by Maine Department staff and by exercise developers from the American Institutes for Research to see if they reflected objectives selected by Maine, and to see if they could be modified where needed to be administered in group sessions using the paced-tape method, and still retain a high degree of comparability. The final decision was to include 23 Citizenship and 7 Writing exercises in the package, plus a 23-item Student Questionnaire to collect background and demographic data considered cogent to the project. Total testing time was less than two hours.

Exercise format was kept virtually identical to NAEP in all cases. The exercises themselves were carefully placed in a quasi-random order in the booklet to avoid problems of inter-item response set. A paced-tape was created for the booklet to minimize effects of slow reading ability among students. Each item was read aloud and ample time was permitted for students to answer before proceeding to the next page.

The issue of confidentiality of information was met by having student identification information, which was necessary to ensure that specified students came to the testing sessions, removed from exercise booklets before the booklets left each school and were sent away for scoring.

A more detailed discussion of the exercise booklet is given in Section III of this report, in terms of exercises and their objectives. A complete MAEP Exercise Booklet is included as Appendix B.

Administration and Scoring

The assessment exercise packages were administered in the first two weeks of May 1972 in 97 schools in the state to 1,749 students aged 17 then attending school. In the interests of having highly standardized testing

conditions and as short a lapse of time as possible between beginning and end of exercise administration, extensive training of exercise administrators was conducted by personnel from Research Triangle Institute.

Personnel employed regularly at the Maine Department in Augusta comprised the exercise administration teams for the project. An exercise manual was developed in detail. After training with the materials, tape recorder, and stimulus tapes, each administrator was provided a schedule of his administration dates and schools, and a complete set of exercise booklets for the schools in his assignment. Finally, check-in procedures were developed to establish a regulated system for editing and accounting for all information and work completed in the field.

Scoring was accomplished by the Measurement Research Center. Again, due to the desire to maintain comparability with the NAEP data, the same scoring criteria were used with the Maine project as were used with NAEP. Closed questions, i.e., those with a specific right or wrong answer or a yes/no, were simply scored appropriately. The many open-ended questions, however, required special care as professional scorers (some who also scored for NAEP) evaluated each student's answer and assigned numerical scores indicating the appropriate NAEP response category code. Thus answers were deemed "satisfactory" by the same criteria used previously when the exercises were administered nationwide in the NAEP project.

A data tape record was then developed for each student which included his responses to each of the exercises, information given by him in the student questionnaire section of the exercise booklet, and certain other information of a more general nature relating to his school or community. This information was sent to RTI for merging with student weight information (from the sample design).

A more detailed discussion of the administration and scoring procedures is given in Section V of this report, and the Exercise Administrator's Manual is included as Appendix D.

Data Analysis Plan

The basic elements of the analysis plan were decided jointly by staff from the Maine Department and personnel from RTI, AIR, and the Department of Utilization/Applications at NAEP. In general terms, comparisons were made of Maine results with appropriate NAEP data, and within-Maine reporting categories were developed as they became statistically feasible. The results were presented as percentages of students who gave the desired response on an exercise, just as with the NAEP project. Groups are compared, using this technique, simply by noting the difference in percentage who succeeded. It must be remembered, however, that a difference between groups does not tell what caused the difference.

Comparisons of results with the NAEP data were made on four comparable dimensions*: the national percentage, the Northeast U. S. percentage, percentages by sex, and percentages by four levels of parent education. These results for NAEP have been published and are readily available.

It was decided at the inception of the project that at least these same four reporting categories should be developed for the within-Maine analyses, plus certain others as feasible:

1. Demographic categories, such as region and size of community.
2. Family and home categories, such as bilinguality and family size.
3. Student-school categories, such as program and activities.
4. School system categories, such as pupil-teacher ratio and school size.

Twenty-three different within-Maine reporting categories were developed based upon data received from students, schools, and department records.

*In the planning stages of the project, it was thought that MAEP results could be compared to NAEP results on the size and type of community. Further evaluation demonstrated that this was not feasible due to an inability to match Maine population groups with the NAEP categories. For example, most of Maine's population lives in communities of less than 10,000 people.

Data were analyzed by AIR and RTI, and results for both the MAEP-NAEP comparisons and the within-Maine comparisons are presented in the Results report. The Results report also covers in detail the manner in which the within-Maine reporting categories were developed, and the way subgroup p-values (percentages) for exercise results in Citizenship and Writing were compared first to national p-values for the MAEP-NAEP comparisons, and then to overall Maine p-values for within-Maine comparisons.

II. SAMPLE DESIGN

General Considerations

The sample design problem for the MAEP project was approached by RTI from a background of substantial experience with the sample design of the National Assessment project. When considering a statewide assessment program, this experience takes on strong meaning as the investigator is faced with far more limited funds to produce the same type of statistically reliable answers, and often far more specific issues and groups for which results must be relevant.

Some considerations of a Statewide Assessment program that might require special attention from a sample design viewpoint are:

1. Should a matrix sampling approach be developed to shorten the length of time a pupil would be tested, i.e., each instrument item in combination with each school and pupil having a positive chance of inclusion in the sample?

2. Should pupils participating in federal compensatory programs or other subpopulations of interest be oversampled so that results for variables relating to these groups can be reported separately?

3. Should the school and pupil sample sizes be such that estimates of the sample variability of key estimates are computable from the sample data, including estimates of the between school and between pupil within school components?

4. How should the sample be allocated across the important regions of the state and types of communities (e.g., Large City, Large City Fringe, and Rural) within these regions and have sufficient schools and pupils to report results of reliable precision for each Region and/or Community Type?

In addition to these types of questions which relate to the sample design and the usefulness of the results obtained from the study, there is a host of technical questions which must be understood by local officials

in order to be able to knowledgeably assist in the decisions. Some of these follow:

1. What are sampling and non-sampling errors?
2. What exactly is the target population to be sampled?
3. How complete and accurate must sampling frame data be?
4. What are the advantages of stratification?
5. How is the sampling frame to be stratified, and by what variables?
6. How do the stratification variables differ from reporting variables?
7. From a sample selection viewpoint, how is the sample to be structured?
8. What stages of sample selection are to be used, i.e., schools or pupils within schools?
9. What sample selection procedures are to be used at each stage of sampling?
10. What are disadvantages of using stages of sampling?
11. What is the distinction between a unit of analysis and a unit of sampling?
12. How should the data be weighted in relation to the sample design in the analysis phase of the study?
13. What equations should be used to estimate population values in relation to sample design?

The amount of time necessary and attention required to the above types of issues depends largely upon the expertise of the State Department personnel most concerned with the project and with sampling theory and practice.

Sample Frame Construction

A general purpose school sampling frame was constructed using the Maine Educational Directory 1970-71* and the Directory of Maine Schools† as primary sources of school information. In accordance with the proposed sample design (a two-stage probability sample of 17-year-old students enrolled in a public or non-public school in the state of Maine), the first-stage sampling frame consisted of all schools, both public and non-public, with 17-year-old pupils‡ enrolled. Special education schools or schools whose enrollment consisted primarily of educable mentally retarded pupils were excluded from the sampling frame. Appendix A describes the codes assigned to each school by the Maine Department of Education.

After completion of the list of schools, the following information was obtained for each school in the sampling frame:

1. Region planning code
2. Region code
3. Grade range
4. Ninth grade enrollment
5. Tenth grade enrollment
6. Eleventh grade enrollment
7. Twelfth grade enrollment
8. Total school enrollment
9. Expenditure per pupil (in dollars)
10. Number of professional staff
11. Zip code for school address

This information was coded and punched on cards. (See Appendix A for card formats.)

---

*Maine Educational Directory 1970-71. State Department of Education, Augusta, Maine.

†Directory of Maine Schools. School Administrative Services, Machine Operations. State of Maine, Department of Education, Augusta, Maine.

