

DOCUMENT RESUME

ED 279 672

TM 870 059

AUTHOR Cronin, Joseph M.
TITLE The Cost of National and State Educational Assessments.
PUB DATE Aug 86
NOTE 17p.; One of 46 papers commissioned by the Study Group on the National Assessment of Student Achievement and cited in Appendix B to their final report "The Nation's Report Card" (TM 870 049). For other papers in this group, see TM 870 050-094.
PUB TYPE Viewpoints (120)

EDRS PRICE MF01/PC01 plus Postage.
DESCRIPTORS *Educational Assessment; *Educational Finance; Educational Testing; Elementary Secondary Education; Federal State Relationship; Financial Support; Information Needs; International Programs; National Programs; *National Surveys; *Program Costs; School Districts; State Programs; *State Surveys; *Testing Programs

IDENTIFIERS International Assn Evaluation Educ Achievement; *National Assessment of Educational Progress

ABSTRACT

The National Assessment of Educational Progress (NAEP) assessed three grade levels and 21-year-olds in the 1960's, at an annual cost of \$6 million. By the early 1980's, adults were no longer tested and funds had decreased to \$3.8 million. Other governmental departments, however, have also been funded for testing and a number of state testing programs have been implemented recently. New York State's comprehensive minimum competency testing program, like the NAEP's, cost \$3.8 million annually. Test construction is very costly. The Council of Chief State School Officers recommends testing during grades 5, 8, and 11. NAEP tests grades 3, 7, and 11. Local school district testing programs generally have goals that differ from those of the state programs. Local boards use tests for purposes of diagnosis, proficiency testing, or ranking. In the 1970's the international project, IEA (International Association for Evaluation of Educational Achievement) became an attention-commanding project; however United States government officials are not encouraged to fully participate in IEA planning. Coordination of IEA and NAEP procedures could reduce costs. Some state programs use NAEP items for purposes of comparison. Costs for each state's testing programs can range from \$105,000 to \$525,000, depending on the number of subjects tested per year; total national costs could range from \$5,250,000 to \$25,250,000. (GDC)

 * Reproductions supplied by EDRS are the best that can be made *
 * from the original document. *

ED279672

The Cost of National and State

Educational Assessments

Joseph M. Cronin

Massachusetts Higher Education

Assistance Corporation

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- This document has been reproduced as received from the person or organization originating it.
- Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

J. M. Cronin

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

Paper commissioned by

THE STUDY GROUP ON THE NATIONAL ASSESSMENT OF STUDENT ACHIEVEMENT

1986

2

BEST COPY AVAILABLE

THE COST OF NATIONAL AND STATE
EDUCATIONAL ASSESSMENTS

Dr. Joseph M. Cronin, President
Massachusetts Higher Education
Assistance Corporation

August 1986

THE COST OF NATIONAL AND STATE EDUCATIONAL ASSESSMENTS

What does it cost to prepare a report card on the education performance of the nation's school children? Could the costs be assumed by the federal government or shared between federal and state authorities? Could local school officials also pay for some appropriate share?

These simple, straightforward questions deserve a careful review. Unfortunately, the precise answers are much more complicated by this country's form of federal-state-local democracy than in the nation-states of the rest of the world. A review of these complexities should precede a final costing out of recommendations.

To evaluate what American children learn in school, it is possible to test and report on a representative sample, male and female, of various children in a given grade. Just as pollsters rely on a few thousand interviews to accurately predict election outcomes, it is fair and reasonable to identify 3,000 children and report their scores to present a national picture of student achievement.

If policy-makers want to report on what children in each state know, this requires a proper statistical sample for each of the fifty states and additional trust territories. For three grades of children (e.g., third, seventh, eleventh) this could require as many as 300,000 children tested.

If it is important to have every school represented in a report, consider that there are 90,000 schools in America. There are as many as 3 million pupils in each grade. No one in the United States has ever imposed the same test on all the schools, or on all of the children. However, most children are tested at one or more grades. All children in thirty or more states take at least one achievement test as required by the state legislature or state education agency. This is a relatively new practice in many states, and states vary enormously in what they are willing to spend per test or per pupil.

Any group planning an evaluation of American education must wrestle with the question - how much do we want to know, how often, and how much evaluation can we afford?

The Costs of National Assessment

The U.S. government has for twenty years financed a program of periodic educational evaluations and assessment. The Congress and educational officials were concerned about the impact of federal educational programs but even more so about what American children in general know about mathematics, science and citizenship as well as the basic reading and writing skills. Thus was created the National Assessment of Educational Progress (NAEP).

