DOCUMENT RESUME

ED 336 267 SE 052 171

AUTHOR	Jacobson, Willard J.; And Others
TITLE	The Second IEA Science Study--U.S. Revised Edition.
INSTITUTION	International Association for the Evaluation of Educational Achievement, New York, NY.
SPONS AGENCY	Center for Education Statistics (OERI/ED), Washington, DC.; National Science Foundation, Washington, D.C.; Spencer Foundation, Chicago, Ill.
PUB DATE	Sep 87
CONTRACT	NSF-8470382
NOTE	54p.; A Presentation at a Meeting of the General Assembly of the International Association for the Evaluation of Educational Achievement (New York, NY, September 16, 1987).
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	*Academic Achievement; Comparative Education; Educational Change; *Elementary School Science; Elementary Secondary Education; *Foreign Countries; International Cooperation; Science Education; *Secondary School Science; *Sex Differences

ABSTRACT

The Second IEA Science Study (SISS) was carried out in 1983 in 24 countries. In the United States a second phase of testing was undertaken in 1986. This document describes the organization of the study in the United States and highlights some of the results and their possible implications for science education. The countries included in the SISS are Australia, China, Canada, England, Finland, Ghana, Hong Kong, Hungary, Israel, Italy, Japan, Korea (Republic), The Netherlands, Nigeria, Norway, Papua-New Guinea, Philippines, Poland, Singapore, Sri Lanka, Sweden, Tanzania, United States, and Zimbabwe. Sections include (1) "Some Results from the Second IEA Science Study"; (2) "Science Education in the 1970s and 1980s: What Changes Have Taken Place?"; (3) "Sex and Science Achievement"; (4) "The Second IEA Science Study and Science Education in the United States"; and (5) "Comments and Reactions." (KR)

THE SECOND IEA SCIENCE STUDY--U.S.

Revised Edition

BY

WILLARD J. JACOBSON
RODNEY L. DORAN
EDITH Y.T. CHANG
EVE HUMRICH
JOHN P. KEEVES

The International Association for the Evaluation of Educational Achievement [IEA]

USA

Second IEA Science Study [SISS]

UNITED STATES

IEA

Teachers College, Columbia University

THE SECOND IEA SCIENCE STUDY--U.S.* **
Revised Edition

---

* A presentation at a meeting of the General Assembly of the International Association for the Evaluation of Educational Achievement (IEA) at Teachers College, Columbia University, 16 September 1987.

** This research has been supported by grants from The National Science Foundation, the Spencer Foundation, and the Center for Education Statistics.

Revised Edition

This material is based upon work supported by the National Science Foundation under Grant No. 8470382. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author and do not necessarily reflect the views of the National Science Foundation.

THE SECOND IEA SCIENCE STUDY--AN OVERVIEW

Willard J. Jacobson

The First IEA Science Study (FISS) was undertaken in 1970. The results were reported in Science Education in Nineteen Countries by Comber and Keeves (1973) and in the national reports of the participating countries. Some of the U.S. results are available in Achievement in America by Richard Wolf (1977).

The Second IEA Science Study (SISS) was carried out in 1983 in 24 countries. In the U.S. a second phase of testing was undertaken in 1986. This presentation will describe the organization of the Study in the United States and highlight some of the results and their possible implications for science education.

COUNTRIES IN THE SECOND IEA SCIENCE STUDY

AUSTRALIA	ISRAEL	PHILIPPINES
CHINA	ITALY	POLAND
CANADA	JAPAN	SINGAPORE
ENGLAND	KOREA (REPUBLIC)	SRI LANKA
FINLAND	THE NETHERLANDS	SWEDEN
GHANA	NIGERIA	TANZANIA
HONG KONG	NORWAY	U.S.A.
HUNGARY	PAPUA-NEW GUINEA	ZIMBABWE

ADMINISTRATIVE STAFF
SECOND IEA SCIENCE STUDY--U.S.

Willard J. Jacobson
National Research Coordinator, Second IEA Science Study
Professor of Natural Sciences, Teachers College, Columbia University

Rodney L. Doran
Associate National Research Coordinator, Second IEA Science Study
Professor of Science Education, State University of New York at Buffalo

Eugene W. Muller and Mark Rinkerman
Data Managers

Janis W. Owen
Administrative Assistant

Arthur E. Schneider

Drena Elfrink-Hetzel

RESEARCH ASSOCIATES
SECOND IEA SCIENCE STUDY--U.S.

Much of the research that is being carried out within the framework of SISS--U.S. is being undertaken by research associates. They are the sources of many of the innovative ideas that have emerged, and they have done much of the hard, tedious work that always seems associated with research.

O. Roger Anderson
Professor of Natural Sciences
Teachers College, Columbia University

Roosevelt J. Baker
Principal
Elementary School 28, Yonkers, NY

Steven L. Beyer
New York City Board of Education

Kevin J. Bleakley
Dean of Faculty
Trinity School, New York City

Arleene Cervasio
Greenwich, Connecticut Board of Education

Marilda S. Chandavarkar
Ministry of Education, Brazil

Edith Y. T. Chang
Second IEA Science Study

Thelma Clive
New York City Board of Education

Michael A. Dryden
Assistant Professor
University of Arizona

Ethelbert Ekeocha
Second IEA Science Study

Eve Humrich
Second IEA Science Study

Joan Jung
Graduate Assistant, Elementary Science Education
Queens College, City University of New York

Ira Kanis
Assistant Director of Science
New York City Board of Education

Joseph Menis
Bar-Elan University
Israel

Elizabeth A. Meng
Adjunct Assistant Professor of Science Education
Teachers College, Columbia University

James M. Micik
Greenwich, Connecticut Board of Education

June K. Miller
Assistant Professor of Science Education
Queens College, City University of New York

NATIONAL COMMITTEE
SECOND IEA SCIENCE STUDY--U.S.

In the United States, as in all other countries, there is a National Committee that advises on broad policy issues related to the National Study. The following have served on the National Committee of the U.S. Study:

Bill Aldridge, Executive Director
National Science Teachers Association

Hans O. Anderson, University of Indiana
Former President, Association for the
Education of Teachers in Science (AETS)

Essie Beck, Jefferson Parish Public School System
Former President, National Science Supervisors Association (NSSA)

Stanley Helgeson, The Ohio State University
Former President, National Association for
Research in Science Teaching (NARST)

Donald W. McCurdy, University of Nebraska
Former President, National Science Teachers Association (NSTA)

Kenneth R. Mechling, Clarion University
Former President, Council for Elementary Science International (CESI)

Douglas Reynolds, The New York State Education Department
Former President, Council of State Science Supervisors (CSSS)

Malcolm Rosier, International Coordinator
Second IEA Science Study

Paul Smith, Exxon Research and Engineering Company
Active in Chemical Education

Wayne Welch, University of Minnesota
Former President, National Association for
Research in Science Teaching (NARST)

Richard M. Wolf, U.S.A. Member, IEA General Assembly

SISS Monographs

The following is a list of the monographs that are being undertaken in SISS--U.S.:

Science Achievement in the United States
Willard J. Jacobson, Rodney L. Doran
Eugene W. Muller, and Mark Rinkerman

An Analysis of Science Curricula in the United States
June K. Miller

The Teaching and Learning of Biology in the United States
O. Roger Anderson

The Teaching and Learning of Chemistry in the United States
Joseph Menis

The Teaching and Learning of Physics in the United States
Marilda S. Chandavarkar

The Teaching and Learning of Elementary School Science
Elizabeth A. Meng

The Teaching and Learning of Science at the Ninth Grade Level
James Micik

Correlates of Science Achievement: A U.S. Study of Fifth Grade Students
Ethelbert Ekeocha

Science Achievement of Advanced Science Students
Arleene Cervasio

Science Achievement of Students not Studying Science
Joan Jung

Science Process Achievement
Ira Kanis

Science Education in the 1970s and 1980s: What Changes have Taken Place?
Edith Y. T. Chang

Sex and Achievement in Science
Eve Humrich

Science Achievement in the United States and in Other Countries: What can be Learned from Other Countries?
Kevin J. Bleakley

Science Achievement of Non-Science Students--A Case Study
Roosevelt J. Baker

Science Achievement in an American School-- A Case Study
James M. Micik

Achievement in Science Education in the United States: Student, Teacher, School, and Community Factors that Correlate with Science Achievement
Steven L. Beyer

Modeling Classroom Environments: An Analysis of Achievement at the Ninth Grade Level
Michael A. Dryden

The Trial Testing of Items and Instruments for the Second IEA Science Study: An Analysis of Results to Verify the Cumulative Hierarchical Nature of Bloom's Taxonomy of Educational Objectives (Cognitive Domain)
Thelma Clive

Science Curricula and Achievement in Science
June K. Miller

Science Achievement in Japan and the United States
Willard J. Jacobson, Shigekazu Takemura, Rodney L. Doran, Shigeo Kojima, Eve Humrich, and Masao Miyake

Science Education in the United States (A Report to Policymaking Groups and to the Public)
Willard J. Jacobson and Rodney L. Doran

There were two phases of testing conducted by the Second IEA Science Study--U.S. Phase I was carried out in 1983 with additional testing in 1984. In 1983/84 the sampling and data collection were carried out by the staff of SISS--U.S., using IEA procedures. Similar procedures were used in the First IEA Science Study (FISS) in 1970. The Second IEA Science Study Phase II was carried out in 1986. The sampling and data collection were conducted by the Research Triangle Institute (RTI) of North Carolina. We are indebted to the more than 20 thousand students and their teachers and administrators who cooperated in the Study.