‡A pupil is classified as a 17-year-old if his birthdate falls between the dates October 1, 1955, and September 30, 1956.

Stratification of the Sampling Frame

For purposes of reporting results by region, the state of Maine was partitioned into the following four regions specified by the Maine Department of Education:

1. Region I--Northern Maine District
2. Region II--Penobscot District and Eastern Maine District
3. Region III--Androscoggin District, Kennebec District, and Midcoastal District
4. Region IV--Southern Maine District and Cumberland District.

Figure 1 shows the eight Regional Planning and Development Districts, defined by the Maine State Planning Office, which were grouped into the four regions.

The schools within each region were stratified into the following groups on the basis of size of school, using eleventh grade school enrollment as the criterion:

Group Number	Group Name	Eleventh Grade Enrollment
1	Mini Schools	000-019
2	Small Midi Schools	020-199
3	Large Midi Schools	200-294
4	Maxi Schools	295-550

The schools within each of the above groups were divided into two classes on the basis of expenditure per pupil (in dollars):

Class "1": Expenditure less than $800 per pupil

Class "2": Expenditure $800 or more per pupil

For sampling purposes, the schools within the larger classes were ordered by eleventh grade enrollment and divided into subclasses containing as equal a number of schools as possible. For analysis purposes, a four-digit substratum code was developed to identify each substratum by region, size of school (group number), expenditure per pupil (class code), and subclass code. These codes were used in labeling the sample schools. Table 1 gives the number of schools in the Maine Assessment of Educational Progress sampling frame by region--size of school--per pupil expenditure.

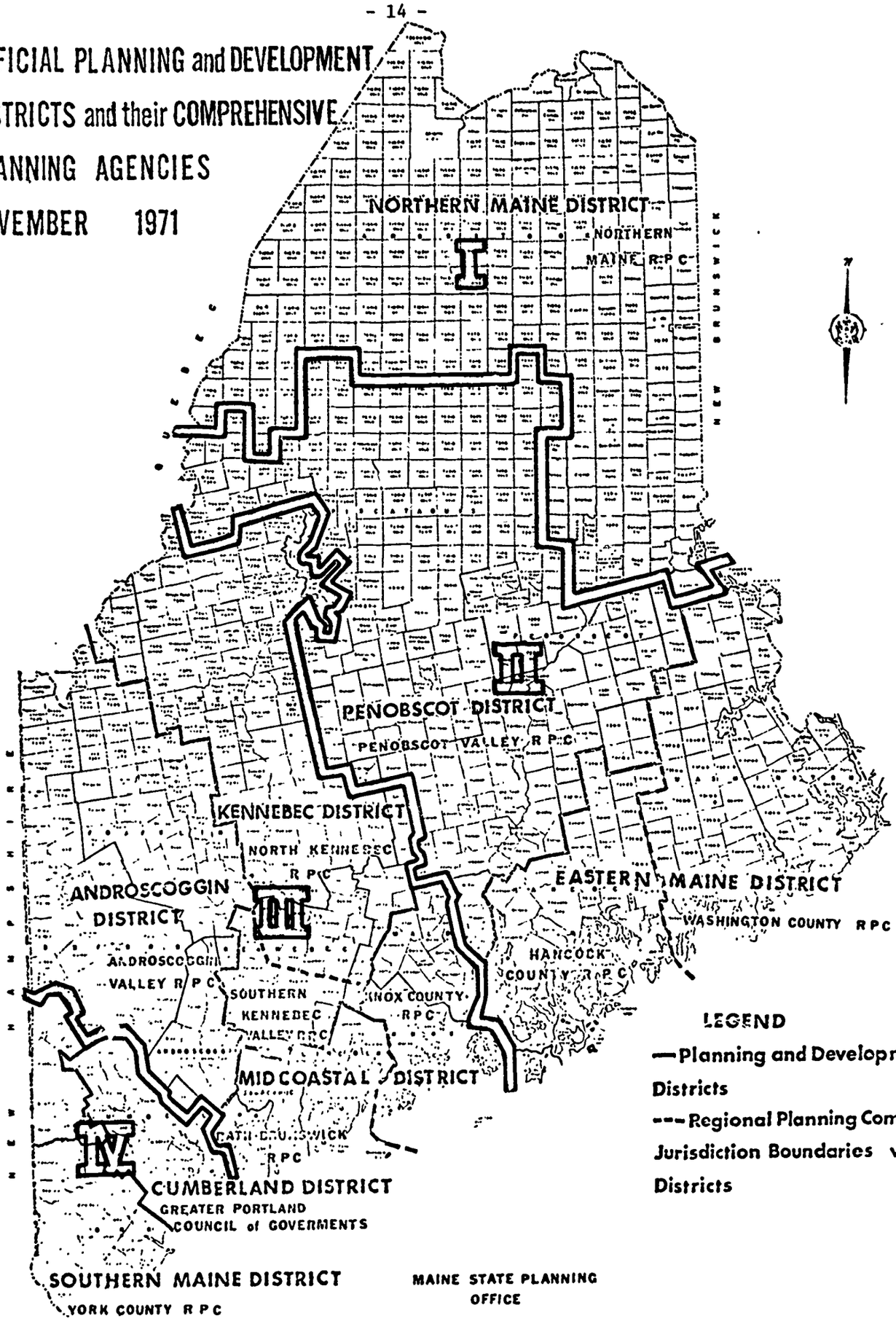


Figure 1. Assessment Regions.

Table 1

Number of Schools in MAEP Sampling Frame by
Region--Size of School--Per Pupil Expenditure Sampling Sub-Strata

	Size of School Strata									
	Mini		Midi				Maxi		Total	
			Non Large		Large					
	Low Per Pupil Expenditure	Non Low Per Pupil Expenditure	Low Per Pupil Expenditure	Non Low Per Pupil Expenditure	Low Per Pupil Expenditure	Non Low Per Pupil Expenditure	Low Per Pupil Expenditure	Non Low Per Pupil Expenditure	Low Per Pupil Expenditure	Non Low Per Pupil Expenditure
I	1	1	8	6	2	0	2	0	13	7
MAEP II	1	6	18	13	1	0	1	1	21	20
Sampling III	9	10	19	24	2	4	4	1	34	39
Regions IV	1	9	7	16	0	2	1	4	9	31
Total	12	26	52	59	5	6	8	6	77	97

Total Number of Schools of Sampling Frame = 174

Sample Allocation

A sample of 2,000 17-year-old pupils was to be selected from among the four regions, with twenty pupils to be selected per sample school for exercise administration purposes. For this sample design, the index of sample precision used was the standard error of an estimated p-value which indicated the sampling variability of the estimated p-values in repeated samplings. In other words, the smaller the standard error, the better the precision of the estimate. For purposes of reporting results by the four regions with a total sample size of 2,000 and 20 pupils selected per sample school, an exercise p-value of 50% had standard error equal to 3.7. For a sample of 2,000 pupils, if 50% of the pupils answered a certain question correctly, the standard error would be 3.7. Using this p-value of 50%, the minimum sample size to be selected from each of the four regions would be 400 pupils. Since 400 x 4 = 1,600 pupils, proportional allocation was used to allocate the remaining 400 pupils among Regions III and IV.

Table 2 gives the number of schools (N) in the sampling frame by region and by size of school as well as the planned number of sample schools (n) in each group. Multiplying each of these planned sample sizes (n) by 20 or 40* gives the resulting pupil sample sizes for each group, recorded in Table 3.

Since the Size of School strata were split into low and non-low per pupil expenditure classes, the school sample sizes shown in Table 3 were also divided accordingly, resulting in the data of Table 4. The actual selection of the sample schools will be discussed in the next section.