During the 1960's the U.S. Government spent as much as six million dollars a year for NAEP assessment of children in three grades (age levels) in addition to the educational attainment level of twenty-one year olds. It seemed important to know how much learning or literacy had been absorbed by all Americans whether they had stayed in school or not. By the end of the 1970's and early 1980's the government reduced the funds for NAEP to \$3.8 million per year. Evaluation of children in school remained a priority. The assessment of young adults, or of older children not in school, fell by the wayside for lack of funds. If the \$6 million of the late 1960's were adjusted for inflation, a full program would cost almost \$13 million in 1986, not \$3.8 million.

The nation does care about adult literacy, about employability of older youth, and about the qualifications of those persons who will join the military. Funds have been appropriated to pay for surveys of literacy and knowledge levels of teenagers and young adults. The Department of Defense, Labor and Commerce needs to know about manpower trends and problem areas. Although the National Assessment of Educational Progress has not had funds to do this work, the U.S. government has spent millions through other programs or departments. For example:

- o A recent survey of literacy levels of young adults ages 21-25 authorized by the Congress, cost \$1.1 million.
- o The National Science Foundation conducted surveys and spent \$1 million on the development of international assessments of math and science, and another \$800,000 on other math and science indicators or special tests.

What is important to note is that the National Assessment of Educational Progress budget has fallen behind inflation, and purchases substantially less information than required in the 1960's. NAEP now assesses the learning of children in-school only rather than its original commitment to evaluate all young people's learning through age twenty-one. Therefore, NAEP is only a portion of what the federal government spends on educational evaluation. Also, other federal agencies (Department of Defense, National Science Foundation) have spent their own funds to evaluate educational effectiveness. Other Department of Education offices spend as much as six or eight million dollars a year on several educational evaluations (other than specific program evaluations such as Chapter I or Vocational Education).

The State Assessment Programs

Each state bears a constitutional responsibility to provide for the education of those children in residence. No federal law, no national curriculum, indeed no national consensus determines what is either taught or tested or when a student should be evaluated.

However, twenty years ago very few states believed in requiring state achievement tests and in spending money for evaluations. Since about 1975, the situation has changed dramatically, with governors and state legislators demanding more and better information about schools:

- o What have the students learned and, in some cases, can they pass a test of minimal competence for graduation to or from the ninth grade, from senior high school?
- o Are the teachers and their courses teaching students the skills and knowledge that they are expected to teach?
- o What do the teachers know, and can they pass a test of basic subject matter knowledge?

Each state varies tremendously in the kind of testing required, the frequency of testing, grade levels tested, uses of the testing information and the amount of state funds allocated for testing.

New York State has the largest and one of the most comprehensive approaches to testing, requiring the following evaluation programs:

- o Regents Examinations in 21 senior high school subjects required as eligibility for a Regents diploma;
- o Regents Competency Tests in reading, writing, and math required for local diplomas;
- o Pupil Evaluation Program (PEP), annual tests required in reading and mathematics in grades 3 and 6 (writing test in grade 5), mainly to identify remedial needs;
- o Preliminary Competency Test - grade 8 or 9 - required for below-average P.E.P. score students and their remediations.

The total cost of these programs is \$3.8 million a year or equal to that of the federal NAEP for the entire nation.

North Carolina spends \$1.4 million on the implementation of its testing programs, or about \$2 per student.

Usually the initial development of a test costs a great deal of money. Texas appropriated \$17 million to devise a test (actually 100 specific tests) on what teachers should know before they teach. The cost of updating a test is also significant since it is not possible to use the same test items over and over again.

The tests required vary from achievement tests developed by national testing companies with national norm references, to state-developed tests for graduation. Some states select the test; others allow local districts to select whatever test they feel appropriate for their students.

Several states decide which grade levels must be tested:

Alabama - Achievement tests in grades 1, 2, 4, 5, 7, 8 and 10, with a graduate exam in grade 11, and aptitude and vocational interest surveys in grade 9;

Arkansas - Criterion referenced tests in 3, 6, and 8, with achievement tests in 4, 7 and 10;

Michigan - Mathematics and reading tests in grades 4, 7 and 10 (science soon to be added);

New Mexico - Achievement tests in grades, 3, 5 and 8, and a high school proficiency exam in grade 10 (retake in grade 11 or 12, if needed);

South Dakota - Achievement tests in grades 4, 8 and 11;

Virginia - Achievement tests in grades 4, 8 and 11, and competency tests.

