DOCUMENT RESUME

ED 266 966 SE 046 478

AUTHOR Hilton, Thomas L.; And Others

TITLE Science Indicators from National Assessment and Other Sources. National Assessment of Educational Progress. Part 2: Analysis of National Assessment Results and Related Data. Final Report.

INSTITUTION Educational Testing Service, Princeton, N.J.

SPONS AGENCY National Inst. of Education (ED), Washington, DC.; National Science Foundation, Washington, D.C.
PUB DATE Oct 85
GRANT G-85-2006-P4
NOTE 57p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Advanced Placement; *Educational Assessment; Elementary Secondary Education; *Mathematics Education; *Science Education; Scores; *Test Interpretation; *Test Results
IDENTIFIERS *National Assessment of Educational Progress; National Teacher Examinations; *Science Indicators

ABSTRACT

This study investigated ways of making present and past results of the National Assessment of Educational Progress (NAEP) and other data sources maximally useful as science indicators and to recommend possible changes in instrumentation and procedures that may enhance future NAEP surveys for these purposes. Primary emphasis was given to reviewing published and unpublished reports that have resulted from the national assessment and the frequent special data collections which have also resulted from it, as well as on integrating the findings into tables and text of special relevance to the 1984 edition of "Science Indicators." In addition, the report files and test score files of the National Teacher Examinations (NTE) were examined for similar evidence and results that may confirm or disconfirm the NAEP results. Other data sources (such as the Educational Testing Service Advanced Placement Program) were also examined. Results and recommendations are reported and discussed separately for national assessment, national teacher examinations, advanced placement scores, and for other data sources. One recommendation (related to NAEP) is that all scales purporting to measuring the same concept should be on a common scale. Among the findings is that the NTE data filed proved to be of limited value as a source of science indicators. (JN)

FINAL REPORT

Science Indicators from National Assessment and Other Sources

National Assessment of Educational Progress:
Part 2: Analysis of National Assessment Results and Related Data - Grant #G-85-2006-P4

Submitted to National Institute of Education

Thomas L. Hilton, Project Director
Penelope Engel
Janet R. Johnson
Barbara Pitcher

Educational Testing Service
October 1985

The research reported herein was performed as part of a grant from the US Department of Education. Grantees undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Department of Education position or policy.

PREFACE

The work reported here was conducted as part of a grant to Educational Testing Service by the National Institute of Education (NIE) for analysis of data from the files of National Assessment of Educational Progress (NAEP) and other sources of data relevant to the status of science students and science teachers in the United States. Funds for these analyses were provided by the National Science Foundation (NSF). The authors of this report are indebted to Richard M. Berry, Program Director, Studies and Analyses, NSF, for taking the initiative in requesting the work, for providing support for the work, and for expert advice to us in conducting the work. Lawrence M. Rudner and Gerald Kulm of NIE were helpful in managing the disbursement of the funds by NIE and in monitoring the preparation of project reports.

At ETS, a number of ETS staff members assisted the authors in locating and processing the required data and library resources, including Ina S. Mullis, Associate Director of NAEP; Lucy Mitchell, Manager of Systems, SHEP; Edward J. Masonis, Program Administrator, SHEP; and Carl H. Haag, Program Administrator, College Board Program Services. Albert E. Beaton served as technical advisor in regard to the NAEP data base. Irene Smith managed the preparation of materials and reports and Faith Thompson assisted in the typing.

Table of Contents

Science Indicators from National Assessment and Others Sources

The preparation of secondary school graduates in the area of mathematics and science currently is the focus of widespread concern in this country. Several educational commissions have pointed to lack of attainment in this area, and the National Science Board's Commission on Precollege Education in Mathematics, Science and Technology has proposed as a primary goal the strengthening of precollege science and mathematics education.[1] In line with its proposal, NSF will devote a full chapter of the 1984 edition of Science Indicators to trends in the preparation of secondary students in mathematics and science. The purpose of the present study was to investigate ways of making the present and past results of NAEP and other data sources maximally useful as science indicators and to recommend possible changes in instrumentation and procedures that may enhance future NAEP surveys for these purposes.

The primary emphasis of the work was on reviewing the large number of reports, both published and unpublished, that have resulted from National Assessment and the frequent special data collections which have resulted from it, as well as on integrating the findings into tables and text of special relevance to the forthcoming 1984 Science Indicators.

---

[1]Report of the National Science Board's Commission on Precollege Education in Mathematics, Science and Technology, Educating Americans for the 21st Century, National Science Foundation, Washington, DC, 20550, 1983, page 5.

Secondly, in recognition of the importance of competent instruction in mathematics and science at the high school level and, thus, the importance of teacher characteristics, the report files and the test score files of the National Teacher Examinations were examined for similar evidence and results that may confirm or disconfirm the NAEP results.

Method

National Assessment Steps

"National Assessment has measured science knowledge and skills of 9-, 13- and 17-year-olds in 1969-70, 1972-73 and 1976-77." So states the NAEP report, Three National Assessments of Science: Changes in Achievement 1969-77. The report goes on to explore the collected data on science knowledge and skills and to provide a context for the trends in precollege scientific achievement and their relative ascendency or decline among students in public and private schools across the nation. In addition, a more recent report, Images of Science from the Science Assessment and Research Project of the University of Minnesota, summarizes the results from the 1981-82 National Assessment in Science. These reports are only two of more than fifty NAEP official publications and NAEP-related publications dealing exclusively with science. The number of similar publications concerned with mathematics is equally extensive.

The purpose of NAEP is to provide information for governmental and educational policy makers. It serves as the nation's report card and has the responsibility to determine what young Americans know and are capable of doing. In addition, the authorizing legislation calls for the periodic reporting of data on changes in the knowledge and skills of such students over time.

The rich data base of NAEP began to be amassed in 1969 and at present contains records of the educational achievement of over 1,000,000 students in several curriculum areas, of which mathematics and science are two.

The NAEP collection of representative national data on educational competence in mathematics and science is an invaluable resource in the development of precollege science indicators. Through matrix sampling, NAEP has always had the capacity for comprehensive coverage of subject matter. In addition, the utilization of a deeply stratified, multistage probability sample design ensures that participating students are selected in such a way that they represent the national population of 9-, 13-, and 17-year-olds.

The current administration of NAEP has dramatically increased the number or background and attitude items responded to by each student. However, there is still a wealth of data from prior assessments which further break down the national sample by:

- o Region (Northeast, Southeast, Central, and West)
- o Type of community (advantaged-urban, disadvantaged-urban, and extreme-rural)
- o Size of community (big city, fringes around big city, medium city, and small places)
- o Grade in school
- o Sex (males, females)
- o Race (Blacks, Whites)
- o Parental education (less than high school graduation, at least one parent who graduated from high school, at least one parent who had some post-high school education)

The following steps were undertaken.

1. The extensive internal collection of NAEP materials were searched for all relevant documents, unpublished papers, and computer files.
2. Organizations and individuals who have been using the NAEP public-use data tapes for independent research projects were called. The purpose of these contacts was to identify relevant results and reports that may not have been disseminated to date.
3. A detailed search of ERIC and other computerized files that may have included relevant abstracts and material was conducted. In conducting the literature searches, we looked not only for trend results based on NAEP data but also for any other trend results based on data that were relevant to the precollege preparation of high school students in science and mathematics.
4. With the aid of ETS staff members associated with NAEP relevant material was identified.
5. Limited analyses of source data were performed to obtain cross-tabulations and other desired descriptive statistics that were not in the output and reports examined.
6. Meetings were held in Washington with NSF staff to present the preliminary results and to discuss possible additional steps, to interpret the results, and to discuss inconsistencies or discrepancies among them.

National Teacher Examinations Steps

In view of the critical importance of competent instruction in the sciences and mathematics at the secondary school level, other ETS sources of

data on the characteristics and qualifications of science and math teachers were examined. An obvious first step was to conduct descriptive analyses from data in the NTE Programs (formerly known as the National Teacher Examinations).

A recent GAO report discussed how little is known about the achievement of math and science teachers and specifically asked: "How much can we learn from the National Teacher Examinations and similar state tests...?" Because of distributional limitations and confidentiality considerations explained below, neither NTE score nor volume data have been applied for policy purposes beyond the individual state level. This exploratory study was probably the first examination of the possible usefulness of NTE data in research on the characteristics and competence of teachers in the U.S.

The annual volume of NTE test takers is approximately 83,000 nationally. Candidates, however, are not equally distributed across the United States. At present, about 30 states require, encourage, or offer as an alternative to course requirements at least some component of these examinations. (See attached chart, in appendix.) However, the number of states informally identified as "NTE States" is much smaller. Those which require the NTE Core Battery and/or Specialty Area Tests or use these tests as a significant alternative to the approved program approach are Arkansas, California, Louisiana, Mississippi, North Carolina, South Carolina, and Virginia. In addition to these, New Mexico, Tennessee, and West Virginia are major users of NTE Program Tests. These ten states obviously do not constitute a nationally representative sample—a fact that is noted below.

In addition to the very uneven use and distribution patterns, policy analysis beyond the state level has been precluded in the past because of a policy of confidentiality on the part of ETS. ETS will not publicly release

NTE test score data by state without client permission. Even volume data have proven difficult to interpret over time (both within states and aggregated) because of the frequent changes in certification options and the varying number of user states from year to year.