School Sample Selection

Simple random sampling was used to select the specified number of sample schools (n) from the schools composing each substratum. Table 5 indicates the number of schools to be sampled from each substratum or subclass

*For the Mini, Small Midi, and Large Midi schools, a sample of 20 pupils was selected; however, for the Maxi schools, a sample of 40 pupils was selected for each exercise administration.

Table 2

Number of Schools (N) and Planned Number of Sample Schools (n)

	Small Midi Schools		Large Midi Schools		Maxi Schools		Mini Schools	
Region	N	n	N	n	N	n	N	n
I	14	14	2	2	2	2	2	2
II	31	15	1	1	2	2	11	6
III	43	18	6	6	5	5	19	6
IV	25	14	2	2	5	5	13	6

Table 3

Pupil Sample Sizes by Groups

	Small Midi Schools Sample 17's	Large Midi Schools Sample 17's	Maxi Schools Sample 17's	Total Sample 17's	Mini* Schools Sample 17's
Region	n'	n'	n'		n'
I	280	40	80	400	40
II	300	20	80	400	120
III	360	120	200	680	120
IV	280	40	200	520	120
Total	1220	220	560	2000	400

*The Mini School Category was ignored in allocating the 400 pupils since in most cases these schools have fewer than twenty 17-year-olds. A decision was also made to select 40 pupils from the Maxi sample schools.

Table 4

School Sample Sizes of MAEP Sample Design by
Region--Size of School--Per Pupil Expenditure Sub-Strata

	Size of School Strata									
	Mini		Midi				Maxi		Total	
			Non Large		Large					
	Low Per Pupil Expenditure	Non Low Per Pupil Expenditure	Low Per Pupil Expenditure	Non Low Per Pupil Expenditure	Low Per Pupil Expenditure	Non Low Per Pupil Expenditure	Low Per Pupil Expenditure	Non Low Per Pupil Expenditure	Low Per Pupil Expenditure	Non Low Per Pupil Expenditure
I	1	1	8	6	2	0	2	0	13	7
MAEP II	1	3	9	6	1	0	1	1	12	10
Sampling III	3	3	8	10	2	4	4	1	17	18
Regions IV	1	2	4	10	0	2	1	4	6	18
Total	6	9	29	32	5	6	8	6	48	53

Total Number of Sample Schools Selected = 101

Number of Sample Schools Not Participating = 4

Total Number of Sample Schools Participating = 97

Table 5

Number of Schools to be Sampled
From Each Substratum

RTI Code	School Selection Code Range	Number Schools in Frame N	Number Schools in Sample N
1111	01	1	1
1121	01	1	1
1211	01-08	8	8
1221	01-06	6	6
1311	01-02	2	2
1411	01-02	2	2
2111	01-02	4	2
2121	01-04	7	4
2211	01-03	6	3
2212	01-03	6	3
2213	01-03	6	3
2221	01-03	7	3
2222	01-03	6	3
2311	01	1	1
2411	01	1	1
2421	01	1	1
3111	01-03	9	3
3121	01-03	10	3
3211	01-03	7	3
3212	01-03	7	3
3213	01-02	5	2
3221	01-03	8	3
3222	01-03	8	3
3223	01-04	8	4
3311	01-02	2	2
3321	01-04	4	4
3411	01-04	4	4
3421	01	1	1
4111	01-02	2	2
4121	01-04	11	4
4211	01-04	8	4
4221	01-03	6	3
4222	01-03	5	3
4223	01-04	6	4
4321	01-02	2	2
4411	01	1	1
4421	01-04	4	4

identified by the unique four-digit RTI code defined in Stratification of the Sampling Frame. An additional two digits added to this four-digit code formed a unique six-digit RTI identification code for each sample school selected. Beginning with "01," the two digits were assigned consecutively (within each subclass) to each school as it appeared in the sample. Random numbers were used to select the required number of schools from each subclass.

A computer listing with the following necessary information was produced:

1. School district identification including Union code (UNC) and school district superintendent's name and address.

2. Ten-digit Maine Department of Education school code (CTY, UNC, AUC, SPN).

3. Six-digit RTI identification code.

4. School identification including town name, school name and address, grade range, estimated number of 17-year-olds, and planned number of 17-year-old pupils in the sample.

Appendix A explains the procedure used to estimate the number of 17-year-olds per school. Card formats for sample school data used to obtain the computer listing are also given in Appendix A, as is a list of all sample schools.

Upon receipt of the computer listing of sample schools, the Maine Department of Education sent to each selected school a Student Listing Form (see Figure 2) with instructions to list the names (as well as sex, grade and birth date information) of all 17-year-old pupils in that school. These forms were returned to the Maine Department of Education and forwarded to RTI for pupil sampling to be discussed in the next section.

At this stage of the assessment project, several sample schools were found to be ineligible.* These schools were dropped from the sampling frame and replacements drawn using simple random sampling as before. The RTI code for each ineligible school was assigned to its replacement, with the fifth

*These schools had no 17-year-old pupils enrolled at that time.

MAINE ASSESSMENT OF EDUCATIONAL PROGRESS

STUDENT LISTING FORMS

Instruction: List all students in your school that were born October 1, 1954 through September 30, 1955 (17 year olds). Please print or type.

Name of School: _____
School Address: _____
Telephone Number: _____
Contact Person if Other Than Principal: _____

STUDENT NAME Last First MI	SEX	GRADE	BIRTH DATE

Figure 2. Student Listing Form.

digit (originally "0") changed to "1," thus denoting a replacement school. At later stages in the project, more ineligible schools were discovered, however, replacements were not drawn for these schools. Table 6 gives the final number of schools (N') in the sampling frame corrected for ineligibles and the corresponding number of sample schools (n') by substratum.

Pupil Sample Selection

All pupil sample selection was done from the lists of 17-year-old pupils submitted by the sample schools.

The pupils within each school were stratified on the basis of grade in the following way:

Stratum I—Tenth graders or less

Stratum II—Eleventh graders

Stratum III-Twelfth graders

A Pupil Sample Selection Form (Figure 3) was designed by RTI staff for use in calculating the number of pupils to be sampled from each of the above pupil strata. Proportional allocation was used to determine these sample sizes.

The following steps were employed in sampling pupils for the Maine Assessment of Educational Progress:

1. All tenth graders or less were numbered consecutively (using a green pencil) on the Student Listing Form illustrated in the previous section (Figure 2).

2. All eleventh graders were numbered consecutively (using a red pencil).

3. All twelfth graders were numbered consecutively (using a black or regular pencil).

4. The number of pupils in each stratum was recorded in the proper place on the Pupil Sample Selection Form and this total checked with the total number of pupil names submitted by the school. All pupil names were assigned a number.

Table 6

Final Number of Schools in Sampling Frame (N') and Number of Sample Schools (n') by Substratum

RTI Sub-stratum Code	School Selection Code Range	Corrected Number in Frame N'	Adjusted Number in Sample n'
1111	01	1	1
1121	01	1	1
1211	01-08	8	8
1221	01-06	6	6
1311	01-02	2	2
1411	01-02	2	2
2111	12	1	1*
2121	02-04	6	3†
2211	01-03	6	3
2212	01-03	6	3
2213	01-03	6	3
2221	01-03	7	3
2222	01-03	6	3
2311	01	1	1
2411	01	1	1
2421	01	1	1
3111	01-03	9	3
3121	01-03	10	3
3211	01-03	7	3
3212	01-03	7	3
3213	01-02	5	2
3221	01-03	8	3
3222	01-03	8	3
3223	01-04	8	4
3311	01-02	2	2
3321	01-04	4	4
3411	01-04	4	4
3421	01	1	1
4111	02	1	1‡
4121	01-02	9	2§
4211	11, 02-04	7	4¶
4221	01-03	6	3
4222	01-03	5	3
4223	01, 12, 03-04	5	4¶
4321	01-02	2	2
4411	01	1	1
4421	01-04	4	4

*Substratum sample size was reduced to one since substratum 2111 contains only one eligible school. This school was originally selected as a replacement school (coded 211112).