TABLE 1
Student Instruments*
Second IEA Science Study--U.S.
Phase I, 1983/84

GRADE	SESSION 1	SESSION 2	SESSION 3
5	Achiev. Test (Int'l. Core) (24 Items)	6 Rotated Achiev. Tests (16 Int'l. Items) (8 Nat'l. Items)	Booklet 3**
9	Achiev. Test (Int'l. Core) (30 Items)	6 Rotated Achiev. Tests (20 Int'l. Items) (10 Nat'l. Items)	Student Questionnaires Opinionnaire
12 Physics	Achiev. Test (Gen. Sci.) (30 Int'l. Items) (5 Nat'l. Items)	Physics Test (30 Int'l. Items) (5 Nat'l. Items)	Science Learning Inventory
12 Non-Sci	Same as Above	Gen. Sci. Test (30 Int'l Items) (5 Nat'l. Items)	Word Knowledge Test

* All tests and questionnaires have been internationally developed and trial tested for validity with U.S. students.

** The format for Booklet 3 was the same for all populations, although additional items were included for Grades 9 and 12.

TABLE 2

Student Instruments*
Second IEA Science Study
Phase II, 1986

GRADE	SESSION 1	SESSION 2
5	Achievement Test Background Questions	Process (Lab) Test
9	Achievement Test Background Questions	Process (Lab) Test
10 First-Year Biology	Achievement Test	High School Science Booklet 2 Student Questionnaire Opinionnaire Word Knowledge Test** Mathematics Test**
11 First-Year Chemistry	Achievement Test	
12 Advanced/ Second-Year Science	Achievement Test (Biology, Chemistry, OR Physics)	

* All tests and questionnaires have been internationally developed and trial tested for validity with U.S. students.

** Separate versions of these tests were developed to accommodate the different levels of expertise among these groups of students.

TABLE 3

Number of Schools and Students by Population and Year

		GRADE 5	GRADE 9	GRADE 12 PHYSICS	GRADE 12 NON-SCI.
1983	SCHOOLS	121	87	158	134
	STUDENTS	2909	1958	2719	2055

		GRADE 5 SET A	GRADE 5 SET B	GRADE 9 SET A	GRADE 9 SET B
1986	SCHOOLS	123	123	119	119
	STUDENTS	1346	1301	1220	1182

	FIRST-YEAR BIO	FIRST-YEAR CHEM	ADVANCED SCIENCE BIO	ADVANCED SCIENCE CHEM	ADVANCED SCIENCE PHYS.
SCHOOLS	118	118	43	40	34
STUDENTS	2582	2344	659	539	485

SOME RESULTS FROM THE SECOND IEA SCIENCE STUDY

Rodney L. Doran

To summarize the results of a study which encompassed 11 different populations involving over 1000 schools and over 20,000 students in a few minutes is impossible. I have chosen to highlight a few major effects in this presentation, namely the anchor effect (the differences in achievement between different grade levels or populations), the relatively low achievement of first-year physics students, an international comparison (comparison of U.S., Japanese, and English results) of advanced science testing, and a sampling of results from the process testing (laboratory practical items) in Grades Five and Nine.

The "anchor" effect. The "anchor" effect is the comparison between students of different grade levels or in different courses on a common set of items. On 20 common items, the Grade Nine students had, on the average, higher scores by 19% than the Grade Five students. On the 15 common items, Grade 12 physics students achieved a 22% higher score than the Grade 9 students. However, the nonscience students had the same mean score as the 9th grade students. In general, these findings were expected. Any other results would have raised serious questions about the validity of the tests, the sampling, or both. However, the comparative scores on anchor items seem to indicate that 12th graders who are not studying science are at the same level of achievement as 9th graders who probably are studying science.

Low achievement in physics. The achievement of first-year physics students in the U.S. is of particular interest because it is the final science course for most science-oriented students. Approximately 15% of U.S. high school graduates have completed such a course. (NCES, 1984) Physics teachers who reviewed the 35 test items of the physics test judged them as generally part of the material they taught. The average opportunity-to-learn (OTL) rating for these items was 69%. Nevertheless, this international physics test was very difficult for these U.S. students. The average score was 34% correct (a mean of about 12 of 35 items correct). A further perspective on the difficulty of the test is the fact that the easiest item for U.S. students was one on which 69% of the students chose the correct answer. This item was in the mechanics area, a large unit in the U.S. physics courses. Further, only four items were answered correctly by more than 50% of the U.S. students. On 11 items, less than 25% of the sample chose the correct answer. On the whole, these results suggest that a further examination of U.S. first-year physics courses and its teaching needs to be conducted.

Some international comparisons. Perhaps of most interest are the comparisons of achievement results of the students studying science at the high school level compared to several other countries. Preliminary results from students in U.S. first-year and second-year courses in biology, chemistry, and physics and 12th grade students not studying science have been compared with results from Japan, England, and an international composite of nine countries. The U.S. students studying biology, chemistry, and physics for a second year consistently performed at lower levels than students from the other participating countries. The U.S. samples performed at a level of about 41% to 44% correct. In the other countries, performance ranged from 48% to 73%

correct. The U.S. nonscience sample had lower means than the other countries, but the difference was about 10%. In the U.S. the second-year physics enrollment is very small--attracting approximately 1.5% of high school graduates. (NCES, p. 18) Second-year chemistry enrollment is approximately 4%, while second-year biology (including advanced biology, botany, and zoology) enrolls about 15% of high school graduates. (NCES, p. 18) Considering the selectivity of these courses, these results are especially disappointing. Many had expected that our very best science students could compare well with their counterparts in other countries. It appears that their expectations were not realized.

TABLE 4

Preliminary Comparisons of U.S. Results with Japan, England, and an International Composite (Percent Correct)

	Biology	Chemistry	Physics	NonScience
U.S. First Year*	34	27	34**	48**
U.S. Second Year	44	41	44	--
Japan	48	62	59	58
England	71	73	58	60
International (Nine Countries)	62	55	51	56
Number of Items	13	13	26	30

*Most of the students in first-year biology were in the 10th grade, while most of the first-year chemistry students were 11th graders. Similarly, most of the first-year physics students and the nonscience students were in the 12th grade.

**These scores were based on data collected in 1983. All other U.S. results were from the 1986 test administration.

Science process achievement. In the First IEA Science Study (FISS), two countries (Israel and Japan) tested with process (manipulative, practical) tasks. (Kojima, 1974) In SISS, six countries (Israel, Japan, Korea, Hungary, Singapore, and the U.S.A.) conducted this testing with youngsters in Grades Five and Nine. These tasks were developed to represent the range of content and skills emphasized at those levels. Each child responded to a set of questions related to three tasks. Two sets ("A" and "B") were prepared. Half the class completed the Set A tasks and the other half completed the Set B tasks. The correlation of scores in these process tests with the achievement tests were .30 and .43 for Grade 5 Sets A and B, respectively. For Grade 9, the values for Grade 9 Sets A and B were .50 and .54, respectively. These values can be interpreted to mean that there is a moderate, positive relationship with paper and pencil achievement tests, but that process tests are substantively unique, separate measures of science performance. The depth of information that will be available is indicated in the results from one item from the Grade Five test (see below). In this item, the students were asked to assemble an electrical circuit, test objects to find whether they conduct electricity, and formulate conclusions about the objects' conductivity. It appears that students are relatively good at assembling and testing, but less able to reason and explain these observations.

ELECTRICAL CIRCUIT AND TESTING (5A3)

	SCORE	PECENT RESPONDING
ASSEMBLY OF CIRCUIT		
Bulb lights & testing area	2 pts	76%
One of the above	1 pt	19%
Unable to assemble circuit correctly	0 pts	3%
No response	0 pts	2%
TESTING OF OBJECTS		
Correct results for 3 or 4 objects	2 pts	70%
Correct results for 1 or 2 objects	1 pt	7%
No correct results	0 pts	19%
No response	0 pts	3%
CONCLUSION/REASON ABOUT CONDUCTORS		
Identify conductors & cite evidence	2 pts	9%
One of the above	1 pt	51%
No correct identification of evidence	0 pts	25%
No response	0 pts	15%

TABLE 5

RESULTS FROM PROCESS TESTING IN U.S.
GRADE 5

		Points	Percent Responding
5A1	Color Changes in Solution		
	Part 1 - describe change	(2 pts)	57%
	Part 2 - explain change	(2 pts)	11%
5A2	Observe Similarities and Differences in Toy Animals		
	Part 1 - describe similarities	(3 pts)	20%
	Part 2 - describe differences	(3 pts)	49%
5A3	Electrical Circuit and Testing		
	Part 1 - Assembly of Circuit	(2 pts)	76%
	Part 2 - Testing of Objects for Conductivity	(2 pts)	70%
	Part 3 - Conclusion and Reason for Conductors	(2 pts)	9%
5B1	Temperature of Water Mixture		
	Part 1 - Measure Temperature of the Water	(2 pts)	86%
	Part 2 - Predict Temperature of New Mixture	(1 pt)	7%
5B2	Diffusion of Substance in Water		
	Part 1 - Observe and Describe Change	(2 pts)	26%
	Part 2 - Explain Change	(1 pt)	26%
5B3	Testing Seeds for Oil Content		
	Part 1 - Describe Stain from Oily Swab	(1 pt)	82%
	Part 2 - Plan for Testing Seeds	(1 pt)	43%
	Part 3 - Identification/Explanation of Seeds with Oil	(2 pts)	41%

TABLE 5, continued

RESULTS FROM PROCESS TESTING IN U.S.
GRADE 9

			Points	Percent Responding
9A1	Determining Electrical Circuit			
	Part 1 -	Assemble Testing Circuit	(1 pt)	90%
	Part 2 -	Perform Test and Record Results	(2 pts)	57%
	Part 3 -	Identify Correct Schematic Diagram	(2 pts)	67%
9A2	Testing Solutions for Ac.ds/Bases			
	Part 1 -	Phenolphthalein - Observe/ Describe Changes	(1 pt)	77%
	Part 2 -	Conclusion from Phenolphthalein Test	(1 pt)	47%
	Part 3 -	Plan for Additional Testing	(1 pt)	64%
	Part 4 -	Observe/Record Results from Testing	(2 pts)	36%
	Part 5 -	Identify/Explain Acid and Base	(2 pts)	17%
9A3	Testing Solutions for Starch			
	Part 1 -	Determine Plan for Starch Testing	(1 pt)	69%
	Part 2 -	Observe/Record Results from Testing	(2 pts)	35%
	Part 3 -	Identify Starch Solutions/ Explain Reasons	(2 pts)	31%
9B1	Determine Density of Object			
	Part 1 -	Finding the Mass of Object	(2 pts)	51%
	Part 2 -	Measuring the Volume of Object	(2 pts)	29%
	Part 3 -	Calculate Density (with correct units) Given Formula	(2 pts)	8%
9B2	Color Chromatography (U.S. Option Item)			
	Part 1 -	Observe Rates of Movement	(1 pt)	74%
	Part 2 -	Describe Changes in Colored Dots	(2 pts)	95%
	Part 3 -	Explain Changes in Black Dot	(1 pt)	12%
9B3	Testing for Sugar and Starch (US Option Item)			
	Part 1 -	Determine Plan for Testing Three Solutions	(1 pt)	40%
	Part 2 -	Perform Tests and Record Results	(2 pts)	21%
	Part 3 -	Identify Sugar Sample and Give Reason	(2 pts)	42%
	Part 4 -	Identify Starch Sample and Give Reason	(2 pts)	40%

SCIENCE EDUCATION IN THE 1970S AND 1980S: WHAT CHANGES HAVE TAKEN PLACE?