The NTE data include a five-year history file that contains personal variables for each candidate, including race, sex, age, undergraduate grade point average (UGPA), undergraduate and/or graduate major field, and of course, the particular NTE test or tests taken. Of the 27 Specialty Area Tests (also in appendix), data from the Mathematics Test and two science tests (the Chemistry, Physics, and General Science Test and the Biology and General Science Test) would be of most interest to the proposed study.

An ETS statistical coordinator familiar with NTE data files and a professional associate examined options for "mining" the Specialty Area Test data for indicators of both quality and quantity of prospective teachers of mathematics and science. Findings of possible value regarding quantity were the number of candidates taking the math test and each of the science area tests as a percentage of those taking all Specialty Area Tests over a ten-year period. Comparative distributions of these test takers by race and sex over five years were considered. The feasibility and value of comparing these volume trends by states were examined, always keeping in mind the obligation of confidentiality.

Regarding the quality of prospective teachers who have taken NTE Tests, we compared the mean NTE Core Battery or Common Examinations scores of those taking the Specialty Area Tests in math and the sciences with those taking other Area Tests in recent years. It should be emphasized, however, that the nature of the population taking the NTE, the importance of confidentiality,

the possible sensitivity of certain results (e.g., racial and regional differences), and the changing character of the tests in recent years place limits on use of the data.

Other Steps

A number of other possible data sources were examined including data from the:

- o ETS Secondary School Admissions Test,
- o Iowa Test of Educational Development,
- o College Board Preliminary Scholastic Aptitude Test (PSAT),
- o Armed Services Vocational Aptitude Battery,
- o ETS Advanced Placement Program,
- o CEEB Achievement Tests, and
- o the High School Equivalency Test.

The use of scores from the Scholastic Aptitude Test and the Graduate Record Examinations had already been considered and acted upon as part of previous contracts between ETS and NSF. With one exception the possible use of each of the additional data sources listed above was rejected, usually because the sample of students taking the tests were not sufficiently representative of all students in the United States or, especially in the case of the PSAT, the results would unnecessarily duplicate the results obtained earlier for the SAT.

Advanced Placement. The one exception was data from the ETS Advanced Placement Program. Although obtained from a small highly self-selected subpopulation, the results from this program were judged to be of sufficient interest to justify detailed examination, provided that caution is exercised in interpreting the results. These results are described shortly.

Results

National Assessment

The primary products of this line of work are the sections of the 1984 Science Indicators that were prepared by Richard Berry of NSF. This publication will be available shortly. The major task as far as NAEP data were concerned was to update the NAEP results previously published concerning national trends in science and mathematics scores.

For trends in science, the authors relied primarily on Three Assessments of Science, 1969-77 (Report No. 08-S-21) for past trend data and on Images of Science (Hueftle, Rakow, & Welch, 1983) for current results. To make sure the later survey was designed and conducted in a way that would provide comparable results, the Images of Science was reviewed by a consultant formerly associated with NAEP. The conclusion of this consultant was that the results of this fourth in the series of national assessments in science met conventional standards in survey technology and that the results could be compared with the results of previous national surveys. His report, a copy of which can be obtained from T. L. Hilton, noted the following:

- More technical information in regard to the procedures used to obtain the data in Images of Science would have been desirable, perhaps in the form of a technical appendix or supplement.
- Appropriate adjustments were made to correct for the fact that the Science Assessment and Research Project (SARP) sample included more larger schools than previous NAEP assessments.
- The SARP sample was approximately one-third the size of previous NAEP samples but the absence of standard errors of the various summary values made it difficult to assess the statistical significance of the results.

- o In the absence of the actual items used in the SARP assessment it was difficult to evaluate their comparability to previous items.

On balance, the consultant judged the SARP report to be of "good quality" and concluded that "if used with some caution [it could] serve well as the fourth point on the science trend line."

Subsequently, the present authors arranged for a review of the science items of the SARP assessment by a subject-matter specialist at ETS who judged them to be of good quality and adequate for their purpose.

Assembling the required NAEP data proceeded without serious problems and, as is evident from Mr. Berry's chapter in Science Indicators, resulted in a substantial addition to that volume. On the basis of this work, the authors would make the following recommendations:

1. That future reports based on NAEP data be documented in minute detail, probably in technical footnotes or appendices. This documentation should include details about exactly which items were used in reporting summary scores and, when possible, copies of the actual items used; details about any statistical operations performed on the items, and details about the sample on which the descriptive statistics were based including standard errors. Despite the obvious care that has been taken in the past in preparing NAEP data for publication, the authors occasionally were uncertain about some of the details.
2. That uniform procedures be adopted in regard to rounding the statistics reported. Some differences between statistics reported in one publication and, presumably, statistics based on the same data in other publications were attributable to different rounding procedures.

3. That manuscripts be proofread with unusual care; a few minor typographical errors were found.
4. That an annual or biennial cumulative index be disseminated, including all reports that are available, with detailed descriptions of each and, preferably, a summary of the salient findings of each. This seems to have been done occasionally but not on a regular basis.
5. That henceforth all scales purporting to measure the same concept be on a common scale. (Presumably this will now be done routinely from now on, through IRT equating.)
6. That uniform procedures be adopted for reporting standard errors. The authors occasionally were in doubt about whether small changes were statistically significant.
7. That each report routinely include information about when the items on which the results are based will be available to other researchers and how they should go about obtaining the items or scales.

National Teacher Examinations

As anticipated at the start of this study, the NTE data files proved to be of limited value as a source of science indicators, primarily for the reasons cited above. The NTE population is highly self-selected; whether a particular individual takes the NTE depends primarily on whether the school district or state in which the prospective teacher is applying for a position requires the NTE, and these district or state requirements change considerably from one year to the next, and differ markedly from one locale to another. Also, there are occasional special administrations of the test that contribute to change in the test population. Lastly, the test score files were not designed for the purpose of annual indicators meaning that considerable data processing is

necessary in order to identify subpopulations that might be reasonably stable from one year to the next.

Despite these known problems the authors considered several alternative ways of constructing useful annual indicators. The first was to focus on school districts that were known to have required the NTE of applicants for teaching positions in the district for a number of years in the recent past. The city of Chicago was an example. This alternative was rejected since the number proved to be too small, especially considering that our interest was only in the subsample of test takers who took the area tests in science and mathematics.

A second alternative focussed on candidates taking the Area Tests of interest and then adjusting these scores for annual fluctuations in the mean ability of the sample by means of the Core Battery scores, this was rejected because only a small fraction of the NTE population take both the Core Battery and any particular Specialty Area Test of interest.

A third alternative, which was pursued, focussed on test takers in a small subset of states which were known to have required the NTE consistently for at least the last five years. As described in the report in the appendix, these states were Arkansas, California, Illinois, Louisiana, Mississippi, New York, North Carolina, South Carolina, Tennessee, and West Virginia. Obviously, this alternative was not fully satisfactory since the sample of available states overrepresents states in the southern part of the United States. However, change in the mean scores of the test takers in these states was judged to be of interest even though the mean scores for these states were not necessarily representative of all the states.

As described in the report in the appendix, the mean scores on the Area Tests related to science and mathematics did change from 1979 through 1984 in nonrandom ways. However, there was no consistent pattern in the changes and the changes were not large. These results could be viewed as encouraging in that there was no evidence of an expected marked decline in the test scores of prospective math and science teachers. But in view of the possible bias in the results, the authors' advice is that the results be interpreted with caution. No recommendation was made that the NTE test scores be used as science indicators.

Advanced Placement Scores

As mentioned above, the third possible source of science indicators that was examined with some care was the data files of the Advanced Placement (AP) Program. AP courses and examinations are given at over 20% of American secondary schools to 15% to 20% of their most able college-bound students. Participants are by definition doing college level work and are high achievers and generally highly motivated. Each examination (with the exception of Studio Art) includes both an objective, multiple choice section and a free response or essay portion. In each subject area, a group of teachers grades the free reponse part of the examination. These teachers, from participating schools and colleges across the country, are organized and directed in their grading by a chief reader, who typically is a college professor. Final grades, based on the student's entire examinations (with free response and multiple choice questions appropriately weighted), are reported on a 5 point scale: 5- extremely well qualified, 4- well qualified, 3- qualified, 2- possibly qualified, 1- no recommendation. Participating colleges normally honor grades of 3 or higher.

As described in the report of this work, which is included in the appendix, changes in the mean summary scores from one year to the next do not necessarily reflect reliable changes. The grade data are based on scores from both the multiple choice portion and the free response parts of the examinations but only the multiple choice portion is equated from one year to the next. Furthermore, cut scores vary from year to year for each examination, reflecting changes in levels of exam difficulty, and the chief readers often attempt to maintain similar percentages of students receiving each of the scores from 1 to 5, from year to year. Thus, changes in mean scores during the last ten years were judged to be unsuitable as indicators of changes in the achievement of students taking the AP examinations.

Examination of the number of students taking the AP exams over the last ten years indicated that interest in biology, chemistry, and physics—relative to other subjects—remained very much at the same level over the last ten years while relative interest in mathematics declined dramatically. These volumes, however, may have been affected by the availability of teachers in mathematics and science and, since there is no feasible way of ascertaining the extent to which this factor may have influenced volumes, the authors' recommendation was that these data be used as science indicators with caution.