†One school in substratum 2121 was determined ineligible at a later stage of the project. Thus no replacement school was drawn.

‡One school in substratum 4111 was determined ineligible at a later stage of the project. Thus no replacement school was drawn.

§Two schools in substratum 4121 were determined ineligible at a later stage of the project. Thus no replacement schools were drawn.

¶One school in substratum 4211 was determined ineligible and a replacement school was selected (coded 421111).

¶One school in substratum 4223 was determined ineligible and a replacement school was selected (coded 422312).

PUPIL SAMPLE DESIGN

FOR SAMPLE SCHOOLS OF MAINE ASSESSMENT OF EDUCATIONAL PROGRESS

Spring 1972

PLACE LABEL HERE

Number of Students to be Sampled __

Stratum Code	Stratum Description	Number of 17 Year Olds	Proportion of 17 Year Olds in Stratum	Proportional* Allocation of Sample	Adjusted† Allocation of Sample	Range of Student Codes to be Used on Labels
I	10th Graders or less					
II	11th Graders					
III	12th Graders					
	TOTAL		1.00			

*Rounded to integers.

†To be used if proportional allocation needs adjustment, for example, at least two pupils should be selected from each stratum.

Name of Sampler: _____

Checked by: _____

Date: _____

Figure 3. Sample Design Form.

5. The proportion of 17-year-olds in each stratum was calculated by dividing the stratum size by the total number of 17-year-old pupils in the school; this value was recorded on the Sample Section Form.

6. Each stratum sample size was calculated by multiplying each respective proportion by the total number of students to be sampled from the school (20 for the Mini, Small Midi, and Large Midi schools, 40 for the Maxi schools). This value was then recorded under the column "Proportional Allocation of Sample" on the Sample Selection Form.

7. When necessary, this calculated sample size was adjusted to satisfy the following criteria:

a. The sample size for Stratum I may never exceed ten, with pupil codes ranging from 00 to 09*.

b. A sample size of at least two must exist for all pupil strata whenever possible.

When the first criterion was not satisfied, the sample size for Stratum I was adjusted to ten and the remainder allocated as equally as possible among Stratum II and Stratum III. When the second criterion was not satisfied (a sample of size one existed), the sample size was adjusted to two and the sample size for Stratum II was altered as necessary to compensate for this increase.

Using the pupil stratum sizes (N_h), the adjusted pupil stratum sample sizes (n_h), and the corresponding pupil codes indicated on the Sample Selection Forms, sample pupils were selected using Simple Random Sampling within each pupil stratum. Using random number tables†, n_h random numbers were selected out of random numbers ranging from 01‡ to N_h. The number 00 and any number larger than N_h were rejected, as well as any number within

---

*The pupil codes for Stratum II ranged from 10-39 and for Stratum III from 40 to 69.

†Fisher, Ronald A. and Frank Yates. Statistical Tables for Biological and Agricultural Medical Research. New York: Hafner Publishing Company, Inc., 1953.

‡The value 1 was used if N_h was only one digit, 01 was used if N_h was two digits, and 001 was used if N_h was three digits.

the indicated range occurring more than once. Thus, n_h unique random numbers were selected, and the corresponding numbers on the Student Listing Form circled in the appropriate color (green, red, or black). The RTI code and the pupil codes were assigned consecutively to these circled pupil names, thus identifying the sample students.

Sample Student Listing Forms (Figure 4) were generated in triplicate by computer, and the sample pupil names (and additional information for each student such as pupil code, sex, grade, and birth date) were typed on these forms by RTI and Department staff. These forms were completed by the exercise administrator after actual exercise administration.

SCHOOL NAME: ALLAGASH
ADDRESS: ST FRANCIS MAINE 04774

THIS PORTION OF THE FORM WILL NOT LEAVE THE SCHOOL SINCE THE NAME OF THE STUDENT IS NOT TO BE IDENTIFIED WITH THE FILE. CUT HERE

MAINE ASSESSMENT OF EDUCATIONAL PROGRESS
17 YEAR OLD STUDENTS SELECTED FOR SAMPLE

MAINE SCHOOL CODE
CTY UNC AUC SPN LEV
02 510 007 01 3

NO	STUDENT NAME	TO BE VERIFIED AND COMPLETED BY SCHOOL: IDENTIFICATION RTI SCHOOL CODE	PUPIL CODE	SEX	GRADE	BIRTHDATE MONTH/YR	BILINGUAL (CHECK ONE) YES	NO	NO	TO BE COMPLETED BY EXERCISE ADMINISTRATOR: PARTICIPATED YES	NO	REASON FOR NOT PARTICIPATING (ABSENT, ETC.)
1		1 112101							1			
2		2 112101							2			
3		3 112101							3			
4		4 112101							4			
5		5 112101							5			
6		6 112101							6			
7		7 112101							7			
8		8 112101							8			
9		9 112101							9			
10		10 112101							10			
11		11 112101							11			
12		12 112101							12			
13		13 112101							13			
14		14 112101							14			
15		15 112101							15			
16		16 112101							16			
17		17 112101							17			
18		18 112101							18			
19		19 112101							19			
20		20 112101							20			

Figure 4. Student Listing Form.

III. THE EXERCISE PACKAGE

Development

The exercise package for the Maine Assessment of Educational Progress was a 32-page booklet printed specially for this assessment project by the Measurement Research Center in Iowa City. It consisted of three basic elements divided into two separate parts. The first part, 27 pages long, contained the 23 Citizenship and 7 Writing exercises selected for administration; and the second part consisted of the 23 items of the Student Questionnaire. A complete exercise booklet is included in this report as Appendix B.

The development of the package began in March, 1972, quite soon after the basic decision had been made concerning the two NAEP subject matter areas to be covered in the assessment. Knowlegeable personnel from AIR, NAEP, and RTI worked with the Maine Department of Educational and Cultural Services in Augusta to select from among released NAEP exercises in Citizenship and Writing those which were best suited to the needs of Maine. Another consideration throughout was that some of the exercises would have to be changed from the NAEP individual administration or interview technique mode to the paced-tape, group mode of administration. Changes which might take place in the interpretation of the results because of this change had to be carefully considered. Finally, administration time became important because of the desire that each student complete a full booklet. Less than two hours, including a "stretch" break, was desirable; and this limited both the number and types of exercises which could be considered.

Integral to these discussions was consideration of the Student Questionnaire section of the booklet. While the assessment planners recognized the desirability of collecting ample demographic and background data so that in the analysis phase there might be more flexibility in delving into the results, they also were aware of both administration time and personal privacy constraints. The 23 questions finally selected reflect several separate yet interconnected concerns:

Figure 4. Student Listing Form.

1. The desire to collect data and relate it to results on factors perhaps unique to Maine which were not covered in the NAEP reporting categories. Examples of this are the bilinguality situation in the state, and the quite different range of community sizes in the state as compared to the NAEP categories.

2. The desire to collect data on dimensions comparable to other studies, including NAEP, so that some sorts of comparisons could be drawn along these dimensions. Examples of these are levels of parents' highest education, reading material in the home, and parents' occupation.

3. The desire to look into some dimensions on which little if any data exists in Maine, such as amount and degree of parental and community support of the school, and the personal satisfaction and happiness of students in school.

The planners realized that due to statistical considerations relating to sample size and precision of estimates, not all of these potential reporting categories would necessarily be used to group Citizenship and Writing results. But they would at least be able to provide guidance for more profitable avenues to explore in subsequent years of the assessment project.