Edith Y.T. Chang

There have been relatively few comparisons of science achievement over time. In the Second IEA Science Study, some of the items that were used in 1970 were used again in 1986. These were called "bridge" items. There were 21 common "bridge" items in Grade 5, 19 in Grade 9 and 18 in Grade 12. The presence of these "bridge" items in the 1970 and 1986 tests makes it possible to compare the achievement scores in 1970 and the 1980s. Information about the 1970 results are reported in Comber and Keeves (1973) and Wolf (1977).

It should be recognized that scores on achievement tests are affected by the response rate and by differences in data collection procedures. In 1986, slightly different procedures were used than were used in 1970, which may have led to a higher response rate in 1986. Thus, it may be that a broader range of students was tested in 1986 than in 1970. Also, the length of the testing period in 1986 was slightly shorter than in 1970, and this may have led to slightly lower mean scores in 1986.

In this paper the scores in 1970 and the 1980s are compared, the changes in the scores in the life and physical sciences are reported, the results on items classified as "process" and "non-process" are examined, and the possible effects of visual illustrations are explained. Finally, some recommendations are made for future studies of achievement over time.

The fifth graders in 1986 scored about the same as the 1970 students on the 21 bridge items. See Table 6. The lack of increase in scores in 1986 may have been affected by the following statistics taken from the student questionnaire: 24.7% of the fifth graders were watching more than six hours of TV per day in 1986 as compared to 7% who watched more than six hours in 1970. But, the correlation was not significant, which suggested that televisions might have been turned on but not watched or that the students might have been learning from some of the programs. The fifth graders in 1986 were also spending one hour less per week on their homework than their counterparts in 1970. But, they appeared to be enjoying school more than their 1970 counterparts.

TABLE 6
Comparison of All Bridge Item Results in Percent Correct for 1970 and 1986
Grade 5 (Population 1)

	1970	1986	1986 minus 1970
N	3504	2838	
Mean	55.9	56.3	+0.4

The ninth grade students in 1986 scored lower than those in 1970. See Table 7. One reason for this decline might be attributed to the decrease in testing time, which was 10 minutes less in 1986 than in 1970, and on the 35 test items as compared to 30 test items in 1970. Another reason may have been the application of IEA procedures for sampling and data collection in 1970 while the Research Triangle Institute (RTI) sample in 1986 was possibly broader because of their more persistent approach to sampling and data collection. Participation of schools in 1986 was also made more attractive by reducing the number of testing periods from three to two. Finally, the 1986 students simply may not have been achieving as well as those in 1970. On the average, they spent about two hours less per week on their homework as compared to their 1970 counterparts and they exhibited a correspondingly lower aspiration for further education of about four years in comparison to the six years of their 1970 counterparts. But, they enjoyed school more than their 1970 counterparts.

TABLE 7
Comparison of All Bridge Item Results in Percent Correct for 1970 and 1986
Grade 9 (Population 2)

	1970	1986	1986 minus 1970
N	2339	2518	
Mean	53.7	49.2	-4.6

The 12th grade students in 1983 who were taking physics and those not taking any science all scored higher than their 1970 counterparts. See Table 8. These differences are in the moderate range; +4.1% for 3P and +2.8% for 3N. At both testings, the 3P students had higher scores than the 3N students (19.1% and 17.1%).

TABLE 8

Comparison of All Bridge Item Results in Percent Correct in 1970 and 1983 for Students Taking Physics (Population 3P) and Students not Studying a Science (Population 3N)

	3P Mean 1970	3P Mean 1983		3N Mean 1970	3N Mean 1983
N	563	2719		1377	2055
Bridge Total	53.2	57.3		35.4	38.2
1983 minus 1970	+4.1			+2.8	
1970 3P minus 1970 3N			+19.1		
1983 3P minus 1983 3N			+17.1		

In 1986 the fifth graders scored the same as their 1970 counterparts on both the life and physical science items. In terms of the differences in achievement scores on life and physical science items, the intervals were the same, suggesting that the fifth graders scored higher on the physical science items in 1986 as was also the case in 1970. See Table 9.

TABLE 9

Life and Physical Science Item Mean
Percent Correct Comparisons in 1970 and 1986
Grade 5 (Population 1)

	1970	1986	1986 minus 1970
Life Science (8 Items)	50.4	51.3	+0.9
Physical Sci. (13 Items)	59.3	59.3	0.0
Life Science minus Physical Science	-8.9	-8.0	

In 1986 the ninth grade scores were slightly lower than their 1970 counterparts on the life and physical science items. But, the ninth grade students scored higher on the life science items in 1986 than on the physical science items, as was the case in 1970. See Table 10.

TABLE 10

Life and Physical Science Item Mean
Percent Correct Comparisons in 1970 and 1986
Grade 9 (Population 2)

	1970	1986	1986 minus 1970
Life Science (7 Items)	59.6	56.3	-3.6
Physical Science (12 Items)	50.5	45.0	-5.5
Life Science minus Physical Science	+9.1	+11.3	

First year physics students (3P) and 12th grade students who were not studying science (3N) were tested in 1983. The 12th grade students in 1983, both those taking physics (3P) and those not taking any science (3N), scored higher than their 1970 counterparts. Both groups scored higher on the life science items than on the physical science items. This result was somewhat surprising for the 12th grade students who were taking physics, where a higher score on the physics items might have been the expected result. See Tables 11 and 12.

TABLE 11
Life and Physical Science Item Mean
Percent Correct Comparisons in 1970 and 1983
for Students Studying Physics (Population 3P)

	1970	1983	1983 minus 1970
Life Science (6 Items)	58.2	62.6	+4.4
Physical Science (12 Items)	50.7	54.6	+3.9
Life Science minus Physical Science	+7.5	+8.0	

TABLE 12
Life and Physical Science Item Mean
Percent Correct Comparisons in 1970 and 1983
for Students not Studying any Science (Population 3N)

	1970	1983	1983 minus 1970
Life Science (6 Items)	42.4	45.4	+3.0
Physical Science (12 Items)	32.0	34.7	+2.7
Life Science minus Physical Science	+10.4	+10.8	

The scores of fifth grade students in 1986 remained essentially the same as in 1970. But, the difference in scores between process and nonprocess items narrowed from 1970. See Table 13.

TABLE 13

Process and NonProcess Item Mean Percent Correct Comparisons in 1970 and 1986, Population 1 (Grade 5)

	1970	1986	1986 minus 1970
Process (11 Items)	53.1	54.9	+1.8
NonProcess (10 Items)	59.0	57.8	-1.2
Process minus NonProcess	-5.9	-2.9	

For the ninth graders, there was a decline on both process and nonprocess items in 1986 as compared to 1970. The difference in scores between the process and nonprocess items was essentially the same in 1986 and 1970. See Table 14.

TABLE 14

Process and NonProcess Item Mean Percent Correct Comparisons in 1970 and 1986, Population 2 (Grade 9)

	1970	1986	1986 minus 1970
Process (12 Items)	53.8	48.8	-5.0
NonProcess (7 Items)	54.0	49.8	-4.2
Process minus NonProcess	-0.2	-1.0	

The 12th grade scores for both students taking physics (3P) and those not studying any science (3N) improved from 1970. The 3P students in 1983 scored the same in both process and nonprocess areas and higher than their 1970 counterparts. The 3N students found the process items easier than the nonprocess items. See Tables 15 and 16.

TABLE 15

Process and NonProcess Item Mean
Percent Correct Comparisons in 1970 and 1983
for Students Studying Physics (Population 3P)

	1970	1983	1983 minus 1970
Process (9 Items)	52.3	56.9	+4.6
NonProcess (9 Items)	54.1	57.6	+3.5
Process minus NonProcess	-1.8	-0.7	

TABLE 16

Process and NonProcess Item Mean
Percent Correct Comparisons in 1970 and 1983
for Students not Studying any Science (Population 3N)

	1970	1983	1983 minus 1970
Process (9 Items)	36.8	39.3	+2.5
NonProcess (9 Items)	34.1	37.2	+3.1
Process minus NonProcess	+2.7	+2.1	

Student scores in 1986 on the items with visual illustrations, when compared to their 1970 counterparts, remained the same or slightly lower. Items were classified as "visual" when the students were required to obtain information from the illustrations (e.g., graphs, tables, and pictorial illustrations). The visual items were classified into three categories:

1. Visual illustrations that clarified the question.

2. Visual illustrations that contained information relating to the questions or answers and that required inference skills.