Other Data Sources

Of a number of other possible data sources examined, one was considered promising. This is the number of math and science courses that secondary students enroll in during their last three years of high school attendance. To investigate the validity of this indicator, the preliminary results of the study of Excellence in High School Education, then underway at ETS, were examined. This study was based on longitudinal data for the High School and

Beyond subjects who were surveyed as sophomores in 1980 and as seniors in 1982. Identical short tests of achievement in mathematics and science were given to the approximately 23,000 students who participated at both times. Transcripts of the high school academic work were obtained for one-half of the subjects after they were graduated. From the transcripts, counts were obtained of the total number of mathematics courses and the total number of science courses in which each student had enrolled. Two outcome measures were of interest: the test scores of the students in their senior year, and the residual gain in test scores from the sophomore to the senior year. (The latter measure is the gain in test score adjusted for difference among the subjects as sophomores.)

The product moment correlations between the number of courses enrolled in and the two outcome measures were as follows:

	Senior test score	Residual gain
No. of math courses	.55	.29
No. of science courses	.49	.23

These correlations—for a large national sample of high school students—point to a strong relationship between enrollment in high school math and science courses and the measured achievement of the subjects. The authors concluded that these correlations, along with similar findings in the research literature, were sufficient evidence of the validity of the number of science and math courses as indicators of probable achievement in these areas during high school, and recommended that the statistics be given serious consideration as science indicators even though data on course enrollments are available only from quite widely spaced studies such as the ETS Growth Study

(1963-69), the 1972 National Longitudinal Study (1972 to present), the Department of Labor National Longitudinal Surveys (1979 to present), and HS&B (1980 to present).

References

Hueftle, S. J., Rakow, S. J., & Welch, W. W. (1983). Images of science. Minneapolis, MN: Minnesota Research and Evaluation Center.

National Center for Education Statistics. (1978). Three national assessments of science: Changes in achievement, 1969-77. Washington, DC: U. S. Government Printing Office.

National Center for Education Statistics. (1979). Three assessments of science: 1969-77: Technical summary. Washington, DC: U. S. Government Printing Office.

National Science Board's Commission on Precollege Education in Mathematics, Science and Technology. (1983). Educating Americans for the 21st Century. Washington, DC: National Science Foundation.

APPENDICES

A - NTE Programs Users

B - Means and Standard Deviations of Scores on the Common Examinations, and Correlations (r_{xy}) Between Area and Common Examinations Scores for Examinees with Composite Scores

C - Memorandum by Barbara Pitcher on NTE Programs Data for NSF, September 11, 1984

D - Advanced Placement Analyses

APPENDIX A

NTE Programs Users

NTE PROGRAMS USERS

	Core Battery	Most Area Tests	ASHA Tests	PPST	CBEST	Commons
Alabama						
Alaska						
Arizona						
Arkansas	Y	Y				
California	X	X	Y		Y	
Colorado	X					
Connecticut			Y			
Delaware			Y	Y		
Florida						
Georgia			Y			
Hawaii			Y			
Idaho						
Illinois						
Indiana						
Iowa			Y			
Kansas				Y		
Kentucky	Z	Z	Y			
Louisiana	Y	Y	Y			
Maine	O	O				
Maryland						
Massachusetts						
Michigan						
Minnesota						
Mississippi		Y	Y			Y
Missouri						
Montana			Y			
Nebraska		X				
Nevada			Y			
New Hampshire	X	X				
New Jersey						
New Mexico	Y					
New York	Z		Y			
North Carolina	O	Y	Y			Y
North Dakota						
Ohio			Y			
Oklahoma						
Oregon			Y			
Pennsylvania	O	O				
Rhode Island			Y			
South Carolina		Y				Y
South Dakota						
Tennessee	Y		Y			
Texas				Y		
Utah			Y			
Vermont						
Virginia	Y	Y				
Washington						
West Virginia	X	X				
Wisconsin						
Wyoming						

Y Required
O Encouraged
X Alternative to course requirements
Z Completed study, currently under consideration for 1984

/mb
4/6/83

APPENDIX B

Means and Standard Deviations of Scores on the Common Examinations, and Correlations (r_{xy}) Between Area and Common Examinations Scores for Examinees with Composite Scores

Supplement to leaflet B1, Interpreting National Teacher Examinations Scores

TABLE 6: Means and Standard Deviations of Scores on the Common Examinations, and Correlations (r_{xy}) Between Area and Common Examinations Scores for Examinees with Composite Scores*

Area Examination	Number of Examinees	Professional Education			Written English Expression			Social Studies, Literature, and the Fine Arts			Science and Mathematics			Weighted Common Examinations Total		
		Mean	S.D.	r_{xy}	Mean	S.D.	r_{xy}	Mean	S.D.	r_{xy}	Mean	S.D.	r_{xy}	Mean	S.D.	r_{xy}
Examinees with Less Than Master's																
Art Education	2,680	55.2	9.8	.726	56.2	10.4	.618	59.3	9.5	.784	57.1	9.3	.648	568	86	.803
Biology and General Science	2,924	57.4	10.6	.778	58.0	10.6	.655	57.8	10.2	.758	66.2	10.7	.849	598	94	.869
Business Education	3,717	50.9	10.5	.852	55.0	10.4	.735	49.6	9.4	.720	52.7	10.0	.755	515	90	.874
Chemistry, Physics, and General Science	519	60.2	10.4	.735	60.8	9.8	.632	61.1	9.8	.715	72.4	10.1	.801	635	91	.823
Early Childhood Education	14,942	55.1	11.0	.870	54.5	10.8	.740	53.1	10.2	.742	53.6	10.1	.746	542	96	.878
Education in the Elementary School	27,638	55.6	11.2	.878	55.1	11.0	.751	54.1	10.4	.737	55.2	10.7	.773	551	99	.888
Education of the Mentally Retarded	7,742	56.9	10.9	.858	54.9	10.5	.703	54.0	10.0	.707	54.8	10.2	.715	555	94	.854
English Language and Literature	5,311	59.8	10.2	.802	64.3	10.0	.754	63.0	10.3	.851	58.7	10.2	.731	608	91	.884
French	504	59.9	9.4	.469	65.0	9.5	.476	63.2	9.6	.511	60.5	9.0	.360	614	82	.516
Home Economics Education	2,651	54.0	10.8	.874	54.9	10.5	.742	52.4	10.0	.761	56.1	10.2	.773	543	95	.891
Industrial Arts Education	1,242	49.3	11.3	.808	49.3	10.3	.682	51.8	10.8	.728	58.3	11.9	.840	522	103	.860
Introduction to the Teaching of Reading	325	59.8	11.6	.874	59.8	11.3	.739	60.5	12.1	.722	57.7	11.8	.727	595	107	.860
Mathematics	3,052	58.4	10.6	.661	60.1	10.5	.603	57.2	10.4	.599	66.7	10.4	.708	604	94	.722
Music Education	4,158	55.8	9.7	.723	59.1	9.8	.639	58.2	8.9	.708	59.8	9.6	.684	578	84	.795
Physical Education	9,378	50.0	10.5	.838	50.4	9.9	.697	50.2	9.0	.702	55.0	10.0	.743	514	88	.867
Social Studies	6,578	56.3	11.0	.794	56.7	10.8	.670	60.8	11.1	.831	57.7	10.7	.719	578	98	.859
Spanish	1,231	52.5	11.3	.193	53.7	13.1	.106	54.4	11.5	.196	53.3	10.7	.160	533	105	.191
Speech-Communication and Theatre	761	56.0	10.1	.800	58.4	10.2	.694	59.2	10.2	.776	56.7	10.1	.769	572	90	.856
Examinees with Master's or Doctorate																
Educational Administration and Supervision	874	62.5	11.0	.872	58.1	11.2	.718	60.2	11.9	.744	59.2	11.1	.681	607	100	.876
Guidance Counselor	351	61.9	10.7	.822	59.9	11.1	.719	60.2	11.1	.757	58.2	10.1	.680	604	97	.836
Media Specialist--Library and A-V Services	239	63.0	10.9	.818	64.0	11.0	.702	66.4	11.2	.738	59.3	10.7	.638	630	97	.834
Reading Specialist	258	66.5	10.0	.845	63.4	10.6	.730	64.5	11.0	.712	58.9	10.6	.686	638	94	.846
Speech Pathology	281	63.8	8.2	.653	63.9	9.1	.563	61.2	8.8	.493	61.0	8.9	.588	625	73	.689

*Based on the performance of examinees tested in the NTE program between November 1, 1976, and August 31, 1979. These are the same groups of examinees whose scores were used as the basis for Tables 4 and 5 in the published leaflet to which this table is a supplement. These statistics, because they are all based on the Common Examinations, may be compared across areas; however, it should be noted that the groups of examinees whose data were used for the table may not be representative of the people who typically teach in those areas.