Some states sample educational performance on a portion of students, as does Minnesota, Utah, Wisconsin (beginning in 1989, all third graders will be tested) and Wyoming.

Some states test all students at one grade - Idaho in grade 8; New Jersey, grade 9 (grades 10-12 if the student hasn't passed); Tennessee in grade 9 (with each student allowed five chances to pass).

Two states have no formal state-wide test program - Iowa, Vermont. Ohio requires some local tests. New York State, as already noted, has one or more tests for grades 3, 5, 6, 8, 9, 10, 11 and 12.

The Council of Chief State School Officers has established a state education assessment center to discuss how to measure student achievement in a comparable way among the states. The chief state school officers agree that there should be a core of achievement data common to each state, measured in a common set of grades, skill and content areas, and both basic and higher order skills. They agree that reading, math, science, English and social studies - and a writing sample - should be evaluated. They affirm that achievement data should be used along with other indicators of success.

The CCSSO Steering Committee prefers testing during grades 5, 8 and 11, the usual end of a unit of school. NAEP uses grades 3, 7 and 11. At least there is agreement on comparisons at three grades, at least three years apart, so that the costs would be roughly the same for national samples of three grades - one elementary, one junior high or middle, one in senior high. These data could supplement a state's own testing program, could become part of or the total statewide testing program in each state.

Local Assessment Programs

Do local school boards follow the same procedures as states? Rarely. Local school district assessments are often very different in purpose than that of state programs or the National Assessment of Educational Progress;

1. Local school officials use tests to diagnose learning difficulties or to search for shortcomings in the curriculum;
2. Tests are used for entry into specialized courses such as college preparatory or technical programs, for promotions or graduations or awards.
3. Tests are used to rank schools or students for certain purposes on a city-wide basis.

Generally in local school districts all students are tested in certain grades. No one is excluded except the severely retarded or handicapped. The sampling techniques used in NAEP are not applicable to most local assessments. Some cities contrast their pupils' performance with those of other cities using the same test, such as the Metropolitan Test of Educational Achievement. Other districts compare the current year test data with that of previous years.

The International Assessments

During the 1970's the International Assessment of Education (IAE) became a project commanding the attention of educators and policy-makers around the world. The achievements of U.S. students in math, science and other subjects have been compared with British, French, Russian, Japanese and students from other nations. Next to comparisons state by state, or 1986 with 1975 or 1965, the performance of U.S. students versus those of other major countries, has commanded considerable attention.

U.S. policy at present does not encourage Department of Education officials to fully participate in the planning of international assessments. U.S. officials are not provided travel money to take part in international assessment conferences in the U.S. or abroad.

Most of the seed money for international comparisons, at least in mathematics and science, comes from the National Science Foundation, an independent agency, and is awarded to educational testing experts based at universities. A long-term commitment of funds to complete a four or five year project has not been made.

If the U.S. government becomes actively committed to periodic international assessments, then one of two action policies should be considered to integrate national assessment data bases with international data exchanges:

- o The International Assessment of Education could be persuaded to use some NAEP procedures, items and formats, or
- o NAEP could agree to use enough IEA formats and procedures to link the results.

The cost of participating more fully in the international discussions and in sharing evaluation strategies, formats, test procedures and items, could cost \$1,000,000 a year. Adopting NAEP procedures and items would, of course, be economical. However, in the interest of international comparisons, it is possible that the United States may need to make additional funds available for a standard international assessment of education that is linked to NAEP.

Current Federal-State Cooperation in Assessments

One of the virtues of the National Assessment of Educational Progress is that, from the start, the NAEP provided information on student learning by region, by gender, by race, by urban-rural-suburban locale but not by state, school district or local school. This protected state and local communities and school faculties from odious comparisons. Such a virtue, commonly agreed upon by professional educators and policy-makers in the 1960's - has become a shortcoming in the 1980's. State officials now want to know:

- o How are our schools performing in contrast with those of other states?
- o How are types of school districts in one state performing in comparison with similar districts elsewhere?
- o How is the school in X community performing in comparison with the school in Y community?

Furthermore, if the state is a poor-achiever educationally or if there are many low-achieving schools, should the state develop a strategy and appropriate funds to remedy those deficiencies? Since many low-scoring schools tend to have low-income families, should that strategy include health, nutrition, early childhood and other services?