Citizenship Exercises

More than four years of work went into the development of the plan and instruments for assessing the attainment of young Americans in the area of Citizenship before the first exercises were administered by NAEP in the spring of 1969 to a random sample of 17-year-old students throughout the country. These objectives and the procedures leading to their development are fully described with clarity in the pamphlet Citizenship Objectives available from NAEP*. For the Maine Assessment project, exercises were selected from all but one of the nine NAEP Citizenship objective areas.

*Education Commission of the States, Denver, Colorado.

It is important to an understanding of the Maine Assessment project to realize that for the NAEP project the task of developing objectives in the field of Citizenship was awarded to the American Institutes for Research in Palo Alto, the same AIR that is part of the Research Consortium for Educational Assessment. Thus selection of exercises, potential effects of any administration changes, and any scoring anomalies which might arise could all be discussed and resolved by personnel intimately familiar with the area.

The pamphlet Citizenship Objectives describes the nine objectives and lists by age group (9, 13, 17, and young adult) both the sub-objectives which state more specifically the kinds of information or skills which define that major objective, and some suggestions or examples of behavior which might be indicative of these objectives being attained. Table 7 shows those objectives, the sub-objectives, and the number of the NAEP exercise which was developed to elicit student behavior to measure the objective. In the Table, the exercise numbers shown are limited to those 23 which were selected for the Maine Assessment project--there are many more released NAEP exercises which were not used. The actual format of all released exercises is given in Report #2 Citizenship: National Results (also available from NAEP), and examination of both this report and the Maine exercise booklet (Appendix B) will reveal the virtual identicality of comparable exercises.

An example will show how an exercise relates to an objective. Referring to Table 7, Objective I is:

> Show Concern for the Welfare and Dignity of Others

One of the seven sub-objectives is labeled "D":

> D. Help other individuals voluntarily

Quoting now from the pamphlet Citizenship Objectives for age 17, suggestions or examples of behavior illustrating attainment of this sub-objective are:

> Although youths might find it difficult to give money or to play an active political role, they are as able as any other age group to help others by their own individual effort. Thus, they help persons who are lost or in trouble (with due

consideration for their own safety); help instruct classmates who have been absent from school; help new students adjust to school; defend younger children against bullies; help or console friends with personal problems; and volunteer to aid others in constructive achievements such as learning to play a musical instrument or finding something in the library.

The exercise which was developed as a result of the above, and which was used in the Maine project, is number five in the exercise booklet (Appendix B) or NAEP exercise A-3 as shown in Report #2:

5.A. Suppose you and some friends were walking by a public park. As you went by, some children of a minority group were stopped from entering the park by a man at the gate who told them, "The park is not for kids like you." Would you feel that you should do something about it?

_____ Yes

_____ No

5.B. What could you do about it if you wanted to? (Maximum of four)

(1) _____

(2) _____

(3) _____

(4) _____

The scoring for this exercise was two-fold. First, part A was tabulated for the number (or percentage) of students answering Yes or No. Then the actions cited for part B were first categorized by content, then classed as Acceptable or Not Acceptable ("Tell the kids to go on in anyway"), and finally reported as number of Acceptable answers. Looking ahead to the Results Report of the Maine Assessment project, it will be seen that two reporting categories were finally developed for this exercise: per cent answering Yes to part A, and per cent able to give two or more acceptable actions.

The scoring of the remainder of the Citizenship exercises will be discussed more fully in Section V of this report as well as in the Results Report of the assessment project. Both the relationship of other exercises

Table 7

Maine Assessment of Educational Progress
NAEP Objectives and Selected Items Designed to Measure Them

	NAEP Exercise* Numbers	Maine Exercise Numbers
CITIZENSHIP		
Objective I. Show Concern for the Welfare and Dignity of Others	--	--
A. Treat all individuals with respect; do not condemn others on the basis of irrelevant personal or social characteristics.	A-3, A-4	5†, 25
B. Consider the consequences for others of their own actions.	--	--
C. Guard safety and health of others.	--	--
D. Help other individuals voluntarily.	A-3	5†
E. Are loyal to country, to friends, and to other groups whose values they share.	A-8	2
F. Understand and oppose unequal opportunity in the areas of education, housing, employment, and recreation.	A-8	2
G. Seek to improve the welfare of groups of people less fortunate than they.	--	--
Objective II. Support Rights and Freedoms of all Individuals	B-4	8
A. Understand the value of constitutional rights and freedoms.	B-1	10†
B. Recognize instances of the proper exercise or denial of constitutional rights and liberties, including due process of law.	B-1	10†
C. Defend rights and liberties of all kinds of people uniformly.	--	--

Table 7 (Continued)

	NAEP Exercise* Numbers	Maine Exercise Numbers
Objective III. Help Maintain Law and Order	C-2	22†
A. Understand the need for law and order.	C-3	26
B. Are conscious of right and wrong behavior; encourage ethical and lawful behavior in others.	--	--
C. Comply with public law and school rules.	--	--
D. Help authorities in specific cases.	--	--
E. Protest unjust rules openly.	--	--
F. Inform themselves about the law.	--	--
Objective IV. Know the Main Structure and Functions of our Governments	--	--
A. Recognize the purposes of government.	D-1	18, 26
B. Recognize the main functions and relations of governmental bodies.	D-2	21†
C. Recognize the importance of political opposition and diverse interest groups.	D-9, D-5, D-6	12, 19, 23
D. Recognize that democracy depends on the alertness and involvement of its citizens, and know how citizens can affect government.	--	--
E. Recognize the structure and operation of political parties.	D-3, G-7	6, 13
F. Know structure of school and student government.	E-1, E-2 D-4	3†, 7†, 29
Objective V. Seek Community Improvement through Active, Democratic Participation	--	--
A. Believe that each person's civic behavior is important, and convey this belief to others.	--	--
B. Recognize important civic problems and favor trying to solve them.	--	--

Table 7 (Continued)

	NAEP Exercise* Numbers	Maine Exercise Numbers
C. Actively work for community improvement.	E-5	11
D. Participate in local, state and national governmental processes.	--	--
E. Apply democratic procedures on a practical level when working in a group.	--	--
F. Display fairness and good sportsmanship toward others.	--	--
Objective VI. Understand Problems of International Relations	F-4	15†
A. Are aware of the problems of international conflict and dangers to our own national security.	--	--
B. Seek world peace and freedom for all peoples.	--	--
Objective VII. Support Rationality in Communication, Thought and Action on Social Problems.	--	--
A. Try to inform themselves on socially important matters and to understand alternative viewpoints.	G-9, G-4	14, 20†
B. Evaluate communications critically and form their own opinions independently.	--	--
C. Weigh alternatives and consequences carefully, then make decisions and carry them out without undue delay.	--	--
D. See relations among social problems and have good ideas for solutions.	--	--
E. Support free communication and communicate honestly with others.	B-4, E-5, G-9, G-10	8, 11, 14, 28
F. Understand the role of education in developing good citizens.	--	--

Table 7 (Continued)

	NAEP Exercise* Numbers	Maine Exercise Numbers
Objective VIII. Take Responsibility for Own Personal Development and Obligations	--	--
A. Further their own self-improvement and education.	H-4	1
B. Plan ahead for major life changes.	H-4	1
C. Are conscientious, dependable, self-disciplined, and value excellence and initiative.	--	--
D. Economically support self and dependents.	--	--
Objective IX. Help and Respect Their Own Families.	--	--
A. Respect the reasonable authority of their parents, or guardians, and help with home duties and problems.	--	--
B. Help younger brothers and sisters to develop into good citizens.	--	--
C. Discuss social matters with their families and respect the views of all family members.	--	--
WRITING		
Objective I. Write to Communicate Adequately in a Social Situation	302	30
Objective II. Write to Communicate Adequately in a Business or Vocational Situation	305, 304	9, 24
Objective III. Write to Communicate Adequately in a Scholastic Situation	310	16
Objective IV. Appreciate the Value of Writing	--	--
A. Recognize the value of writing for social, business and scholastic needs.	--	--

Table 7 (Continued)

	NAEP Exercise* Numbers	Maine Exercise Numbers
B. Write to fulfill these needs.	308, 306	4, 27
C. Get satisfaction from having written something well.	307	17

*See NAEP Reports #2 and #6 (Citizenship) and Reports #3 and #5 (Writing) for exercise text, format, administration procedure, and p-values.