3. Visual illustrations that did not explain the question or the answer.

On "visual" items the fifth graders in 1986 scored about the same as their 1970 counterparts. Over the same time period, the fifth grade student scores increased on the "nonvisual" items. The 1986 fifth graders also scored essentially the same on visual and nonvisual items, while their 1970 counterparts scored higher on the visual items. See Table 17.

TABLE 17

Mean Percent Correct Comparisons of Items With and Without Visual Illustrations in 1970 and 1986 Grade 5 (Population 1)

	1970	1986	1986 minus 1970
Visual (9 Items)	54.7	55.6	+0.9
NonVisual (12 Items)	52.4	56.8	+4.4
Visual minus NonVisual	+2.3	-1.2	

The 9th grade students in 1986 scored lower on the visual items when compared to 1970. The visual items were the group with the largest decline for the 1986 ninth graders when compared with 1970.

The discrepancy between items with and without visual illustrations was also the largest in the ninth grade sample. The classification of the visual illustrations showed that most of the visual items required inference skills. This, in Piagetian terms, might require formal operations, and might be difficult for most 9th grade students, some of whom might still be in the concrete operational stage. See Table 18.

TABLE 18

Mean Percent Correct Comparisons of Items With and Without Visual Illustrations in 1970 and 1986 Grade 9 (Population 2)

	1970	1986	1986 minus 1970
Visual (11 Items)	48.8	42.4	-6.4
NonVisual (8 Items)	60.8	58.8	-2.0
Visual minus NonVisual	-12.0	-16.4	

The 12th grade students studying physics (3P) and those not studying any science (3N) in 1983 scored higher than the 1970 students on both the visual and nonvisual items. See Tables 19 and 20. Both groups performed better on the visual items than the nonvisual items.

TABLE 19

Mean Percent Correct Comparisons of Items With and Without Visual Illustrations in 1970 and 1983 12th Grade (Population 3P)

	1970	1983	1983 minus 1970
Visual (10 Items)	56.4	60.2	+3.8
NonVisual (8 Items)	49.2	53.6	+4.4
Visual minus NonVisual	+7.2	+6.6	

TABLE 20

Mean Percent Correct Comparisons of Items With and Without Visual Illustrations in 1970 and 1983 12th Grade (Population 3N)

	1970	1983	1983 minus 1970
Visual (10 Items)	37.6	40.0	+2.4
NonVisual (8 Items)	32.8	36.1	+3.3
Visual minus NonVisual	+4.8	+3.9	

The average test scores improved from 5th grade to 9th grade and 9th grade to 12th grade. Some test items, called "anchor items," were used with more than one population. This made it possible to compare the scores at different grade levels. Some of these items were also used in 1970 and are referred to as "bridge/anchor" items. There were nine bridge/anchor items between Grades 5 and 9 and seven bridge/anchor items between Grades 9 and 12. The gain in scores from fifth grade to ninth grade was greater in 1970 than in 1986. See Table 21.

TABLE 21

Comparison of Bridge/Anchor Item Scores
in Mean Percent Correct for 1970 and 1986
Between Grade Five (Population 1) and Grade 9 (Population 2)

GRADE	1970	1986
5	55.1	53.8
9	71.4	65.9
Grade 9 minus Grade 5	+16.3	+12.1

The gain in scores from 9th grade to 12th grade physics (3P) was about the same in 1983 as in 1970.

In 1983 and 1970, 12th grade students who were not studying science (3N) scored about the same on common bridge/anchor items as did students in the 9th grade. This suggests that students who do not study science do not raise their scores on science items that were also used in the 9th grade. See Table 22.

TABLE 22

Comparison of Bridge/Anchor Item Scores in Mean Percent Correct for 1970, 1983, and 1986 Between Grade 9 (Population 2) and Grade 12 (Populations 3P and 3N)

GRADE	1970	1983
9	42.8	45.6
12 (3P)	64.0	66.6
12 (3N)	44.0	45.8
12 (3P) minus Grade 9	+21.2	+21.0
12 (3N) minus Grade 9	+1.2	+0.2

Word knowledge ability declined but science ability improved. A comparison between the Word Knowledge Test and science achievement showed a decline in the 1983 student verbal ability from 1970, which is consistent with SAT trends over the same time period. Verbal scores have been shown to be a good predictor of science achievement. Therefore, it was interesting to see the improvement of the student science achievement scores despite no change (Grade 5) or declining (Grade 9 and 12) verbal ability. See Tables 23, 24, 25, and 26.

TABLE 23

Comparison of 1983 and 1970 Word Knowledge Test Results with Science Achievement Test Results in Mean Percent Correct Grade 5 (Population 1)

	Word Knowledge Test						Science Achievement					
	Boys		Girls				Boys		Girls			
	1983	1970	1983	1970	1983 Boys minus Girls	1970 Boys minus Girls	1983	1970	1983	1970	1983 Boys minus Girls	1970 Boys minus Girls
	73.5	73.0	73.5	72.5	0.0	+0.5	64.2	58.3	58.3	53.8	+5.9	+4.5
1983 minus 1970	+0.5		+1.0				+5.9		+4.5			

TABLE 24

Comparison of 1983 and 1970 Word Knowledge Test Results with Science Achievement Test Results in Mean Percent Correct Grade 9 (Population 2)

	Word Knowledge Test						Science Achievement					
	Boys		Girls				Boys		Girls			
	1983	1970	1983	1970	1983 Boys minus Girls	1970 Boys minus Girls	1983	1970	1983	1970	1983 Boys minus Girls	1970 Boys minus Girls
	67.3	70.6	66.8	69.9	+0.5	+0.6	61.2	57.5	53.7	50.7	+7.5	+6.8
1983 minus 1970	-3.3		-3.1				+3.4		+3.5			

TABLE 25

Comparison of 1983 and 1970 Word Knowledge Test Results with Science Achievement Test Results in Mean Percent Correct Population 3P

	Word Knowledge Test						Science Achievement					
	Boys		Girls				Boys		Girls			
	1983	1970	1983	1970	1983 Boys minus Girls	1970 Boys minus Girls	1983	1970	1983	1970	1983 Boys minus Girls	1970 Boys minus Girls
	67.1	68.8	67.4	70.0	-0.4	-1.2	60.9	56.0	51.2	44.5	+9.7	+11.5
1983 minus 1970	-1.7		-2.6				+4.9		+6.7			

TABLE 26

Comparison of 1983 and 1970 Word Knowledge Test Results with Science Achievement Test Results in Mean Percent Correct Population 3N

	Word Knowledge Test						Science Achievement					
	Boys		Girls				Boys		Girls			
	1983	1970	1983	1970	1983 Boys minus Girls	1970 Boys minus Girls	1983	1970	1983	1970	1983 Boys minus Girls	1970 Boys minus Girls
	56.9	60.8	58.9	62.0	-2.0	-1.2	40.8	39.6	36.2	32.6	+4.6	+7.0
1983 minus 1970	-3.9		-3.1				+1.2		+3.6			

Some procedural recommendations for future studies of science achievement over time.

1. Keep the instruments the same. While the temptation to improve items is great, if comparisons are to be made with testing results of an earlier time it is essential that the items not be changed.

2. It has been suggested that the 1980s reproduction of the instruments was clearer and easier to read than the 1970 version. Again, while the temptation to improve the presentation of the instruments is great, every effort must be made to repeat exactly the original instruments with respect to the format of print layout, the size of print and illustrations, the sequence of items, and the sequence of the distractors.

3. The selection of the sample and the method of persuading teachers and schools to participate should be the same.

4. The administration of the tests should be the same in terms of the testing time and the use of trained, outside administrators versus selected, in-school, test administrators.

5. Even if all and every aspect is kept the same, items will change in difficulty over time, for example, an item requiring students to estimate the amount of time for a rocket ship to reach the moon proved to be much easier in 1970 than in the 1980s.

6. In the interpretation of the comparisons in achievement over time, one must consider the impact of changes in the retentivity of students in schools. For example, if retentivity of a school system doubles then surely there will be an impact on the scores.

7. The last consideration would be changes in the composition of a school population. For example, an influx of students who do not speak the language of instruction will affect the results.

This study is unparalleled because of the sheer number of total bridge items kept over the past 16-year span, the size of the total samples, and the number of participating countries.

SEX AND SCIENCE ACHIEVEMENT

Eve Humrich

In looking at sex differences in science achievement in the U.S., a few important issues have come to light. The first and most obvious is that sex differences have been found at every grade level and in every subject area in the written science achievement tests. This sex difference always favors males. Secondly, the fact that some female students have female science teachers does not appear to increase their level of achievement. In fact, in some cases, having a female teacher has a negative correlation with science achievement. Finally, we do have some more optimistic news. It would seem that there is no sex difference to be found in the results of the manipulative tests of process skills which were administered to both fifth and ninth graders in Phase 2 of our study.

With the Second IEA Science Study, we hoped to find a smaller difference between the sexes in science achievement than had been found in the first study (Comber & Keeves, 1973; Wolf, 1977). The reasoning behind this hypothesis was that not only had 13 years gone by, but also that those years saw the rise of the Feminist Movement, and saw many wives and mothers forced to become working wives and mothers. The idea that girls "get through" school only to marry and raise a family has changed dramatically, and it has become an economic necessity for many families to have both spouses earning an income. In an age of science and technology, therefore, it was assumed that a proportionate number of females would be seeking education and employment in the areas of expanding opportunities. It was hoped that, realizing this growing trend, schools would become more effective in educating girls in science. Apparently, this has not been the case.