Several of the major national surveys of education attainment have offered to the states the chance to piggy-back on a national sampling. For example, the National Opinion Research Center in Chicago in the late 1970's in sampling for the High School Class of 1980 - and Beyond offered states the option of purchasing the sample for their own analysis, and adding additional students to the survey if their sample wasn't statistically large enough to analyze as a state. This option is again available with the most recent study, National Education Longitudinal Study (NELS) 1988.

Also, the Educational Testing Service which now administers NAEP has offered states the chance to use NAEP items in a state. This practice actually began in the 1970's when the Education Commission of the States administered NAEP for the federal government. Several states have in fact adopted the NAEP in whole or in part. Wyoming uses NAEP items at a cost of about \$100,000. Georgia is one of eight states in the Southern Regional Education Board pilot program, but it has also chosen to use the full NAEP test program. The new Massachusetts Educational Assessment Program is modeled after NAEP and includes two-thirds of the items used by NAEP in 1986.

Other states have considered the use of NAEP but decided to search for less expensive test packages. ETS produces NAEP and develops the items for NAEP (also for the Scholastic Aptitude Test, achievement tests and various professional licenses), and is a highly professional organization. But, while ETS is non-profit, the test packages are costly, particularly in relation to the costs to states on a per-pupil basis for their own testing program.

The Cost of Assessment Programs

One of the most important determinants of cost is the frequency of testing in each subject area. How often must one test, and does each subject need to be assessed each year?

The most expensive option is to assess each subject every year - five subjects a year. No state and hardly any state official favors this option.

The least expensive is to test one subject or cluster of skills a year in each of three grades. Reading, writing and English would be combined as one cluster of skills, and mathematics as another cluster. This proposal has genuine support from a number of state educational experts.

NAEP's current model is to assess reading every two years along with two other skill areas (math and science, or writing and social studies). In other words, three skills are assessed every other year. No reports are offered in the off years. Another way to be used by NAEP is an educational assessment every year of at least two subjects: reading, writing/English in years 1, 3 and 5; math and science in the even years.

Another crucial cost variable is the number of students which must be tested. If policy-makers simply wish to sample the nation, male and female, urban, suburban, rural, by major regions and major racial groups (in three grades) - 20,000 students will suffice. If students are sampled by these categories within each of the fifty states, this requires 2,000 students per state, 6,000 students per state in the three grade levels. For fifty states this is 300,000 students a year. NAEP procedures require fifty minutes per student (not all subjects, not all questions).

A CCSSO conference of evaluation experts in March of 1986 costed out the major assessment options based on the experience of the Southern (SREB) states replicating the NAEP procedures:

<u>Number of Subjects, Frequency</u>	<u>Annual Cost Per State</u>	<u>Total for the Nation</u>
1. Every subject every year (highest)	\$525,000	\$25,250,000
2. One subject each year (lowest)	105,000	5,250,000
3. NAEP model - three subjects every two years	157,500 (315,000 over two)	7,850,000
4. Two subjects each year	210,000	10,500,000

The costs should be adjusted upwards to reflect the eight trust territories whose inhabitants are included in federal education programs as well.

NAEP in 1986 is allowed \$3.8 million a year or \$7.6 million every two years. Use of the NAEP model with a state-by-state sample would just about double the cost.

There are other costs which must be considered:

1. To test all 17 year olds both in and out of school would require the identification of samples in the community, including those youth who are not in school. This requirement could cost another million dollars a year.
2. The testing of higher order skills is an extremely difficult but important costing problem. Which higher order skills will be tested - creativity, scientific problem-solving? Should simulation equipment, including computers, be used in the testing procedure? Laboratory equipment? Essay answers? Review of the problem-solving procedures or invention? Several members of the National Academy of Sciences committee working on these issues feel that it is not possible to test higher order skills reliably using multiple choice tests. ETS has allocated \$30 million to the development of computer-based tests to test higher order skills, but this is a long-term project. The IEA tests include some items as will the science NAEP tests in the future. But the cost of writing and correcting such items are much higher than the usual test items.
3. The testing of writing can be as much as five to ten times as expensive as other tests such as multiple choice tests. Education Testing Service is planning ways to contain the costs. However, it is generally thought that two or more assessors should read and evaluate each essay or paragraph.
4. The assessment of 21 year-olds, an initial objective of NAEP, would cost an additional \$1-2 million each testing time. This age bracket is spread about in college, in the military, in the work-place, even some in jail. Yet age 21 is the traditional threshold to adult maturity. Testing this age cohort provides a chance to evaluate the literacy impact of community college, adult education, work-place education, G.E.D. programs, etc.