Supplement to leaflet D1, Interpreting National Teacher Examinations Scores

TABLE 6: Means and Standard Deviations of Scores on the Common Examinations, and Correlations (r_{xy}) Between Area and Common Examinations Scores for Examinees with Composite Scores*

		Professional Education			Written English Expression			Social Studies, Literature, and the Fine Arts			Science and Mathematics			Weighted Common Examinations Total		
Area Examination	Number of Examinees	Mean	S.D.	r_{xy}	Mean	S.D.	r_{xy}	Mean	S.D.	r_{xy}	Mean	S.D.	r_{xy}	Mean	S.D.	r_{xy}
Examinees with Less than Master's																
Art Education	2,274	54.9	10.2	.717	55.7	10.4	.649	58.5	10.0	.778	55.9	9.8	.662	562	90	.799
Biology and General Science	2,671	58.0	10.3	.750	58.3	10.5	.665	58.0	10.4	.750	58.4	10.1	.831	599	91	.855
Business Education	3,240	51.3	10.6	.849	54.6	10.4	.742	49.2	9.8	.735	52.2	10.1	.750	514	92	.876
Chemistry, Physics, and General Science	486	60.5	10.1	.656	60.9	10.8	.605	61.4	10.3	.672	70.3	10.0	.790	632	91	.769
Early Childhood Education	15,788	55.5	11.0	.875	54.4	10.8	.740	52.4	10.4	.733	53.2	10.2	.761	541	96	.883
Education in the Elementary School	24,613	55.8	11.1	.872	54.8	10.9	.745	53.5	10.7	.732	54.6	10.7	.775	549	99	.883
Education of the Mentally Retarded	7,428	56.7	11.0	.844	54.3	10.6	.696	52.7	10.3	.703	53.8	10.4	.728	548	96	.848
English Language and Literature	4,589	60.4	9.9	.761	64.4	9.9	.743	63.7	10.4	.845	58.8	10.1	.714	612	90	.862
French	442	59.9	10.0	.512	65.6	10.1	.498	63.3	10.5	.551	59.9	9.6	.486	614	90	.572
Home Economics Education	2,230	54.2	10.7	.857	54.3	10.6	.727	51.6	10.0	.740	55.1	10.1	.772	538	93	.883
Industrial Arts Education	1,084	49.7	11.6	.786	49.2	10.3	.661	51.6	11.0	.739	57.6	11.8	.837	521	104	.850
Introduction to the Teaching of Reading	386	58.7	11.7	.868	58.4	11.6	.744	57.5	12.6	.731	56.8	11.4	.739	579	109	.861
Mathematics	2,548	58.3	10.4	.614	60.0	10.5	.587	57.1	10.5	.562	65.3	9.9	.672	600	92	.684
Music Education	4,150	55.8	9.8	.691	58.9	10.2	.641	58.1	9.6	.708	58.6	9.9	.696	574	87	.782
Physical Education	9,480	50.2	10.6	.838	49.8	9.8	.685	49.3	9.3	.694	54.1	10.0	.746	509	89	.865
Social Studies	5,577	56.7	10.7	.788	56.8	10.9	.688	61.4	11.2	.834	57.4	10.5	.737	581	97	.870
Spanish	733	54.8	11.5	.193	57.0	13.4	.109	56.9	12.1	.218	54.2	11.3	.157	554	108	.195
Speech-Communication and Theatre	616	56.0	9.9	.797	58.5	10.4	.712	58.4	10.5	.754	55.9	9.7	.691	568	89	.850
Examinees with Master's or Doctorate																
Educational Administration and Supervision	918	61.8	10.4	.850	56.9	10.8	.690	58.8	11.4	.723	58.5	10.8	.675	597	95	.864
Guidance Counselor	388	61.2	10.3	.820	58.9	11.5	.685	59.9	11.7	.735	56.9	10.6	.690	596	98	.831
Media Specialist--Library and A-V Services	203	61.9	11.0	.847	64.3	11.8	.770	66.1	11.7	.758	59.9	11.0	.692	627	102	.856
Reading Specialist	372	60.2	9.2	.822	63.1	10.3	.728	64.1	10.5	.672	59.7	9.5	.620	638	85	.826
Speech Pathology	263	63.5	7.8	.585	63.3	9.3	.547	61.5	10.1	.466	60.1	8.5	.569	621	74	.632

*Based on the performance of examinees tested in the NTE program between November 1, 1978 and June 30, 1981. These are the same groups of examinees whose scores were used as the basis for Tables 4 and 5 in the published leaflet to which this table is a supplement. These statistics, because they are all based on the Common Examinations, may be compared across areas; however, it should be noted that the groups of examinees whose data were used for the table may not be representative of the people who typically teach in those areas.

APPENDIX C

Memorandum by Barbara Pitcher on NTE
Programs Data for NSF, September 11, 1984

Memorandum for: TOM HILTON

cc: Robert Altman
Penelope Engel
Jane Faggen
Marlene Goodison
Catherine Havrilesky
Alice Irby
Ed Masonis
Carolyn Massad
Craig Mills
Lucy Mitchell
Nancy Petersen
Dawn Robinson
Janice Scheuneman
Billie Slaughter
Frances Swineford

Subject: NTE Programs Data for NSF

Date: September 11, 1984

From: Barbara Pitcher BP
20-P, Ext. 5967

Reference: My memorandum to Lucy Mitchell, on NTE Programs Data for NSF, issued August 15, 1984

The work described in the referenced memorandum has been completed and we have used the resulting tapes to run counts, select samples and obtain means and standard deviations as requested.

Enclosed are tables showing the numbers of examinees, means and standard deviations for all nine of the tests we used.[1] They are:

Biology and General Science (03 BGS)
Chemistry, Physics and General Science (07 CPS)
English Language and Literature (04 ELL)
Mathematics (06 MAT)
Social Studies (08 SS)
Business Education (10 BE)
Early Childhood Education (02 ECE)
Education in the Elementary School (01 EES)
Physical Education (09 PE)

[1] NSF was primarily interested in science and mathematics tests (BGS, CPS and MAT). English (ELL) and Social Studies (SS) were added for contrast in other subject areas. The other four tests were also included in the computer run because they are relatively large-volume tests of particular interest to NTE Programs staff. There are 16 additional Specialty Area tests.

As you know, the samples used for this analysis were selected from NTE Programs' files, as described in the referenced memorandum. That is, a year was considered to run from July-June (except for 1978-79 since November 1, 1978 is the oldest test date currently on the history file).[2] Examinees with multiple scores for the same test within a defined year were represented by the first score within that year. Repeater scores across years were included in the appropriate years. For example, suppose a person took the same test in November 1979, February 1980, November 1980, November 1981 and April 1982. That person would be represented three times in the extracted files, as follows:

1979-80 November 1979 (NOT February 1980)
1980-81 November 1980
1981-82 November 1981 (NOT April 1982)

The samples were further restricted to include examinees tested at centers in states that had consistent certification requirements across the years 1978-79 to 1983-84. These states, specified for us by NTE Program Direction (Ed Masonis) are as follows: Arkansas, California, Illinois, Louisiana, Mississippi, New York, North Carolina, South Carolina, Tennessee and West Virginia. In the case of Illinois and New York, the numbers of examinees were consistently small, since only certain school districts in Chicago and New York City required the tests - not the entire states.

Testings from national and special administrations are included. The following table shows the dates of the national administrations that occurred during the time-period covered by this study. Note that the first year, defined as November 1, 1978 - June 30, 1979 includes no summer national administration and that the third year, July 1, 1980 - June 30, 1981, includes two summer national administrations (July 19, 1980 and June 20, 1981). Since the first year was not a full year, it was dropped from the study.

Year	National Test Dates				
78-79		Nov. 11, 1978	Feb. 17, 1979		
79-80	July 21, 1979	Nov. 10, 1979	Feb. 16, 1980		
80-81	July 19, 1980	Nov. 8, 1980	Feb. 21, 1981		June 20,1981
81-82		Nov. 14, 1981	Feb. 20, 1982	April 17, 1982	
82-83		Oct. 30, 1982		April 30, 1983	
83-84		Nov. 12, 1983		April 14, 1984	

[2] Since the first year, 1978-79, included no summer administration and essentially no special administrations, it was decided to drop it from the study and to include only five years each covering a full year, from July 1 through June 30.