The sex difference in science achievement remains the same. Although some previous studies have found no significant difference between girls and boys at Grade 5, in the IEA survey we found a difference of 5.2%. At the ninth grade level, where many studies have found sex differences in science achievement to become more noticeable, we found an increase of only 1% from the fifth grade level, making the difference 6.2%. This increase was not considered to be significant (using a 2% criterion). The literature is divided in reporting sex differences at higher grade levels. Some find that the differences peak at the junior high level, then decrease through high school (Vestin, 1975; Steinkamp & Maehr, 1984). Other studies report a continuing widening of the "gender gap" all through high school (Backman, 1972; Keeves, 1973; Haertel et al., 1981; Zerega et al., 1986). The results of the SISS are reported on the next page (Table 27).

As can be seen, for the non-science population, the gap does shrink to only 3.8% at Grade 12. For the science populations, the outcomes vary. The smallest sex difference can be found in first-year biology (3.1%), which helps it retain its notoriety as a "feminine" science. The difference in achievement for both first-year chemistry and physics is slightly greater (5.4%), but still in the same range as the fifth and ninth grade differences.

For second-year, or advanced science, the sex differences increase by approximately 2% in each subject area, with biology remaining the subject with the smallest difference between sexes. Overall sex differences remain fairly constant, in the 5% to 7% range, with the exceptions of Biology 1 and the non-science population (Figure 1).

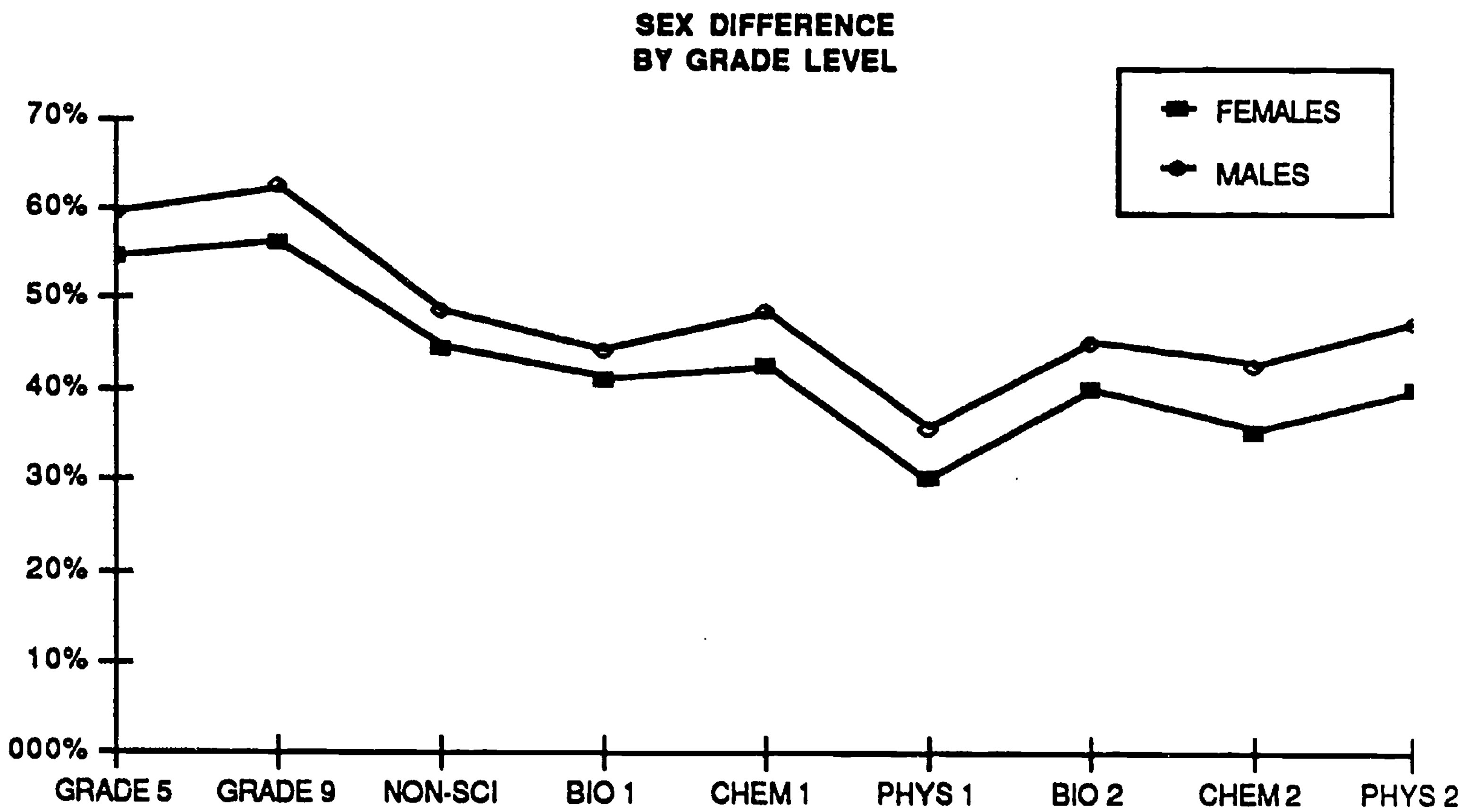
TABLE 27

Population Descriptions and Percent of Each Sex Scoring Correctly

GRADE	NUMBER FEMALES	NUMBER MALES	PERCENT SCORING CORRECTLY FEMALES	PERCENT SCORING CORRECTLY MALES	PERCENT SCORING CORRECTLY DIFFERENCE
5	1426 (51.0%)	1372 (49.0%)	54.6%	59.8%	5.2%
9	1204 (48.1%)	1301 (51.9%)	56.4%	62.5%	6.2%
NON-SCI.	1131 (56.0%)	890 (44.0%)	45.0%	48.8%	3.8%
BIO 1	1321 (53.5%)	1148 (46.5%)	41.4%	44.5%	3.1%
CHEM 1	1091 (47.6%)	1201 (52.4%)	43.3%	48.7%	5.4%
PHYS 1	997 (37.1%)	1687 (62.9%)	30.5%	35.9%	5.4%
BIO 2	369 (58.5%)	262 (41.5%)	40.2%	45.4%	5.2%
CHEM 2	231 (43.6%)	299 (56.4%)	35.7%	42.8%	7.1%
PHYS 2	101 (21.4%)	372 (78.6%)	40.2%	47.6%	7.4%

NOTE: Mean scores in this table are based on the complete tests taken by the students. This includes national option items, rotated items for Grades 5 and 9, and both the core and specialty tests for the non-science population.

FIGURE 1



Comparing our U.S. results with some of those obtained internationally, we do find inconsistencies (Table 28). In England, significant sex differences were found for all populations tested except for physics (A-level) (Keys, 1986). In Israel, Tamir (1987) reports a significant sex difference in 12th grade physics, but not in either biology or chemistry. In Japan, there was a 1% difference between the sexes at the 5th grade level, and 5% at the ninth grade level on the SISS core test (Miyake, 1985). In physics, the sex difference in Japan was also only 1%, but for the non-science population it was slightly over 10% (Jacobson, et al., 1986). The question now is can we discover what England and Japan are doing right in physics and what Israel is doing right in biology and chemistry?

TABLE 28

International Sex Differences

LEVEL	U.S.	ENGLAND	JAPAN	SWEDEN
POP 1*				
BOYS	64.4	59.9	66.1	60.9
GIRLS	59.4	55.7	64.0	57.5
DIFF:	5.0%	4.2%	2.1%	3.4%
POP 2*				
BOYS	65.3	59.7	68.8	64.3
GIRLS	59.3	52.1	64.5	58.9
DIFF:	6.0%	7.6%	4.3%	5.4%
NON-SCI**				
BOYS	46.2	NA	66.7	NA
GIRLS	41.4		56.3	
DIFF:	4.8%		10.4%	
BIOLOGY***				
BOYS	45.4	64.6	NA	53.6
GIRLS	40.2	62.4		56.5
DIFF:	5.2%	2.2%		-2.9%
CHEMISTRY***				
BOYS	42.8	71.4	NA	47.5
GIRLS	35.7	65.9		46.6
DIFF:	7.1%	5.5%		0.9%
PHYSICS***				
BOYS	47.6	58.4	NA	53.3
GIRLS	40.2	57.4		46.9
DIFF:	7.4%	1.0%		6.4%

NOTE: Mean scores in this table are based solely on international items.

---

* Core test plus international rotated items.
** Core test only.
*** Specialty science test only.

NA Data not available

We were curious to find whether female teachers played the part of positive role models to girls in science. Cross-tabulations were run of mean scores on the achievement tests for girls and boys having male or female teachers. Results are presented in Table 29.

TABLE 29

Influence of Teacher Sex on Student Achievement
(In Percent Correct)

LEVEL	MALE TEACHER	FEMALE TEACHER	DIFFERENCE
GRADE 5 (24 ITEMS)			
BOYS	58.3	57.9	0.4
GIRLS	52.5	52.9	-0.4
DIFFERENCE	5.8	5.0	
GRADE 9 (30 ITEMS)			
BOYS	59.0	57.7	1.3
GIRLS	53.0	50.3	2.7
DIFFERENCE	6.0	7.4	
BIOLOGY 1 (30 ITEMS)			
BOYS	43.3	47.3	-4.0
GIRLS	41.3	42.7	-1.4
DIFFERENCE	2.0	4.6	
BIOLOGY 2 (30 ITEMS)			
BOYS	42.7	49.7	-7.0
GIRLS	38.7	42.0	-3.3
DIFFERENCE	4.0	7.7	
CHEMISTRY 1 (30 ITEMS)			
BOYS	48.7	48.7	0
GIRLS	43.7	43.3	0.4
DIFFERENCE	5.0	5.4	
CHEMISTRY 2 (30 ITEMS)			
BOYS	42.3	38.0	4.3
GIRLS	34.7	33.3	1.4
DIFFERENCE	7.6	4.7	
PHYSICS 1 (35 ITEMS)			
BOYS	36.6	32.6	4.0
GIRLS	31.7	28.3	3.4
DIFFERENCE	4.9	4.3	

As can be seen, at the fifth grade level teacher sex had no significant influence on student achievement. At the ninth grade level, male teachers seemed to have a more positive influence on student achievement than did female teachers. This was especially true for girls by about 3%, while for boys it was only about 1%.