†Differences in administration between NAEP (interview mode) and Maine (paced tape group mode), or in scoring (local vs. MRC), could prevent clear comparison of NAEP results with Maine results.

to objectives, and the general mode of scoring and reporting are quite similar to the example given.

Writing Exercises

Just as with the Citizenship exercises, more than four years of work went into the development of the plan and instruments for assessing the achievement of young Americans in their writing ability and use of it as a mode of communication. Citizenship, Science, and Writing were the three subject matter areas included in the first NAEP nationwide assessment in 1969. There are four basic writing objectives specified in the pamphlet Writing Objectives available from NAEP:

1. Write to Communicate Adequately in a Social Situation
2. Write to Communicate Adequately in a Business or Vocational Situation
3. Write to Communicate Adequately in a Scholastic Situation
4. Appreciate the Value of Writing

The task of developing these objectives and their associated behavioral manifestations was awarded in 1965 to the Educational Testing Service in Princeton, New Jersey. The procedures of staff and committee members are well discussed in the pamphlet, as is the rationale for the development of the objectives themselves. In a manner similar to that already discussed for Citizenship, Table 7 shows these objectives, sub-objectives where applicable, and the NAEP exercises selected for use in the Maine Assessment project. The NAEP publication Report #3 Writing: National Results shows the format of all released exercises in this area, as well as national results, and again, examination of this report and the Maine exercise package (Appendix B) shows the identicality.

A useful classification of the seven writing exercises used in Maine is by the type of task to be performed. There are two main classes which can be identified: performance exercises and self-report exercises. Referring to Table 7, the exercises in Objective IV are of the self report variety. Number 4 and 27 for Maine ask whether the respondent has written a mail order

or a message for someone in the past 12 months, and number 17 inquires if the respondent has ever done certain kinds of creative writing not required as a school assignment.

The performance classification has two sub-categories represented by two exercises each. The essay performance set consists of number 16 requiring the respondent to write a composition about a famous person, and number 30 requiring the respondent to give explicit directions on how to do something. The nonessay performance set consists of number 9 requiring the respondent to write a short business letter to order merchandise, and number 24 asking the respondent to fill out an application blank.

The self-report exercises were scored in a conventional manner by simple tabulation of the number of Yes responses. The performance exercises, however, required the services of professional scorers to assess a "quality level" for each exercise response and then compare that level with a set of standard previously scored responses to assign a score. The same criteria were used in the Maine project as had been used in NAEP. This procedure will be discussed in more detail in Section V of this report.

Student Questionnaire

The introduction to this chapter discussed some of the concerns felt by the Maine Assessment project planners when designing the exercise package: unique-to-Maine data were needed, comparable-to-other-studies data were needed, and data on "new" dimensions were needed. Yet there were time, personal privacy, and statistical constraints to be faced.

Among the original desires of the planners was to consider the effects on results of the region of the state the student was from, the size and/or type of community, parent occupation, level of parental education, and school program the student was enrolled in. These are dimensions common to most assessment programs, and their effects have generally been well researched. Their inclusion was designed to provide a degree of validation of their effects in Maine as well as some degree of comparability with other studies. A further and significant benefit would be possible descriptive results

showing how the different aspects were distributed in Maine and some of their potential interactions.

Maine assessment planners felt that there might be some currently hidden effects due to the suspected prevalence of bilingual students in Maine, or at least families in which more than one language was habitually spoken. These effects might not only be related to output (Citizenship and Writing) results of the assessment, but they might also be interrelated with other descriptive statistics. Planners were similarly concerned with parental support of a student's efforts in school, and even with some dimensions of a student's satisfaction with himself and his role, position and success in school. There could be some factors which would emerge from these types of questions which were well worth the time and effort to discover them; and the minimum result from their inclusion would perhaps be suggestions as to how better to ask questions relating to these types of issues. Accordingly, the 23 questions which appear in the Student Questionnaire section of the exercise package (Appendix B) were carefully worded, considered, and placed in the booklet.

Package Assembly

After exercise selection had been completed by the planners, and after the precise wording of the Student Questionnaire items had been agreed upon, the 30 exercises and 23 questions were assembled into an Exercise Package by personnel at Research Triangle Institute experienced in packaging NAEP materials. This was accomplished in close cooperation with Measurement Research Center who would print and score the booklets as they had been doing with NAEP. It was important that the format be as close to identical with NAEP as possible, that the sequence of exercises be appropriate, and that the answer marking and scoring areas be as free from causing errors as possible.

Of particular interest was the paced-tape which was created to go along with the booklet. Virtually every word in the booklet was read aloud page by page so as to forestall differential reading ability from affecting the results. Students were not permitted to turn a page until cued by the tape.

After the Maine Department of Educational and Cultural Services was informed of the names of each of the schools selected in the sample (there were 97, listed in Appendix A), these schools were asked to complete a student listing form containing the names of all 17-year-old students enrolled. Name, sex, grade, and birthdate were requested for each student. These lists when completed were forwarded to RTI, and a sample of pupils was drawn from each of the participating schools. The names of the pupils selected in the sample were sent to and retained by each participating school while an identification code was created to relate each student to a particular exercise booklet. Labels were printed for each exercise booklet. The linkage of identification code to student name was retained by the schools, thus assuring confidentiality of responses. Labeled booklets were grouped by school and by exercise administrator (see Section V) for actual administration in May, 1972.

IV. SUPPLEMENTARY INFORMATION

Principal's Questionnaire

In the planning phase of the assessment project it was decided that by means of a short questionnaire each principal of each sampled school would provide some additional information about his school at the time of the administration of the exercises. A copy of the Principal's Questionnaire that was developed is in Appendix C. This information was planned to be used for two purposes:

1. Some of the information would be used as an accuracy or verification check against similar questions being asked of students in the Student Questionnaire. Examples of this are questions relating to the size of community of attending students' homes, school curricular distribution among students, degree of parental support, and percentage of bilingual students. This type of information could also be used, if desired, to characterize the school along these dimensions if analyses using such characterizations became desirable.

2. Some of the information could be used to explore a dimension of principal's perception of the adequacy, innovative and liberal qualities, and support and use of his school facilities and personnel. Questions along these lines dealt specifically with innovative techniques, a wide range of school aspects in the areas of facilities and personnel, and policies. It was hoped that such a pool of data could lead to a measurement of a leader's perception of what he has to work with, how well he has succeeded, and where he might go in the future. These perceptions could have a bearing on leadership style, which in turn will affect performance of both teachers and students.

Because of the experimental nature of the questionnaire and the fact that it was only a first tentative step into the domain, it was kept to only two pages and required only a few moments to complete. The unit of analysis for the Maine Assessment project was to be the student rather than the school.

Accordingly, the information taken by means of this questionnaire was not planned to be a definite part of the analysis of student achievement data.

Department Records

In order to take up a minimum of the sampled school principal's time on assessment day for his school, as much information as possible needed in the project was taken from records maintained in Augusta by the Maine Department of Educational and Cultural Services. Department officials were extremely cooperative in making this information available even to the extent of making special computer runs on their data bank to isolate needed data. Thus for the Maine project, no necessary data had to be collected from a principal about this school. A listing of departmental data collected is in Table 8.