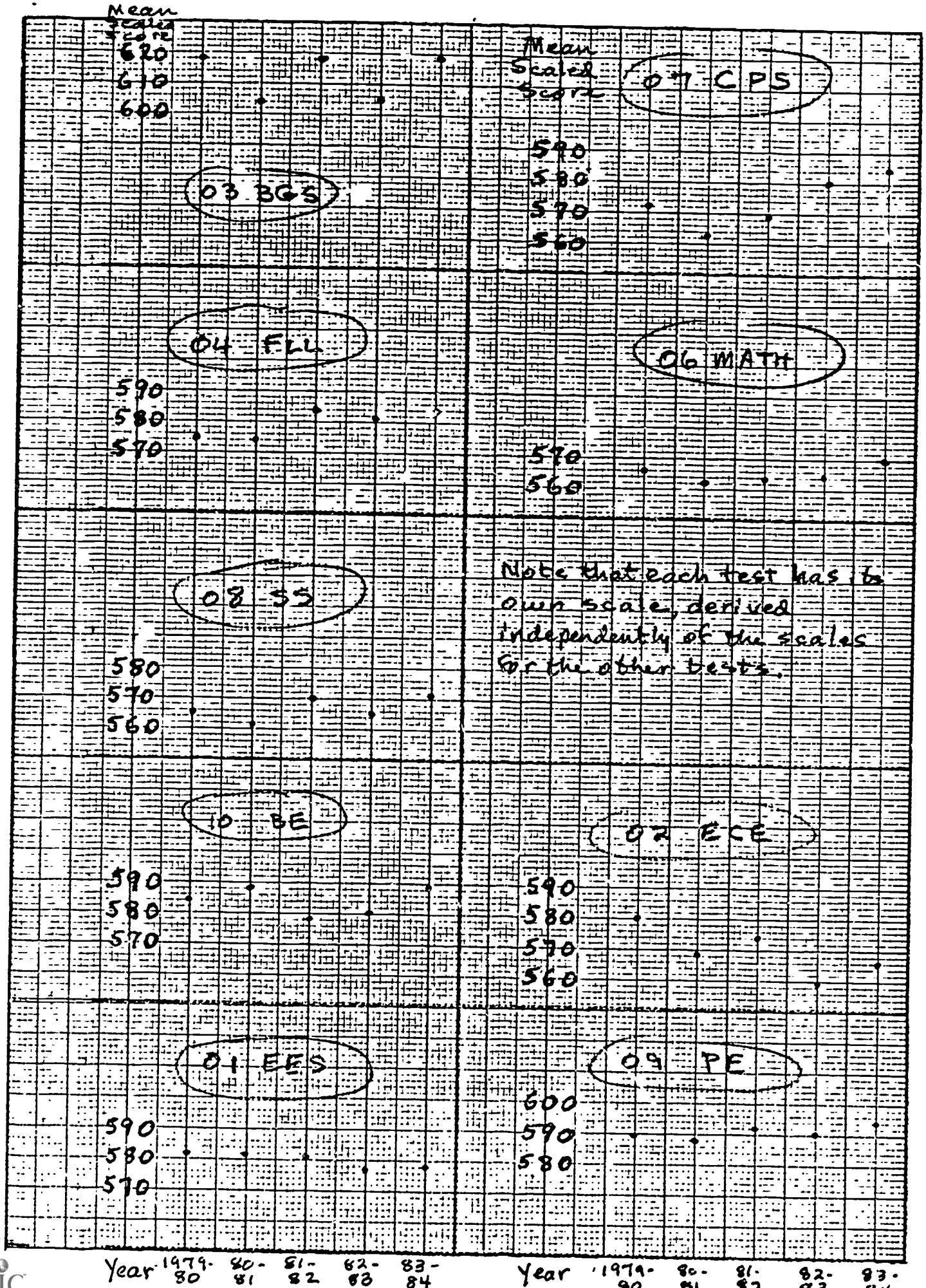
Additional tables, showing distributions of examinees by background information on sex, race and educational level, are also included for the first five of the nine tests (the ones of primary interest to NSF). These tables seem to raise more questions than they answer. Why, for example, is the percentage of examinees in the category, "Native American, Eskimo or Aleut" relatively high in the first three years and low in the last two? (A new processing system was used beginning in October 1982. Collection of background data is now done at the testing site and coded on the answer sheet. Prior to October 1982 background information questions were in the registration materials and were coded on a separate registration form. But should this difference in the way information was collected have caused the differences that showed up in these tables?)

One-way analyses of variance across years were done for each of the first five tests (BGS, CPS, ELL, MAT and SS). The results indicate that one would not ordinarily expect to find differences of the size found here among samples of this size randomly drawn from the same population (BGS, ELL, SS .01 level, CPS and MAT .05 level but not .01 level).

It would seem that, without additional data and/or analyses, one should be very cautious about drawing conclusions about trends one thinks one sees in the data. There are numerous factors that could influence these data, such as the time of year when people were tested, the supply and demand of the teaching positions in these fields (and note that the examinees who took these tests were presumably applying for teacher certification; they did not necessarily become teachers), the amount and recency of preparation in coursework related to the subject area. A serious limitation of the data in NTE Programs' files is the lack of information about when a person's training was acquired, for instance. Some may be currently undergoing their training; others may have had a lapse of several years between acquiring their training and taking the test.

Test	Year	Ns	Mean	SD
03 BGS	79-80	1,144	618	89
	80-81	1,545	604	89
	81-82	1,048	618	87
	82-83	1,013	605	87
	83-84	1,211	619	84
07 CPS	79-80	223	572	93
	80-81	343	563	84
	81-82	238	569	93
	82-83	249	580	90
	83-84	327	584	96
04 ELL	79-80	2,144	575	91
	80-81	2,668	574	91
	81-82	1,843	584	94
	82-83	1,550	581	94
	83-84	1,522	584	96
06 MAT	79-80	1,298	566	89
	80-81	1,820	562	89
	81-82	1,359	563	83
	82-83	1,449	565	80
	83-84	1,933	570	81
08 SS	79-80	2,480	565	91
	80-81	3,237	561	93
	81-82	2,106	570	93
	82-83	1,736	565	92
	83-84	1,852	571	93

Test	Year	Ns	Mean	SD
10 BE	79-80	1,040	584	85
	80-81	1,275	588	85
	81-82	846	578	87
	82-83	629	580	84
	83-84	669	589	84
02 ECE	79-80	5,425	580	106
	80-81	5,787	568	108
	81-82	3,836	574	103
	82-83	3,170	559	106
	83-84	3,422	566	106
01 EES	79-80	8,890	582	97
	80-81	11,570	582	98
	81-82	9,370	581	93
	82-83	8,083	577	90
	83-84	8,628	578	88
09 PE	79-80	3,109	589	85
	80-81	3,611	588	87
	81-82	2,635	592	81
	82-83	2,2[illegible]	590	84
	83-84	2,1[illegible]	594	81



BP 9-7#11-84

NTE PROGRAMS Specialty Area Test in Biology and General Science (03 BGS)
Summary Data Based on Selected Subgroups (See notes) of Examinees

Year		1979-80 (July-June)	1980-81 (July-June)	1981-82 (July-June)	1982-83 (July-June)	1983-84 (July-June)
Mean		618.	604.	618.	605.	619.
S.D.		89.	89.	87.	87.	84.
Number of Examinees		1,144.	1,545.	1,048.	1,013.	1,211.
Distribution by Educational Level						
1 Freshman	(Freshman, Sophomore, and Junior were not given as response options. Examinees at these levels are in the "Not Coded" category.)				1. (0.1%)	0 ...
2 Sophomore					3. (0.3)	0 ...
3 Junior					13. (1.3)	7. (0.6%)
4 Senior		389. (34.0%)	404. (26.1%)	294. (28.1%)	259. (25.6)	282. (23.3)
5 Bachelor's Degree		608. (53.1)	915. (59.2)	582. (55.5)	459. (45.3)	507. (41.9)
6 Enrolled in Graduate School	(This response option was first used in October 1982.)				138. (13.6)	235. (19.4)
7 Master's Degree		109. (9.5)	158. (10.2)	120. (11.5)	119. (11.7)	160. (13.2)
8 Doctoral Degree		6. (0.5)	17. (1.1)	11. (1.0)	11. (1.1)	14. (1.2)
0 Not Coded		32. (2.8)	51. (3.3)	41. (3.9)	10. (1.0)	6. (0.5)
Distribution by Sex						
1 Male		450. (39.3)	582. (37.7)	419. (40.0)	385. (38.0)	482. (39.8)
2 Female		694. (60.7)	963. (62.3)	629. (60.0)	607. (59.9)	692. (57.1)
0 Not Coded	(Records without sex code were coded clerically or by computer)				21. (2.1)	37. (3.1)
Distribution by Responses to Question, "How do you describe yourself?"						
1 Black, Afro-American, or Negro		147 (12.8)	197 (12.8)	129 (12.3)	126 (12.4)	127 (10.5)
2 Mexican American or Chicano		7 (0.6)	2. (0.1)	3 (0.3)	8 (0.8)	5 (0.4)
3 Native American, Eskimo or Aleut		66 (5.8)	114 (7.4)	47 (4.5)	2 (0.2)	4 (0.3)
4 Oriental or Asian American		14 (1.2)	22 (1.4)	7 (0.7)	8 (0.8)	10 (0.8)
5 Puerto Rican		3 (0.3)	5 (0.3)	3 (0.3)	1 (0.1)	0 ...
6 Other Hispanic or Latin American		3 (0.3)	5 (0.3)	3 (0.3)	10 (1.0)	6. (0.5)
7 White		775 (67.7)	1,031 (66.7)	691 (65.9)	739 (73.0)	954 (78.8)
8 Other		10 (0.9)	15 (1.0)	14 (1.3)	14 (1.4)	12 (1.0)
0 Not Coded		119 (10.4)	154 (10.0)	151 (14.4)	105 (10.4)	93 (7.7)

NTE PROGRAMS Specialty Area Test in Chemistry, Physics, and General Science (07 CPS)
– Summary Data Based on Selected Subgroups (See notes) of Examinees