Encouragingly (although stereotypically), both girls and boys scored higher in the biology achievement tests if they had a female teacher. At the first-year level, the difference was 3% overall, and at the advanced level, the difference was 5% overall. At both levels, the achievement difference was more pronounced for boys than for girls. In Biology 1, boys with female teachers scored 4% higher than did boys with male teachers (girls scored 1% higher). In Biology 2, boys with female teachers scored 7% higher than those with male teachers (girls scored 3% higher).

The chemistry tests produced different results. At the first-year level, teacher sex appeared to make no difference to science achievement for either girls or boys. At the advanced level, male teachers had a more positive influence on achievement for students of both sexes, although the difference was greater for boys. Similar results were obtained in Physics 1, with students of male teachers achieving higher than students of female teachers.

Unfortunately, there were too few female physics teachers involved in the survey to provide sufficient data for valid comparisons.

The overall results were unexpected. We had hoped to find a relationship between teacher sex and student achievement, but none was detected. Instead, there appeared to be a relationship between teacher sex and subject area. Female teachers were much stronger positive influences for both girls and boys in biology than they were in chemistry or at the early grades. Biology, therefore, seems to retain its "femininity" not just in terms of there being a smaller sex difference in achievement in this area than in the physical sciences, but also in terms of female teachers encouraging higher achievement for both girls and boys. It may be that female teachers feel more comfortable with biology than they do with the physical sciences, or that they have a stronger background in this area. Further research is being conducted to evaluate teacher variables such as courses taken in college and length of time spent teaching the various science subjects.

We feel we have uncovered a possible path to encouraging girls in science. In Phase 2 of our study, a manipulative process test was administered to both fifth and ninth grade students. According to some of the literature, girls do not do well in manipulative or discovery-type tasks (Harding, 1973); however, other research supports the opposite hypothesis. Johnson & Murphy (1986) have discussed their own Assessment of Performance Unit (APU) findings related to the superior process skills of girls, and also described similar findings by the British Columbia Science Assessment in 1979.

In the U.S., we were pleased to find evidence supporting the hypothesis that girls do excel in manipulative process tasks. The results are presented in Table 30.

TABLE 30

Manipulative Process Test
(Percent Scoring Correctly)

ITEM	GRADE 5 GIRLS	GRADE 5 BOYS	GRADE 9 GIRLS	GRADE 9 BOYS
SUBTEST A				
A1	50.0%	52.2%	70.7%	76.9%
A2	64.7%	61.3%	39.6%	40.3%
A3	62.1%	67.0%	44.0%	46.3%
MEAN (A)	60.1%*	61.1%	50.0%	52.8%
SUBTEST B				
B1	44.5%	47.2%	38.2%	44.6%
B2	44.4%	45.2%	71.8%	67.5%
B3	60.4%	52.7%	45.3%	44.0%
MEAN (B)	49.8%*	48.5%	49.0%	49.8%

* WEIGHTED SCORES, SINCE EACH ITEM HAS DIFFERENT POINT VALUE.

As can be seen, for fifth grade students there was no significant difference between boys' and girls' scores on either subtest of the process tests. At the ninth grade level there was a 3% difference favoring the boys on Subtest A and no difference on Subtest B. The items with the greatest difference favoring boys were two items (one at each grade level) concerning bulbs and batteries, and one item requiring the measurement of mass and volume and the calculation of density. Girls achieved higher than did boys on an item requiring the planning of an experiment to test seeds for oil content, and on an item dealing with paper chromatography.

If we fall back on the argument that girls are more concrete and boys are more abstract in their thinking styles, it could be hypothesized that these "hands-on" tests favor girls, or at least compensate for their "inability" to think abstractly. However, interpretation of data plays a significant part in the SISS test, and girls do very well on the sections that demand interpretation and/or explanation. Further evidence to support female skills in reasoning abilities are the results obtained by Mattheis et al. (1985) on a paper and pencil type process test, the Test of Integrated Process Skills (TIPS), which was administered to over 7,000 junior high school students in Japan and in North Carolina in the U.S. No sex difference was found in either country.

Perhaps science process skills are not quite so sex-role stereotyped as is science content. Kelly (1978) and Kahle (1983), among others, have suggested that science be presented to students in a more equitable, although _not_ a more simplified manner. In standardized achievement tests, it may not be the intelligence of girls that is being assessed, but merely their ability to recall facts which they (in many cases) consider boring and/or useless in terms of future education or career goals. Since socialization plays an important part in defining sex-roles (Baker, 1980), and girls are able to perceive the expectations of others in their environment (Bornstein, 1982; Sadker & Sadker, 1982), they may realize that their parents, teachers, guidance counselors, and peers feel they either should or should not participate in certain courses. In fact, it has been hypothesized that there is subtle pressure against girls in science all through elementary and junior high school (Lipkin & Sadker, 1982). They must take science, but everyone around them projects the feeling that they are not expected to do well. It may be that not much effort is made by teachers to entice girls into finding science interesting and enjoyable (Dickson, 1986). The "hidden curriculum" of the classroom (Bornstein, 1982) places girls in an insulated class-within-a-class. Perhaps then, girls are differentially eased out of the fun part of science, such as laboratory experiences. They might tend, therefore, to fall back on what they have always learned is appropriate classroom activity for females--studying textbook facts. This may put them at a disadvantage since they will have an insufficient foundation of practical experience on which to base the facts they have memorized. How easy is it, and how much pleasure can a person get out of memorizing formulas that she will never get a chance to use?

Manipulative process testing allows girls to participate in the fun part of science. They are tested individually so no boy can take over the experiment or tell them they are doing something wrong. Girls may feel freer to indulge in risk-taking when no one else is ready to pounce on their mistakes. Teaching science via process tasks may also be the way to encourage girls to study science. This does not mean continuance of present-day laboratories, but rather individual exercises in observation, recording data, stating of hypotheses, etc., with no right or wrong answers to be looked up in the back of the book.

Further research is necessary both in the area of manipulative tasks and in the area of written tests such as the TIPS. If the trend of equal achievement of girls and boys on process skills holds true, we may find that teaching science in an integrated, process-oriented manner is, indeed, the more equitable method being searched for by many researchers in the field.

THE SECOND IEA SCIENCE STUDY AND SCIENCE EDUCATION IN THE UNITED STATES

Willard J. Jacobson

> I assume that some educational research is important. I wish there were some way that we teachers could benefit from it. (A Fifth Grade Teacher)

We believe that the major purpose of the Second IEA Science Study (SISS) is to improve the education in science of the children and young people in the countries taking part. We are trying to prepare our papers, monographs, and other materials in such ways that teachers can use them. All of our writers and researchers have had experiences in teaching in schools. We have the dictum, "Write so that the teacher in the room next door can benefit from the work you have done."

There are great advantages to being an international study; we can learn from each other. (Rosier, 1987) (Klein and Rutherford, 1985) In international studies we can study the effects of a wider range of factors than are possible in any national study. For example, our colleague, Pinchas Tamir of Israel, has studied the effect of prior learning on achievement in various sciences, and this is of interest and importance to those in other countries who plan science programs. Tamir has found that in Israel ". . .prior learning of biology did not affect achievement in topics other than biology." For chemistry, ". . . prior study of chemistry did not contribute to achievement of physics, but made considerable contributions to achievement in biology and in chemistry." For physics, ". . . prior learning of physics contributed to achievement in physics. chemistry, and earth science." (Tamir, 1985, pp. 17-18) These findings from Israel provide empirical evidence that can be used in planning science programs in other countries, such as the United States, where some reorganization of science programs may be taking place.

In SISS, 11 different populations have responded to science achievement tests, completed questionnaires and opinionnaires, and fifth and ninth graders have taken process lab tests involving the handling and manipulating of science equipment and materials. The basic results from this Study will be reported in a volume, Science Achievement in the United States. The item analyses and other data will be reported in this volume. Much deeper and searching analyses will be offered in the monographs that are listed in this paper. One of these monographs will be written for consideration by the public and science education policymakers. The following identification of four questions raised by the Study and brief discussions of what might be done to deal with these problems are indicative of the materials that will be prepared for the public.

Why are the science achievement scores in biology, chemistry, and physics in the United States substantially lower than in such countries as England and Japan? What are some things that can be done to raise science achievement in the United States? In the United States, students who had had one year of biology, chemistry, or physics, and students who were not studying science were tested. Compared to students tested in England and Japan, their

FIGURE 2

Preliminary Analysis
COMPARISON OF ACHIEVEMENT SCORES ON INTERNATIONAL BIOLOGY TEST
(13 ITEMS)