Summary of Data Collected

In the compilation of a data record on tape for each student, there were two basic classes of information retained. The first was his scored responses to each section of each of the Citizenship and Writing exercises discussed in Section III of this report. This amounted to 87 different bits of information for each sampled student, because most of the exercises have more than one scored response. For the purposes of this assessment, this classification of information was considered the product, the output, or the dependent variable.

Also on each student's data record were his responses to each of the student questionnaire items (23 items, 32 responses), his school principal's responses to each of the 8 questions on the Principal's Questionnaire (33 responses), and 21 items of information about his school taken from departmental records in Augusta. The manner in which they were grouped for reporting purposes is covered in the Results Report.

Table 8

School Information Collected From Departmental Files

Data	Comments
1. School Identification	Name, address, and 11-digit Maine identification code showing location, classification, etc.
2. Region	Planning and Development Districts (8) grouped into 4 regions
3. Grade Range	1-12, 8-12, or 9-12
4. Number of secondary teachers	Per school
5. Average teacher salary	Secondary only, including full time equivalents
6. Grade enrollments	9, 10, 11, and 12
7. Per-Pupil expenditure	To be adjusted for school size in the analysis
8. Professional staff	Entire school, including specialists
9. Zip Code	For relating to U.S. Census data
10. Eight writing-related courses	Taught-Not taught
Handwriting-Penmanship Linguistics Grammar Composition Theory of Composition Expository Writing Creative Writing Communication Skills	
11. Citizenship education	Taught-Not taught

V. ADMINISTRATION AND SCORING

Time Frame of the Maine Assessment Project

Considerable background work was accomplished building the concept of an assessment of educational progress for the State of Maine during the winter of 1971-72; and a proposal submitted to the state by RTI on behalf of the Research Consortium for Educational Assessment resulted in a contract being executed on 7 March, 1972, for the assessment. Promptly that month a meeting was held in Maine to make decisions regarding the adoption, selection and packaging of the Citizenship and Writing NAEP materials; the stratification of the school sampling frame in four regions defined by the Department research and planning staff; procedures to be used to obtain the cooperation of sample schools; and an overall schedule of the assessment activities. Before the end of March the sample schools had been selected and each of these had been requested to submit a list of names of all enrolled 17-year-olds.

April saw the completed lists of students in the schools sent to RTI, the sample selected, and the names of the sampled students sent back to the schools. The exercise booklet was developed jointly by RTI, AIR, and MRC based on the March decisions, and when finally approved by all parties concerned, it was sent to MRC for printing. Final decisions were made regarding exercise administrator training sessions and the actual conduct of the assessment sessions. The RTI Project Director met with a group of Superintendents of Education in Maine to discuss the analysis plans of the assessment project. The paced-tape was developed and recorded. All materials, tapes, tape recorders and labels were taken to Maine for the training session held in Augusta late in April. Ten exercise administrators had been assigned by the Department, and RTI field supervisors experienced in the administration of NAEP exercise conducted the training session. The exercise manual (Appendix D), developed in detail specially for this assessment, was covered quite carefully, with special consideration given

to administrative procedures, the student questionnaires, the principal questionnaires, and the stimulus tape procedures. Each exercise administrator was given a set of materials for his particular school as well as note pads and pencils.

Assessment began on Monday, May 1, as full scale data collection was launched throughout the state of Maine. An RTI staff member on location assisted the Maine coordinator in monitoring field operations and assigning personnel to make-up administrations. He also completed several school administrations as well as schooling department personnel in check-in procedures to establish a regular system for editing and accounting for all work completed in the field. Except for one school, assessment was completed by May 16. Data from all schools were checked to insure that all packages, completed and uncompleted, were accounted for. Materials slated for return to RTI were gathered, boxed and shipped, while completed exercise booklets were packaged by schools and shipped to MRC. All arrangements were completed by the end of the month.

June was occupied mostly by MRC preparing to score, and later scoring the exercises. Before any items were actually scored, each exercise booklet, student questionnaire and principal's questionnaire was checked by MRC staff for correct coding, e.g., illegal codes, inappropriate information, and other problems. For closed questions, answer codes were transferred to a special area in the exercise booklet to be keypunched. For open-ended questions, professional scorers evaluated each student's answers and assigned numerical scores indicating the appropriate NAEP response category code. The data was then processed by placing each item score(s) on a magnetic tape. The weight data (by pupil) produced by RTI was later merged with the pupil information on the completed tape. A program check was also made to insure that pupil codes for both data sets matched.

Concurrent with the work at MRC, programmers at RTI were creating the structure of the data file using the tape format developed at MRC of each pupil's items. Close cooperation was necessary so that when the tape arrived at RTI, it could be processed using software that had already been

developed. Thus the format defining the respective record of data for each pupil in the sample that participated was being developed at the same time as the software to use it in the initial phase of the data analysis.

During the middle of July as the analysts were beginning to get the feel of the assessment data, a meeting was held in Augusta to reach agreements regarding the basic elements of the analysis plan. It was decided that comparisons would be made of Maine results with appropriate NAEP categories, and that within-Maine reporting categories based on the background and student information collected would be developed as statistically feasible. Furthermore, Department planners agreed that the Department should undertake the responsibility of preparing an interpretive, action-oriented analysis of the RTI final statistical report, specifically pointed toward issues and problems in Maine.

August and September were both fully occupied at RTI with data analysis as the MAEP-NAEP comparisons were completed and preliminary results were brought to Augusta and discussed with Department planners and researchers. The wealth of background data was examined and culled for useful variables. This latter task required running numerous frequency counts on the grouping variables and, in conference with department personnel, making decisions on which variables to group and in what manner in order to retain both meaning and statistical precision. The results of these months of work is the subject of the Results Report of the project.

Summary of Field Operations

Table 9 shows the results of the field operations conducted in the first two weeks of May. The data shown reflect the status of the assessment information after it was collected and sent to MRC for scoring, but before the completed student tape records were examined for completeness and accuracy.

The original sample drawn for the study included 106 public and private schools. Of these, a total of four private schools declined participation and five other schools reported no eligible 17-year-old students. Assessment

Table 9

Maine Assessment of Educational Progress
School and Pupil Sample Losses

		Number	Percent
I.	Sample Schools		
	A. Total Number of Sample Schools	106	100
	B. Number of Schools Having No Eligible Students†	5	4.7
	C. Number of Eligible Sample Schools	101	100
	D. Total Number of Eligible Schools Not Participating	4	3.9
	E. Total Number of Eligible Schools Participating	97	96.1
II.	Principal's Questionnaires		
	A. Total Number of Questionnaires Returned by Participating Eligible Schools	97	100
III.	Sample Students		
	A. Planned Number of Sample Students Selected from Participating Schools	2076	100
	B. Actual Number of Students Participating	1808	87
	C. Number of Students Not Participating	268	13
IV.	Breakdown of Non-Response by Reason		
	A. Number of Students Not Participating Because of:		
	1. Absence from School	176	65.6
	2. Refusal	32	11.9
	3. Ineligible by Birthdate	16	5.9
	4. Withdrawal from School (No longer enrolled)	14	5.2
	5. Transferral	6	2.2
	6. Suspension	5	1.8
	7. Conflicting Activities	5	1.8
	8. Other Tests	4	1.4
	9. Late Arrivals	4	1.4
	10. Work Study Programs	1	.3
	11. Attendance at Off-Site Classrooms	1	.3
	12. Independent Study Program	1	.3
	13. Cheating	1	.3
	14. Unknown Reasons	2	.7
		268	100.0

*The data for the survey were collected from May 1, 1972 to May 16, 1972 from 17-year-old students attending both public and non-public schools on NAEP Citizenship and Writing Exercises.

†Schools thought to have 17-year-old students at sampling time because they had at least one of the grades 9, 10, 11, 12 but when actually contacted by Maine State Department for study participation the contrary was found.

packages were administered in all of the remaining 97 schools. Also a school questionnaire was completed and returned by the principal of each of the participating schools.