Year		1979-80 (July-June)	1980-81 (July-June)	1981-82 (July-June)	1982-83 (July-June)	1983-84 (July-June)
Mean		572.	563.	569.	580.	584.
S.D.		93.	84.	93.	90.	96.
Number of Examinees		223.	343.	238.	249.	327.
Distribution by Educational Level						
1 Freshman	(Freshman, Sophomore, and Junior were not given as response options. Examinees at these levels are in the "Not Coded" category.)				0 . . .	0 . . .
2 Sophomore					0 . . .	0 . . .
3 Junior					0 . . .	6. (1.8%)
4 Senior		58. (26.0%)	69. (20.1%)	43. (18.1%)	46. (18.5%)	59. (18.0)
5 Bachelor's Degree		119. (53.4)	195. (56.9)	143. (60.1)	113. (45.4)	130. (39.8)
6 Enrolled in Graduate School	(This response option was first used in October 1982.)	33. (14.8)	60. (17.5)	41. (17.2)	37. (14.9)	65. (19.9)
7 Master's Degree		7. (3.1)	10. (2.9)	5. (2.1)	45. (18.1)	55. (16.8)
8 Doctoral Degree		6. (2.7)	9. (2.6)	6. (2.5)	3. (1.2)	11. (3.4)
0 Not Coded					5. (2.0)	1. (0.3)
Distribution by Sex						
1 Male		128. (57.4)	194. (56.6)	130. (54.6)	140. (56.2)	177. (54.1)
2 Female		95. (42.6)	149. (43.4)	108. (45.4)	102. (41.0)	140. (42.8)
0 Not Coded	(Records without this code were coded clerically or by computer)				7. (2.8)	10. (3.1)
Distribution by Responses to Question, "How do you describe yourself?"						
1 Black, Afro-American, or Negro		19. (8.5)	25. (7.3)	21. (8.8)	27. (10.8)	22. (6.7)
2 Mexican American or Chicano		1. (0.4)	1. (0.3)	2. (0.8)	1. (0.4)	0. . .
3 Native American, Eskimo or Aleut		19. (8.5)	19. (5.5)	15. (6.3)	3. (1.2)	3. (0.9)
4 Oriental or Asian American		3. (1.3)	4. (1.2)	8. (3.4)	6. (2.4)	5. (1.5)
5 Puerto Rican		0. . . .	0. . . .	0. . . .	1. (0.4)	2. (0.6)
6 Other Hispanic or Latin American		2. (0.9)	1. (0.3)	3. (1.3)	0. . . .	0. . . .
7 White		145. (65.0)	233. (67.9)	147. (61.8)	181. (72.7)	270. (82.6)
8 Other		6. (2.7)	10. (2.9)	3. (1.3)	5. (2.0)	3. (0.9)
Not Coded		28. (12.6)	50. (14.6)	39. (16.4)	25. (10.0)	22. (6.7)

NTE PROGRAMS Specialty Area Test in English Language and Literature (04 ELL)
Summary Data Based on Selected Subgroups (see notes) of Examinees

Year		1979-80 (July-June)	1980-81 (July-June)	1981-82 (July-June)	1982-83 (July-June)	1983-84 (July-June)
Mean		575.	574.	584.	581.	584.
S.D.		91.	91.	94.	94	96.
Number of Examinees		2,144.	2,668.	1,843.	1,550.	1,522.
Distribution by Educational Level						
1 Freshman	(Freshman, Sophomore and Junior were not given as response options. Examinees at those levels are in the "Not Coded" category.)	654 (30.5%)	691. (25.9%)	549. (29.8%)	2. (0.1%)	1. (0.1%)
2 Sophomore					0.	1. (0.1)
3 Junior					20. (1.3)	22. (1.4)
4 Senior		1,098. (51.2)	1,493. (56.0)	940. (51.0)	561. (36.2)	504. (33.1)
5 Bachelor's Degree	(This response option was first used in October 1987.)	310. (14.5)	394. (14.8)	289. (15.7)	206. (13.3)	288. (18.9)
6 Enrolled in Graduate School		15. (0.7)	22. (0.8)	16. (0.9)	262. (16.9)	219. (14.4)
7 Master's Degree		67. (3.1)	68. (2.5)	49. (2.7)	14. (0.9)	14. (0.9)
8 Doctoral Degree					13. (0.8)	8. (0.5)
9 Not Coded						
Distribution by Sex						
1 Male	(Records without the sex code were coded clerically or by computer.)	422. (19.7)	523. (19.6)	387. (21.0)	274. (17.7)	261. (17.1)
2 Female		1,722. (80.3)	2,145. (80.4)	1,456. (79.0)	1,241. (80.1)	1,222. (80.3)
0 Not Coded					35. (2.3)	39. (2.6)
Distribution by Responses to Question, "How do you describe yourself?"						
1 Black, Afro-American, or Negro		285. (13.3)	302. (11.3)	167. (9.1)	192. (12.4)	180. (11.8)
2 Mexican American or Chicano		13. (0.6)	8. (0.3)	17. (0.9)	10. (0.6)	6. (0.4)
3 Native American, Eskimo or Aleut		134. (6.3)	222. (8.3)	135. (7.3)	7. (0.5)	6. (0.4)
4 Oriental or Asian American		25. (1.2)	16. (0.6)	10. (0.5)	9. (0.6)	9. (0.6)
5 Puerto Rican		5. (0.2)	8. (0.3)	5. (0.3)	1. (0.1)	5. (0.3)
6 Other Hispanic or Latin American		7. (0.3)	11. (0.4)	9. (0.5)	5. (0.3)	8. (0.5)
7 White		1,394. (65.0)	1,743. (65.3)	1,175. (63.8)	1,163. (75.0)	1,166. (76.6)
8 Other		30. (1.4)	34. (1.3)	20. (1.1)	11. (0.7)	13. (0.9)
9 Not Coded		251. (11.7)	324. (12.1)	305. (16.5)	152. (9.8)	129. (8.5)

NTE PROGRAMS Specialty Area Test in Mathematics (06 MAT)
— Summary Data Based on Selected Subgroups (See notes) of Examinees

Year		1979-80 (July-June)	1980-81 (July-June)	1981-82 (July-June)	1982-83 (July-June)	1983-84 (July-June)
Mean		566.	562.	563.	565.	570.
S.D.		89.	89.	83.	80.	81.
Number of Examinees		1,298.	1,820.	1,359.	1,449.	1,933.
Distribution by Educational Level						
1 Freshman	(Freshman, Sophomore, and Junior were not given as response options. Examinees at these levels are in the "Not Coded" category.)				1. (0.1)	0.
2 Sophomore					4. (0.3)	1. (0.1)
3 Junior					12. (0.8)	27. (1.4)
4 Senior		400. (30.8)	433. (23.8)	337. (24.8)	363. (25.1)	368. (19.0)
5 Bachelor's Degree		684. (52.7)	1,058. (58.1)	752. (55.3)	590. (40.7)	762. (39.4)
6 Enrolled in Graduate School	(This response option was first used in Oct. 1982)				236. (16.3)	428. (22.1)
7 Master's Degree		166. (12.8)	258. (14.2)	209. (15.4)	214. (14.8)	322. (16.7)
8 Doctoral Degree		12. (0.9)	16. (0.9)	10. (0.7)	12. (0.8)	15. (0.8)
0 Not Coded		36. (2.8)	55. (3.0)	51. (3.8)	17. (1.2)	10. (0.5)
Distribution by Sex						
1 Male		527. (40.6)	750. (41.2)	560. (41.2)	595. (41.1)	814. (42.1)
2 Female		771. (59.4)	1,070. (58.8)	799. (58.8)	817. (56.4)	1,044. (54.0)
0 Not Coded	(Records without sex code were coded clerically or by computer.)				37. (2.6)	75. (3.9)
Distribution by Responses to Question, "How do you describe yourself?"						
1 Black, Afro-American, or Negro		188. (14.5)	237. (13.0)	142. (10.4)	156. (10.8)	188. (9.7)
2 Mexican American or Chicano		12. (0.9)	12. (0.7)	8. (0.6)	13. (0.9)	16. (0.8)
3 Native American, Eskimo or Aleut		61. (4.7)	103. (5.7)	75. (5.5)	6. (0.4)	11. (0.6)
4 Oriental or Asian American		29. (2.2)	40. (2.2)	29. (2.1)	36. (2.5)	51. (2.6)
5 Puerto Rican		7. (0.5)	4. (0.2)	1. (0.1)	4. (0.3)	2. (0.1)
6 Other Hispanic or Latin American		3. (0.2)	9. (0.5)	16. (1.2)	18. (1.2)	8. (0.4)
7 White		826. (63.6)	1,165. (64.0)	833. (61.3)	1,056. (72.9)	1,463. (75.7)
8 Other		31. (2.4)	24. (1.3)	20. (1.5)	27. (1.9)	31. (1.6)
0 Not Coded		141. (10.9)	226. (12.4)	235. (17.3)	133. (9.2)	163. (8.4)

NTE PROGRAMS Specialty Area Test in Social Studies (08 SS)
Summary Data Based on Selected Subgroups (See notes) of Examinees