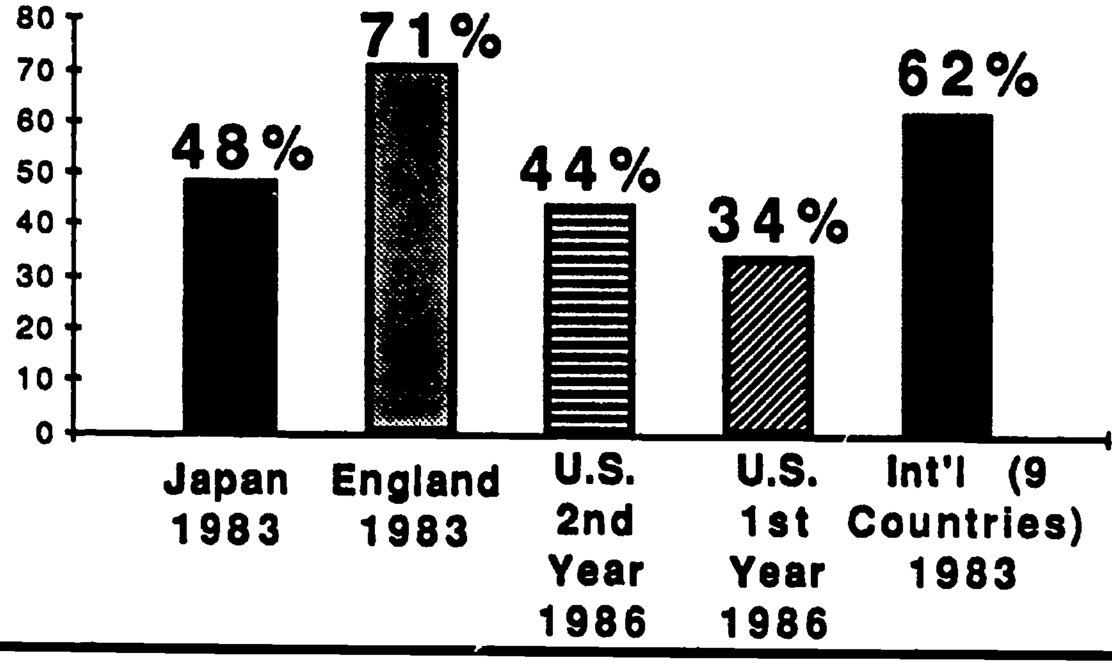


FIGURE 3

Preliminary Analysis
COMPARISON OF ACHIEVEMENT SCORES ON INTERNATIONAL CHEMISTRY TEST
(13 ITEMS)

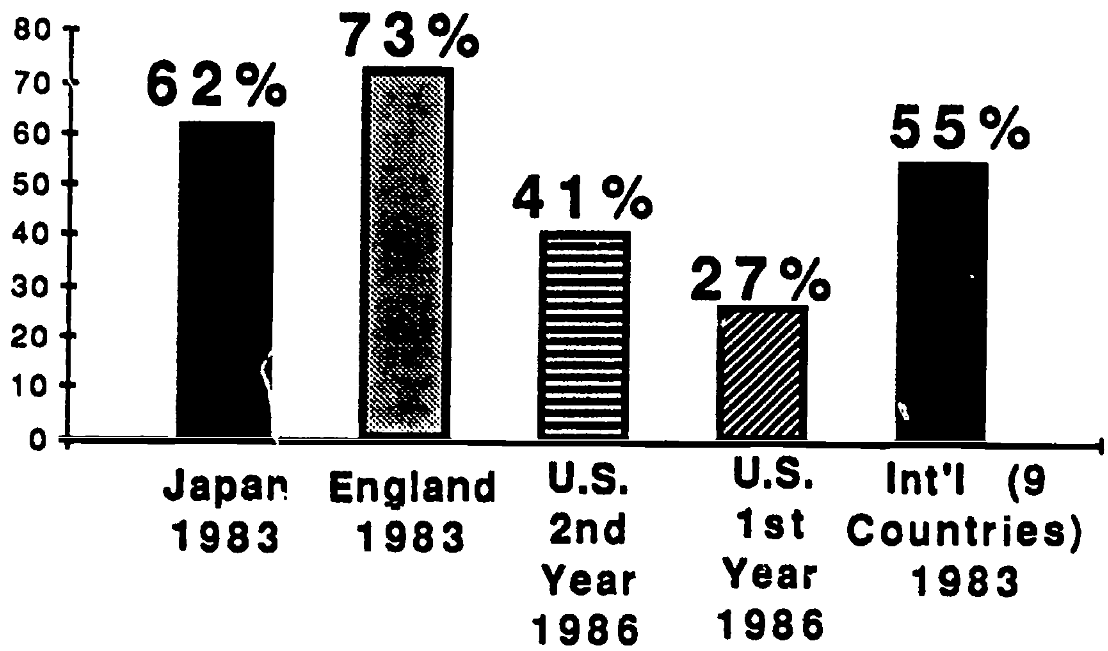


FIGURE 4

Preliminary Analysis

COMPARISON OF ACHIEVEMENT SCORES ON INTERNATIONAL PHYSICS TEST

(26 ITEMS)

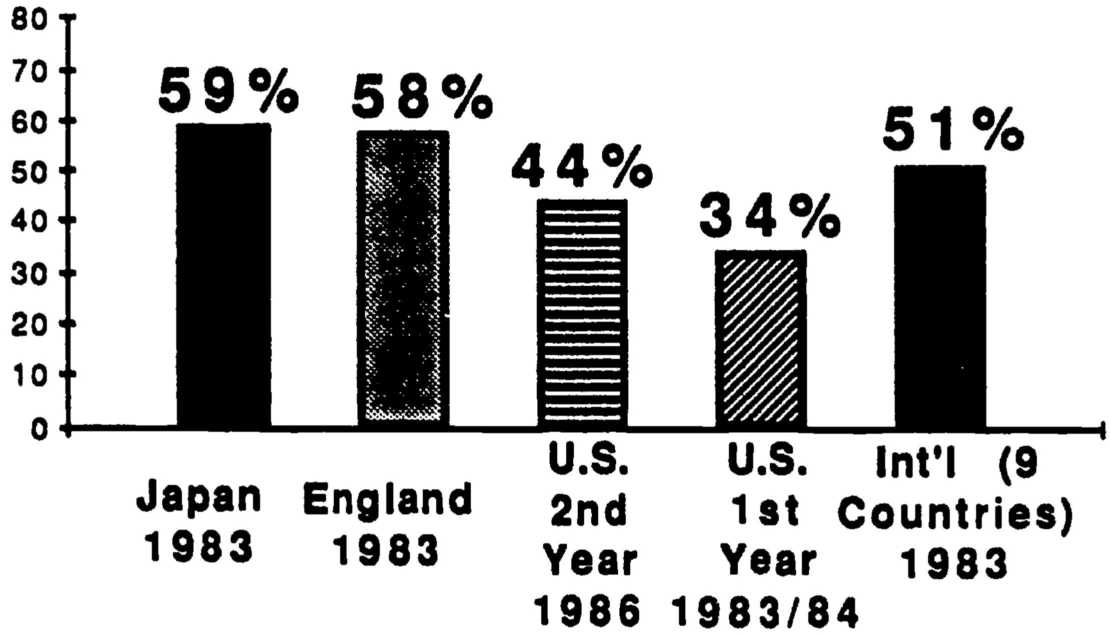
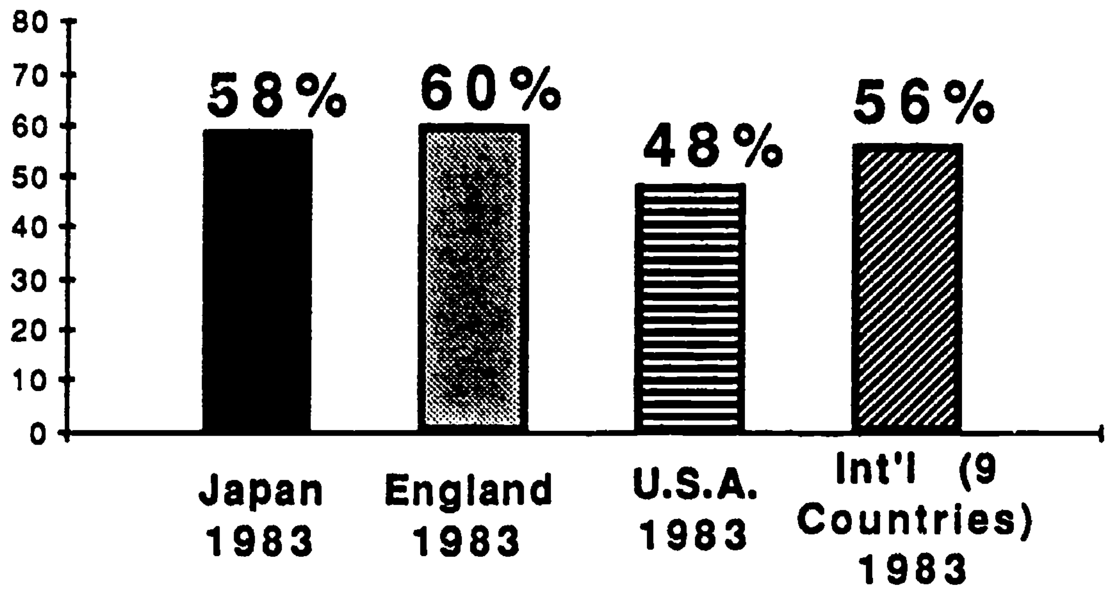


FIGURE 5

Preliminary Analysis

COMPARISON OF ACHIEVEMENT SCORES OF 12TH GRADE STUDENTS NOT STUDYING SCIENCE

(30 ITEMS)



scores were low. (See Figures 2 through 5.) But, in the United States biology is usually studied for one year in the 10th grade and chemistry for one year in the 11th grade. Since the Japanese and English students were tested in the last year of the secondary school, usually the equivalent of U.S. 12th grade, this is a partial explanation for the gap in achievement. In England, approximately half the population of 16- to 18-year-olds continue their education--one/fifth in school and four/fifths in colleges of further education (CFEs) (Keys, 1984). The students specializing in the sciences may study a science for three or more years.

Some American students, usually advanced placement students, study a science for a second year. Also these usually are the most academically gifted and science-oriented students. But, while scoring somewhat higher than first-year students, these second-year science students still did not do as well as students in England, Japan, and other countries. See Table 4 on Page 14 for the comparative scores.

What can be done to raise the achievement in science of U.S. students? This is a very complicated question that will be dealt with at greater length in other monographs. Briefly, we shall suggest three different directions for our search for ways to raise science achievement.

One approach is to organize our science programs very much as we have in the past but do more science and do it better. Our European friends call our science programs "the layer cake" approach. Usually, we offer science in one-year layers, and a decreasing number of students study a science each successive year. In the European metaphor, this suggestion would be an attempt to improve science achievement by thickening and improving the layers. In a sense this is what Japan has done. The Japanese organization of science curricula is very similar to the American. But, apparently, more students take two or more years of a science in Japan than in the U.S. Also, the Japanese school year is considerably longer. In Japan we see science programs somewhat like the U.S. science programs but with more time allocated to science.