In all, 2076 students were sampled from the eligible enrollments of the 97 sample schools. Of this number, 87 percent or 1808 students participated. Thirteen percent or 268 did not.

The most common reasons for non-participation were absence from school, personal refusal, ineligibility and withdrawal from school. One-hundred and seventy-six or 65.6 percent of the non-participating students were absent and unable to participate. Another 11.9 percent or 32 of the non-participating students declined the invitation to take part in the study, and 5.9 percent or 16 non-participating students were found to be ineligible due to the fact that their birthdates were not within the prescribed range. Fourteen students or 5.2 percent of those not assessed were no longer enrolled at the time of administration.

Less common factors in non-participation were transferrals, suspensions, conflicting activities, other tests, and late arrivals. Six of the non-participating students or 2.2 percent had transferred to other schools after being selected as part of the sample. The number of students involved in school activities conflicting with the assessment schedule and students suspended from school totaled five each or 1.8 percent each. Four students had prior test commitments and four arrived late at the assessment site, each group accounting for 1.4 percent of all non-respondents.

The four remaining categories of non-response are: work study programs which detained one student, attendance at off-site classroom which accounted for one sample student, independent study programs which also rendered one student unavailable and the act of cheating in which one particular student was discovered and subsequently discounted. Each of these categories accounted for only .3 percent of those students not assessed. In addition two other students or .7 percent were not available for package administration. The reasons for their absence, however, could not be ascertained.

In review of all field operations it is felt that each objective in regard to school participation, percentage of response, interest on the part of school administrators and students alike, and the overall quality of completed data was met and satisfactorily fulfilled.

Exercise Scoring

Students in Maine participating in the assessment were asked either to select a correct answer from among a set of alternatives, or to construct an acceptable answer in the form of a self report, a written short answer, or an essay or set of directions.

Scoring of the exercises depending mainly on recall, i.e., those which required the student to select from among alternatives, was relatively simple. For example, the question, "When might a state have more senators than it has representatives?" has only one correct answer among the four from which to select. Similarly, when asked if he would be willing to have a person of a different race live next door, the student merely had to select either "willing to" or "prefer not to." A third type of exercise in this scoring category required that the student respond either Yes or No to a self report item, and then give some sort of verifying information. For example, "During the past 12 months, did you leave a written message for some person?" was followed by "About how many times?" In scoring this type of question, a Yes response was coded only if the second part of the question was also answered, in an attempt to avoid using capricious or socially desired responses, or responses with no thought behind them.

The distinction between selected responses like these, and constructed responses, becomes quite clear when we realize that with the latter we must identify in advance the criteria by which we will distinguish between acceptable and unacceptable responses. For selected responses, there is simply (in most cases) a correct answer. For constructed responses written to see if an objective has been met, criteria and standards must be specified and each response judged accordingly. Thus in the reporting of these exercises, it must be emphasized that when the word "acceptable" is used in

describing a response, points of reference are objective standards and criteria established by the NAEP project.

The criteria of adequacy or acceptability of responses for the open-ended questions in the Maine Assessment project were those of the National Assessment project. Responses to these exercises were categorized and scored (or judged) by trained persons in accordance with previously written instructions. For many exercises these categories are of interest in themselves and will be reported with the exercise in the Results report. An answer was judged acceptable if and only if it demonstrated the achievement or understanding sought in the exercise. The results will be presented in terms of the percentage of students giving the correct response, or demonstrating the desired achievement or understanding.

Scoring the Citizenship exercises in accordance with these scoring concepts presented the trained scorers with the problems of detecting capricious answers, recognizing repetitions along the same theme, and judging depth of understanding. For example, in offering up to four ways to avoid future wars, the student had to write his response. Any plausible way to avoid war was accepted without any implication as to general agreement about its effectiveness. Even opposite ideas like maintaining a strong defense and disarmament were considered equally acceptable. However, a common unacceptable response would be having or keeping peace, or not fighting. Without an explanation of how, this was too vague rather than incorrect.

Writing exercises of the performance type, whether essay or non-essay, presented problems relating to the quality of the response. One exercise asked the respondent to write a letter ordering a pair of sea horses from an advertisement; this was judged on adequacy of information supplied. Another exercise required a written set of directions on how to do something. Quality of the response began to enter the scoring as responses were judged as being acceptable and specific, or acceptable but not specific, or not acceptable. The Famous Person essay, however, was scored and judged solely on the basis of quality.

The Famous Person essay (see Appendix B, number 16 for the complete text of the exercise) required that the student write an essay about a famous person whom he admired. He was asked to select a particularly admirable characteristic or quality of that person and write 200-250 words describing that characteristic and include an illustration of it from the person's life. The item was scored by the holistic technique developed by the Educational Testing Service and used by the Measurement Research Center for the NAEP project.

Holistic scoring is a technique for ranking papers. Each response is given a rating of 0 to 8 depending on how well written it is. There are no fixed rules to specify what "well written" means, although readers are free to discuss why a given response is good or bad. The standards used were those of NAEP, and each response was scored twice in such a way that neither reader knew the score given by the other person. In addition, a set of Xerox copies of National Assessment papers (with old scores removed) was placed randomly throughout the Maine papers and read again at the same time the Maine papers were read. The readers were all experienced in holistic scoring at the high school level, and many were previous NAEP scorers.

The "second reading" of the NAEP papers provided a way to compare the NAEP holistic reading with the Maine holistic reading. Table 10 gives the data for the papers read. Unfortunately, the sample size is small because about half the NAEP copies were so light a second reading could not be done.

The second column of Table 10 tells the score given in the first holistic reading. The samples that were selected for this study were ones in which the two original readers gave identical scores. Thus, paper #02656 originally received two scores of 1, and paper #04642 originally received two scores of 8. The third column of Table 10 gives the average of the two reader's scores during the Maine Reading. For most score points the Maine readers were reading slightly more leniently than the original readers. This was due to the fact that the papers were, on the whole, a little poorer than those found in National Assessment. For high level papers, the Maine readers seemed to be a little stricter than the National

Table 10

Comparative Scores on First and Second Readings,
NAEP Famous Person Essay Exercise

I. D. Number	First Reading Score NAEP Reading	Second Reading Score Maine Reading	
02656	*1	*1	
02486	1	2	
35108	1	1	1.4
02448	1	2	
40086	1	1	
13241	2	2	
02280	2	2	
38021	2	2.5	2.6
20672	2	3	
40036	3	3.5	
09903	3	2	
16061	3	4.5	3.5
04886	3	4	
03296	3	3.5	
12924	4	5	
33265	4	4	
00880	4	6	4.6
41262	4	3.5	
00311	5	5.5	
20452	5	5.5	
22852	5	6	5.4
41614	5	5	
07838	6	6.5	
41250	6	4.5	
33473	6	6.5	6.2
20828	6	7.5	
12314	7	6	
21901	7	6.5	
37637	7	6.5	6.5
00323	7	7	
31487	8	7.5	
02113	8	8	
04410	8	8	7.8
04642	8	8	

*Average score of two readings.

Assessment readers, although it is difficult to feel confident of this since the sample size is so small. (In any re-reading, the only possible result is that "16" papers would get a lower second score, since it is the end point. Since no reader can give a higher value than 8, the second total score must necessarily be ≤ 16.)

Results for this exercise will be presented in terms of percentages of students who earned a combined score (0-16) of 4 or more, 10 or more, and 13 or more. These three categories constitute the Low, Medium, and High response categories respectively.

Appendix E gives all scoring categories for the open-ended exercises used in the Maine Assessment project. Also included are comments and comparisons relative to NAEP scoring categories which would assist in arriving at a full understanding of responses.