Year		1979-80 (July-June)	1980-81 (July-June)	1981-82 (July-June)	1982-83 (July-June)	1983-84 (July-June)
Mean		565.	561.	570.	565.	571.
S.D.		91.	93.	93.	92.	93.
Number of Examinees		2,480.	3,237.	2,106.	1,736.	1,852.
Distribution by Educational Level						
1 Freshman	(Freshman, sophomore, and junior were not given as response options. Examinees at these levels are in the "Not Coded" category.)				1. (0.1)	2. (0.1)
2 Sophomore					4. (0.2)	0.
3 Junior					15. (0.9)	21. (1.1)
4 Senior		853. (34.4)	873. (27.0)	630. (29.9)	499. (28.7)	527. (28.5)
5 Bachelor's Degree		1,231. (49.6)	1,807. (55.8)	1,107. (52.6)	706. (40.7)	691. (37.3)
6 Enrolled in Graduate School	(This response option was first used in Oct. 1982)				240. (13.8)	340. (18.4)
7 Master's Degree		290. (11.7)	405. (12.5)	272. (12.9)	246. (14.2)	239. (12.9)
8 Doctoral Degree		18. (0.7)	15. (0.5)	18. (0.9)	6. (0.3)	19. (1.0)
0 Not Coded		88. (3.5)	137. (4.2)	79. (3.8)	19. (1.1)	13. (0.7)
Distribution by Sex						
1 Male		1,261. (50.8)	1,572. (48.6)	1,047. (49.7)	861. (49.6)	869. (46.9)
2 Female		1,219. (49.2)	1,665. (51.4)	1,059. (50.3)	840. (48.4)	932. (50.3)
0 Not Coded	(Records without sex code were coded clerically or by computer.)				35. (2.0)	51. (2.8)
Distribution by Responses to Question, "How do you describe yourself?"						
1 Black, Afro-American, or Negro		417. (16.8)	477. (14.7)	266. (12.6)	276. (15.9)	280. (15.1)
2 Mexican American or Chicano		28. (1.1)	29. (0.9)	15. (0.7)	11. (0.6)	11. (0.6)
3 Native American, Eskimo or Aleut		152. (6.1)	258. (8.0)	134. (6.4)	17. (1.0)	15. (0.8)
4 Oriental or Asian American		15. (0.6)	26. (0.8)	13. (0.6)	8. (0.5)	17. (0.9)
5 Puerto Rican		2. (0.1)	8. (0.2)	5. (0.2)	2. (0.1)	0.
6 Other Hispanic or Latin American		14. (0.6)	14. (0.4)	11. (0.5)	5. (0.3)	9. (0.5)
7 White		1,522. (61.4)	1,920. (59.3)	1,252. (59.4)	1,193. (68.7)	1,299. (70.1)
8 Other		36. (1.5)	44. (1.4)	24. (1.1)	28. (1.6)	26. (1.4)
Not Coded		294. (11.9)	461. (14.2)	386. (18.3)	196. (11.3)	195. (10.5)

APPENDIX D

Advanced Placement Analyses

Advanced Placement Analyses

TABLE I. AND CHART I.

SUMMARY: Advanced Placement (AP) courses and examinations are given at over 20% of American secondary schools to 15 to 20% of their college-bound students. Participants are by definition doing college-level work and are high achievers and generally very highly motivated. Table I., attached, shows the numbers of examinations given in eight subject areas from 1975-1984. Chart I., also attached, graphically depicts the changes. These data provide some indication of the importance of the various subject areas to high-achieving students over the last ten years. (Table I. and Chart I. should be considered in conjunction with Table II., which indicates the relative importance of each subject to all others each year for the last ten years.)

Table I. and Chart I. indicate that the volumes of all eight AP examination subject areas shown have increased over the ten-year period. Those with the highest volumes (American History and English) increased the most in terms of both absolute numbers and rates of increase (211% and 177% respectively). (Chemistry experienced the third largest rate increase (176%), although it remains a relatively low volume program. The other sciences (Physics and Biology) had relatively low rates of increase (162 % and 136% respectively), and Mathematics (Calculus), currently the third largest in volume, had the next to smallest rate of increase (134%).

Taking an AP course usually suggests students' particular interest in a subject area and their intention to pursue that subject as a major field of study in college. However their interest is likely enhanced by the fact that advanced standing is given by many colleges and universities to those who score well. Since some students take more than one AP course, all test-takers are not necessarily future majors in the subject of each test.

CAVEATS: The AP exams included in Table I. and Chart I. were selected to encompass the math and all the science exams as well as several other exams for comparative purposes. These exams were relatively stable programatically throughout the ten year period. That is, no major changes took place in these programs which would have caused large and lasting volume changes. A possible exception is English, which expanded in 1980 from English Literature and Composition to include an English Language and Composition exam as well. A significant volume increase occurred at that time, although the rate of increase declined in subsequent years.

In addition to students' interest in a subject area, volumes may be affected by the availability of teachers to teach the courses. There has been some speculation that the small increases in the math and particularly the science volumes may be due to a relative scarcity of available teachers.

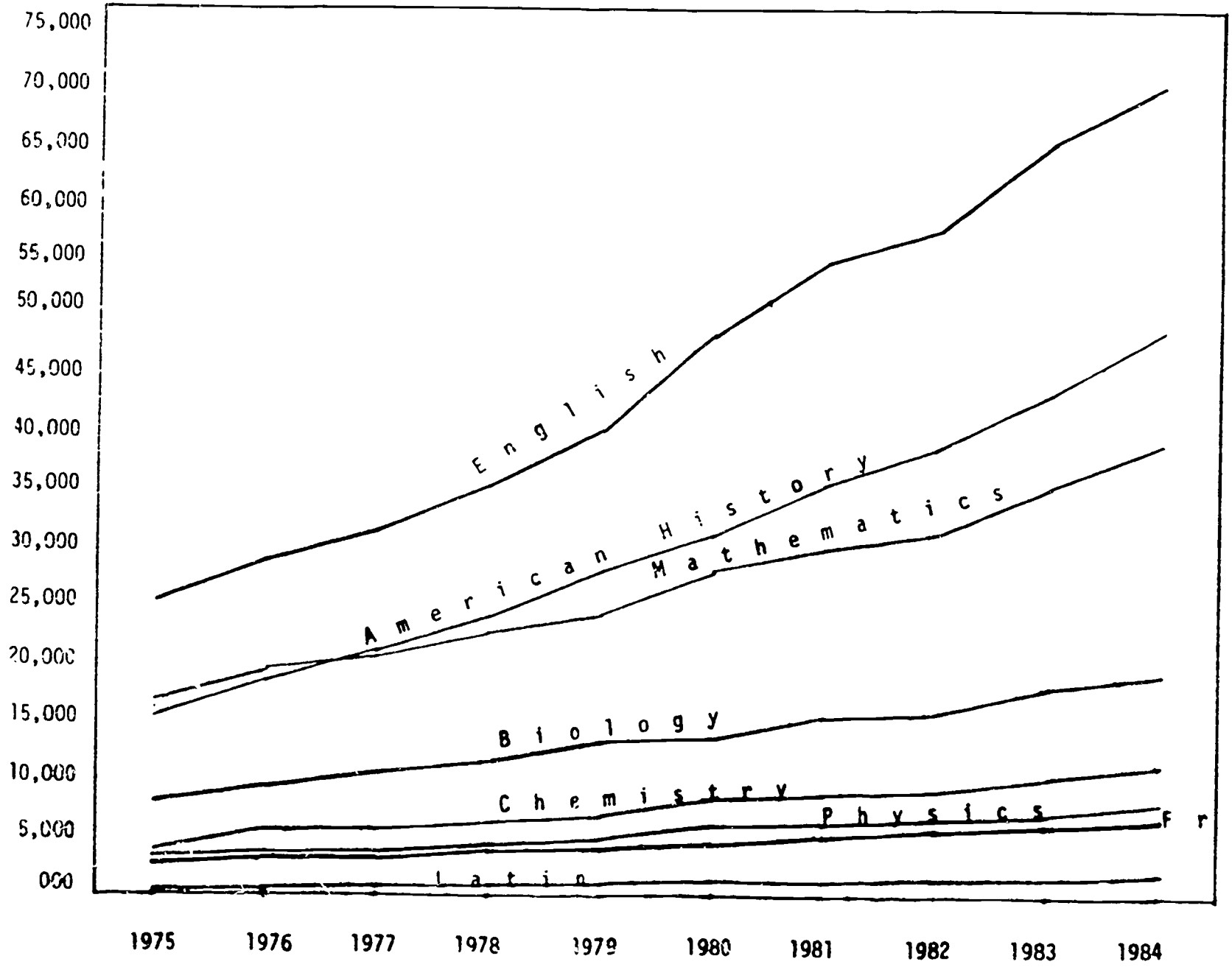
TABLE I.

ADVANCED PLACEMENT EXAMINATION VOLUMES
in Selected Subject Areas
1975-1984

	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	% inc. '75-'84
Am. History	16,068	18,718	21,325	24,444	28,222	32,098	35,999	38,286	43,844	49,939	211
Biology	8,206	9,482	10,530	11,342	12,835	13,549	15,199	15,947	17,804	19,387	136
Chemistry	4,181	5,341	5,559	6,270	7,016	8,209	8,877	9,476	10,291	11,539	176
English	25,656	29,503	32,142	36,334	41,975	49,125	55,010	58,591	65,260	71,263	177
French Lang.	3,029	3,374	3,601	3,994	4,409	4,920	5,352	5,782	6,129	6,843	126
Latin/Vergil	624	745	841	880	1,016	1,122	1,261	1,311	1,529	1,704	173
Math (Calc.)	17,090	19,065	20,317	22,510	24,727	27,879	30,558	31,918	35,489	39,962	134
Physics	3,200	3,663	4,196	4,556	5,039	6,222	6,481	6,804	7,376	3,390	162

SOURCE: "Advanced Placement Examination Volume Changes." (Table prepared by Educational Testing Service, Princeton, Princeton, New Jersey, 1984, for the Advanced Placement Program of the College Board.)