Costs of Local Participation in a National Assessment Program

Will local schools want national educational assessment data? Some will want this information for comparative purposes. However, many school districts do not sample; rather, they test all pupils - for diagnosis or to qualify students for grade promotion or for deciding who shall receive what diplomas.

There are 16,000 school districts in the nation and 87,000 schools. A minimal sample would require 800,000 students. Assessment of all students in three grades would require testing of more than ten million students.

The local sample (800,000 students) might cost the nation from \$12 million (for one subject, every year) to \$60 million (every subject, every year) or almost \$20 million a year for the NAEP model - assuming that the number of students sampled is 2-1/2 times that needed in the state sample.

Multiply these figures by 12 (for a total universe of three grades) - \$144 million for one subject a year to a high of \$720 million for evaluating every subject every year.

It is not clear that a national profile or state-by-state comparisons require this detail. The expense of local assessment is very substantial. Also, the purposes are often very different - school decision-making about pupils rather than legislative decisions about education programs and priorities.

Could States Share the Cost?

States have begun determining how to obtain comparable state data which is more meaningful than the college aptitude test data "report card" now released by The College Board and U.S. Secretary of Education each Fall.

The states affiliated with the Southern Regional Education Board (SREB) have agreed to provide \$25,000 per state to test one subject at one grade level. SREB and ETS together provided \$35,000 for the project per state. In 1986 the pilot project was expanded to two subjects at a total cost per state of \$55,000, suggesting some economies of scale when a second subject is added.

State officials have already demonstrated a willingness to join in cooperative ventures in educational policy-making and assessments. Increasingly, state policy-makers have agreed to the principle that educational quality and financing should be periodically.

States now pay from \$30,000 to \$50,000 a year to participate in the policy research, development and legislative projects of the Education Commission of the States. They pay an average of \$6,000 to be members of the Council of Chief State School Officers.

States might be willing to assume some of the costs of national assessment, especially the administrative costs of a cooperative federal-state evaluation, which include:

1. Selecting a sample of public (and private) school students;
2. Administering the test including training the supervisors;
3. Scoring the tests according to a uniform, centralized guide;
4. Data processing, analysis and reporting on a timely basis.

The above costs might amount to three to five dollars a student. The scoring of data is thought to be \$.80 to \$1.00 per student except for writing sample (\$2.00 to \$3.00 per student).

Most states would continue their own examinations in the Spring of each year. It is not realistic to think that North Carolina or New York will trade in their base line data, their Regents or graduation tests, for a new sampling by state. Many state tests are required by state laws. These practices will continue, with or without a redesign of NAEP.

It is useful to offer state (and local) agencies the option of acquiring additional NAEP items, a larger data base, and other analyses for special purposes. Many states see in NAEP assessments useful information for identifying and acting on state trends; this use of NAEP should be encouraged. State officials very much want to participate in the design, planning, financial discussions, and implementation of data-sharing and cost-sharing options.

Summary and Conclusions

1. The U.S. Department of Education has sponsored a National Assessment of Educational Progress for twenty years, but spends substantially less (\$3.8 million) in 1986 than in the late 1970's, especially when adjusted for inflation.
2. The National Science Foundation spends much more money on international math and science comparisons than the Department of Education, which has not yet made any long-term commitment to the International Educational Assessment.
3. One state, New York, spends as much on state testing and competency programs as the U.S. Department of Education spends for NAEP.
4. Although a majority of states require or offer tests, they differ widely in purpose, frequency, and grade level.
5. State education officials would be willing to share certain of the administrative responsibilities and costs of a cooperative national-state assessment program. Such a project might not replace state testing but a cooperative program would be useful to federal and state officials.
6. A federal-state evaluation could cost approximately twice as much as the existing NAEP programs, almost \$8 million a year or \$16 million every two years.
7. The most important cost variables are the frequency of testing various subjects. There is some consensus that three grade levels (in school) are enough but that testing out-of-school 17 and 21 year-olds is also useful. Add at least \$1 million for the 21 year-olds and \$1 million for international assessment.
8. Extending assessment to every school and school district would raise the costs very substantially. Local schools often use test data for their own purposes of diagnosing or promoting students within a school rather than assessing broad educational goals.

This paper is intended as a preliminary review of cost issues. Other assumptions, variables or strategies could raise or lower the costs dramatically. The issues of testing higher order skills deserve special attention, require additional test development, and may, as in the case of testing actual writing skills, cost much more than multiple-choice tests. The costing of a national assessment, or International/National/State Assessment must follow upon important decisions on the purposes, scope, frequency and types of assessments needed by policy-makers over the next five or more years.