CHART I.
ADVANCED PLACEMENT EXAMINATION VOLUMES
in Selected Subject Areas, 1975-1984



SOURCE: "Advanced Placement Examination Volume Changes." (Table prepared by Educational Testing Service for the Advanced Placement Program of the College Board. Princeton, N.J., 1984.)

TABLE II.

SUMMARY: Advanced Placement (AP) courses and examinations are given at over 20% of American secondary schools to 15 to 20% of their college-bound students. Participants are by definition doing college-level work and are high achievers and generally very highly motivated. The attached chart shows the percentage of all of the AP examinations that were given in each of 13 subjects from 1975-1984. This thus gives an indication of the relative importance to high-achieving students of the various subjects over a ten-year period. Students' interest in taking AP courses is enhanced by the fact that advanced standing is given by many colleges and universities to those who score well. Taking an AP course usually suggests students' particular interest in a subject area and their intention to pursue that subject as a major field of study in college. Since some students take more than one AP course, however, all test-takers are not future majors in the subject of each test.

The attached table shows that, as indicated by AP exam volumes, student interest in biology, chemistry, and physics, relative to other subjects, has remained very much the same over the last ten years, while their relative interest in mathematics has declined dramatically, more than for any other subject. Math (Calculus), in fact, dropped from subject of second greatest interest (English being first) in 1976 to third, surpassed also by American History, in 1977 and thereafter. Biology, Chemistry, and Physics were in 4th, 6th, and 8th place in 1975 and 1984. Unfortunately trend data are not available for AP Computer Science, the exam for which was administered for the first time in the Spring of 1984.

CAVEATS: Certain temporary fluctuations in AP volumes may be partially explained when specific programs have been added or dropped. For example: a third Art examination (in Drawing) was added in 1980 and the Art total increased slightly from 1.0 to 1.1%. Also in 1980 an English Language and Composition exam was added. A slight decrease in the Literature and Composition volume occurred, but Total English increased one percentage point. The German program added a Language exam in 1980 (in addition to Literature), and the Total German percent increased by .3 points to .8% of all AP test-takers. Total German dropped down to .7% in 1983 at the same time that German Literature was discontinued. In addition, increases were noted in Music Total in 1978 (from .3% to .7%) at the time the Music Theory exam was added, and in Spanish Total in 1977 (from 1.8% to 2.8%) when the Spanish Language exam was included (along with Spanish Literature).

These possible program-induced fluctuations, however, should not be strong enough to explain a steady trend over a ten-year period within a subject area.

TABLE II.

ADVANCED PLACEMENT EXAMINATIONS TAKEN 1975-1984
PERCENT OF TOTAL BY SUBJECT

	'75 %	'76 %	'77 %	'78 %	'79 %	'80 %	'81 %	'82 %	'83 %	'84 %
American History	18.7	18.9	19.6	19.9	20.2	20.0	20.2	20.3	20.8	20.8
*Art	1.0	1.0	.9	1.0	1.0	1.1	1.1	1.2	1.1	1.1
Biology	9.7	9.6	9.7	9.3	9.2	8.5	8.5	8.4	8.4	8.1
Chemistry	4.9	5.4	5.1	5.1	5.0	5.1	5.0	5.0	4.9	4.8
*English	29.9	29.8	29.5	29.6	30.1	30.6	30.9	31.0	30.9	29.7
European History	5.1	5.3	4.9	4.8	5.0	5.1	5.2	5.3	5.5	5.3
French	3.5	3.4	3.3	3.3	3.2	3.1	3.0	3.1	2.9	2.9
German	.7	.6	.6	.6	.5	.8	.8	.8	.7	.7
Latin	.7	.8	.8	.7	.7	.7	.7	.7	.7	.7
Math	19.9	19.3	18.7	18.4	17.7	17.4	17.2	16.9	16.8	16.7
*Music	.4	.4	.3	.7	.7	.6	.4	.3	.3	.3
Physics	3.7	3.7	3.8	3.7	3.6	3.8	3.6	3.6	3.5	3.5
*Spanish	1.8	1.8	2.8	2.9	3.1	3.2	3.4	3.4	3.5	3.8
Total Percent	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	98.4%**
Total Number of Examinations Taken	85,786	98,898	108,870	122,561	139,544	160,214	178,159	188,933	211,160	239,666

* See Caveats section attached.
** Total for 1984 does not equal 100% because Computer Science (not shown) was included for the first time.

SOURCE: "Advanced Placement Examination Volume Changes." (Table prepared by Educational Testing Service for the Advanced Placement Program of the College Board. Princeton, New Jersey, 1984.)

TABLE III.

SUMMARY:

Advanced Placement (AP) courses and examinations are given at over 20 percent of American secondary schools to 15-20 percent of their college bound students. Participants are by definition doing college-level work and are high achievers and generally very highly motivated. Each examination (with the exception of Studio Art) includes both an objective (multiple choice) section and a free response or essay portion. In each subject area a group of teachers grades the free reponse part of the examination. These teachers, from participating schools and colleges across the country, are organized and directed in their grading by a chief reader, who is a college professor. Final grades, based on students' entire examinations (with free response and multiple choice questions appropriately weighted), are reported on a 5-point scale: 5-extremely well qualified, 4-well qualified, 3-qualified, 2-possibly qualified, 1-no recommendation. Participating colleges normally honor grades of 3 or higher.

The attached table shows the mean grades in all subject areas from 1975-1984. No interpretation of these scores is provided here, as ETS does not recommend that these data be used for Science Indicators or for any other trend analysis purpose. This table is being provided for information purposes only, as NSF requested an update on the Lyle Jones article of 1981. (See below.)

CAVEATS:

ETS does not recommend the use of AP mean grade data for trend analyses. The grade data are based on scores from both the multiple choice portion and free response parts of the examinations. However, only the multiple choice portion has been equated. The scores from both portions of the exam are weighted and combined and placed on scales that range from a minimum of 0-9 for the Mathematics and Physics C exams to a maximum of 0-200 for Spanish Literature, History of Art, and the Music exams. Cut scores are established at four different points along these scales to designate a grade of 1, 2, 3, 4, or 5. Cut scores frequently vary from year to year for each examination, reflecting changes in levels of exam difficulty, and they also differ across examinations. In addition, the chief readers often attempt to maintain similar percentages of students receiving each of the scores from 1-5 from year to year, unless there is a reason not to.

For all these reasons, when considered in conjunction with the fact that mean grades are not equated from year to year, it is clear that the use of AP grades for trend analyses would be inappropriate and of very little value. A special study that allowed the use of mean grades for trend analysis purposes was conducted by ETS in the early 1980's for Lyle Jones, who published his findings in a widely distributed article in Science, entitled "Achievement Test Scores in Mathematics and Science." These results showed little average change in math or science achievement from 1973-1979.

Scaled data from the equated portions of the math and science AP exams are not routinely prepared, and doing so would require a project of longer duration than is available under this contract. However, such a project could be considered for the future.

TABLE III.

Advanced Placement Mean Grades, 1975-1984

	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984
American History	3.1	3.1	3.1	3.2	3.12	3.08	3.06	3.09	3.11	3.10
History of Art	2.9	3.0	3.0	3.2	3.24	3.41	3.20	3.20	3.16	3.13
Biology	3.3	3.3	3.4	3.4	3.33	3.31	3.35	3.31	3.30	3.25
Chemistry	3.1	3.1	3.1	3.1	3.09	3.03	3.02	3.01	3.05	3.02
Engl. Comp./Lit	3.1	3.1	3.1	3.1	3.10	3.08	3.07	3.05	3.07	3.05
European Hist.	3.1	3.1	3.2	3.2	3.09	3.15	3.03	3.15	3.10	3.13
French Language	3.1	3.1	3.1	3.1	3.13	3.06	3.01	3.05	3.03	3.01
German Literature	3.0	3.0	3.1	3.0	3.13	3.11	3.11	3.10	---	---
Latin/Vergil	3.4	3.4	3.3	3.3	3.28	3.22	3.23	3.23	3.12	3.09
Math Calc. - AB	3.0	3.0	3.0	3.0	3.03	3.02	3.06	3.08	3.09	3.13
Calc. - BC	3.4	3.3	3.4	3.3	3.28	3.19	3.32	3.33	3.39	3.38
Music Theory	3.0	3.0	3.2	3.1	3.18	3.09	3.15	3.13	3.05	3.04
Physics B (Gen'l.)	2.9	2.9	2.9	2.9	2.96	2.93	3.05	2.97	2.91	2.95
C (Mech.)	3.4	3.4	3.4	3.4	3.38	3.36	3.37	3.42	3.31	3.44
C (Elec. & Mag.)	2.9	3.4	3.3	3.4	3.38	3.37	3.25	3.26	3.28	3.36
Spanish Language	3.1	3.1	3.1	3.2	3.31	3.61	3.45	3.48	3.36	3.49

SOURCE: This table was prepared by Educational Testing Service in September 1984 from annual charts (1975 through 1984), entitled "Distribution of Candidate Grades - Advanced Placement Examinations." (The Advanced Placement Program is a program of the College Board.)