Another approach could be to emulate the European science programs which often are more selective and there is greater specialization. In England, for example, at about the age of 16 students begin to specialize. The student who becomes involved in their science and mathematics stream concentrates almost entirely -- science and mathematics. Some students of English science education believe that English students who have completed the science-mathematics stream have reached a level of science achievement comparable to American college students.

It might be more consistent with the American ethos to develop a variety of approaches to the improvement of science achievement. The range of secondary school science could be broadened to include such sciences as astronomy, electronics, microbiology, ecology, integrated science, computer science, and science and technology courses, and the interests of some students would be aroused in some of these subjects. The amount of time devoted to science study can be increased through special programs on Saturdays and summers, such as science honors programs, ecological field experiences, natural history camps, and participation in science programs in museums, zoos, and other specialized educational institutions. Usually, these are voluntary activities in which students choose to take part. Teachers are encouraged to develop special expertise in areas ranging from chronobiology to ornithology; in return, they bring some students with them as they develop their own expertise. The steps that are taken in this approach should arouse interest, release energy, and stimulate creativity of students and teachers and will be discussed at greater length elsewhere.

SISS--U.S. findings indicate that all is not well with science in U.S. high schools. In a large country with a heterogenous population that takes pride in variety, different approaches to the improvement of secondary school science can be tried and tested. The results of survey studies, such as SISS--U.S., can be suggestive of ideas that should be tried.

Why do females have lower science achievement scores than males? What can be done about this "gender gap?" The differences in science achievement scores between males and females appear to be pervasive. The "gender effect" is evident in the science achievement scores of all countries who have released their data, with the possible exception of Population 3 in England and Israel. Disturbingly, the gap between science achievement of males and females seems to increase as the students pass through school. In the U.S., the gender gap is about 5% in 5th grade, 6% in Grade 9, and about 11% in 12th grade physics on general science content. The extent of the gender gap is exemplified by the results for the second-year chemistry test taken by advanced science students who have had a second year of chemistry. With these students, boys did better than girls on 24 items, girls did better than boys on 2 items, while on 4 items the scores were essentially the same (Menis, J. 1987, p. 64). On the first- and second-year specialty tests, there is a consistent gender gap of 5% to 7%. While the gender differences appear to be pervasive, there are differences between countries. The gender differences in Japan are less than those in the U.S. for every population except Population 3 students not studying science (Jacobson, W. & and Takemura, S., 1986, p. 55).

While the gender differences in science achievement appear to be pervasive, there are indications in the SISS--U.S. results that suggest how science achievement of girls can be improved. There were essentially no differences between the scores of males and females on the science process exercises. In these exercises, every student is given the same test booklet and science materials and equipment. None of the students have to compete with other students for the use of a few pieces of essential equipment or for materials of which there is not enough for everyone. Given equal access to equipment, materials, and teacher-time, as there is on the SISS science process lab exercises, females seem to achieve as well as boys. Perhaps, we should make certain that these conditions of free access for everyone exist in our science classrooms and laboratories!

It may be that in many cultures males have more of certain kinds of experiences than females. With Population 1, the item on which there was the greatest difference between males and females was one on which students were asked to predict the direction a beam of light would be reflected by a plane mirror. Of the females, 35% answered correctly and of the males, 61% answered correctly. Why is there this difference? It may be that males have more of such relevant experiences as reflecting light with mirrors, bouncing balls off walls, or playing billiards. Certainly, all children should be encouraged to have rich experiences in a wide variety of science-related activities. But, in school we should make certain that all students, female as well as male, have a wider variety of experiences in handling, manipulating, and experimenting in science activities.

While we may have limited influence or control over the experiences that males and females have out-of-school, in school we can make every effort to see to it that all students can take part in the science activities we plan. To achieve this, we must have the equipment and materials so that every student can do the observing, measuring, experimenting, viewing with a microscope, and all the myriad activities that are essential for learning science. Our research suggests that this is essential. It is easy for teachers and supervisors to agree, but a Japanese science educator who has visited many American schools said, "Almost all American science educators are for 'hands-on' science but go visit the schools." In the science classrooms you know, do all students have equal opportunities to handle, manipulate, and mess around with science equipment and materials?

Why were so few students able to successfully complete some of the process laboratory exercises? Although provided with a scale to find the mass of an object, shown the steps to measure the volume of an object, and given the formula to calculate density, only 8% of the students were able to find the density of an object. Although generally successful in measuring the temperature of equal amounts of water in two cups, only 7% were able to predict the resulting temperature when the water from each of the two cups was poured into a larger cup? Why are so few of our students able to carry out these operations? What can be done to make certain that more of our students will master these skills?

In order to learn some concepts and processes, it is probably necessary to study them more than once. The ways that various science concepts and processes are organized in science programs are reported in a monograph entitled An Analysis of Science Curricula in the United States (Miller, 1986, pp. 60-74). There are some concepts, such as "simple machines," that are studied at the 5th, 9th, and 12th grade levels with increased sophistication in the upper grades. There is some evidence that different and sometimes intriguing approaches are used in the development of concepts associated with simple machines. However, the "solar system" receives high coverage in Grades 5 and 9, but there is very little coverage of it in Grade 12 physics. Are there concepts related to the solar system that should be considered at a more advanced level in the 12th grade?

In the Second IEA Science Study, there were some test items, called "anchor" items, that were used with more than one population. At each level where the students had had a chance to study the science concepts related to a test item at an earlier level, there were improvements in achievement from Grades 5 to 9 and from Grades 9 to 12. For example, on a simple machine item involving a seesaw, the percentage of students answering the items correctly rose from 43% in Grade 5 to 75% in Grade 9. On 20 common items, Grade 9 students had scores that were, on the average, 19% higher than those of Grade 5 students. On another anchor item involving the tension in a string, the percentage of students answering correctly increased from 29% in the 9th grade to 59% in Grade 12 physics. On 16 common items, Grade 12 physics students achieved a 22% higher score than 9th grade students. As may be expected, students will raise their achievement score on an item if given an opportunity to study science concepts related to it. What happens if students do not study science? On 16 common items, 12th grade students who were not studying science had the same mean score as 9th grade students. Apparently, at the upper secondary school level, students who do not study science do not learn much school science.

An implication from the results of this Study is that we should think through very carefully what our goals of science education should be and then plan a variety of activities to help students achieve these goals (Miller, 1986, pp. 79-80). The spiral curriculum is the traditional curriculum development approach to planning science activities that deal with the same or similar concepts throughout the K-12 science program. But, if the same concepts are developed repeatedly in the same way, the repetition can be deadly. Instead, the same concepts should be approached in a variety of ways which help students to see and reflect on concepts from a variety of perspectives. An example of how this might be done is based on one of the items in the 9th grade lab exercise.

If we assume that one of the specific goals of a science program should be to learn how to determine the density of an object, there are several ways it can be approached. The SISS lab exercise leads the student into finding the mass of an object using a spring scale, the volume by measuring the amount of water displaced in a graduated cylinder, and offering the formula by which density can be calculated. The students are asked to give the units for mass, volume, and density. Only 8% of the 9th graders undertaking this exercise used the appropriate calculation procedures, found the correct value, and used the right units. Surely, there must be ways to help students to higher achievement than this.

It is suggested that a concept such as density can be approached in a variety of ways at different stages of the K-12 science program. The classic approach uses the overflow can and might be used somehow in the middle grades. It has been suggested that students should be asked to formulate procedures to find the density of a human body. At another time students might be asked to find the density of a liquid. A variety of such approaches can be used at different stages to gain a better understanding of density and how it can be determined. Decide what is important and have students approach what is important in a variety of ways at different times.

Are there any evidences of growth in science achievement that might be due to the impact of the science curricula development programs of the 1960s? In a meta analysis of 25 years of research comparing the performance of students in the "new" curricula with those in more traditional courses, "The new secondary science curricula, we found, were consistently more effective" (Kyle, Shymansky and Alport, 1982). It has been found that the achievement scores were raised 9 to 14 percentile points when students were placed in classrooms using new programs. The new programs have been found especially effective for females and urban students.

After an analysis of the achievement of students exposed to different curricula, it was found that students did better on those parts of the curricula that were distinctive (Walker and Schaffarzick, 1974, p. 108). The elementary science curriculum development projects emphasized science processes. It is of some interest to see whether there has been an improvement in the science process skills in the 13 or 15 years since the First International Science Study (FISS) in 1970.

Have the new elementary science programs been effective in raising science achievement scores? One of the features of some of the new elementary science programs was the emphasis on such science processes as observation, classification, measurement, stating hypotheses, controlling variables, operational definition and designing experiments. One of the new programs was Science--A Process Approach which was sponsored by the American Association for the Advancement of Science (AAAS). Most of the other new elementary science programs also featured the development of science process skills. New programs had great influence on science courses of study developed by communities and states and upon science textbooks. Have there been any gains in science process skills in the years that followed?

Several science educators read the items in the SISS tests and independently categorized them as "process" or "nonprocess" items. Eleven of the "bridge" items that were used in both 1970 (FISS) and in the 1980s (SISS) were classified as "process" items. From 1970 (FISS) to SISS I in 1983 there was a mean gain of 6.6%, which is substantial.

From 1970 (FISS) to 1986 (SISS II) there was a mean gain of 1.8%. It may be of interest and importance to note that there was a decline in mean score from 1970 to 1986 on nonprocess items. See Table 31.

TABLE 31

Gains on Process and NonProcess Scores
From 1970 to 1983 and 1986
Grade 5

	1970-1983	1970-1986
PROCESS ITEMS	+6.6%	+1.8%
NONPROCESS ITEMS	+4.2%	-1.2%

Thus, there were greater gains in the mean scores on process items than on nonprocess items. The gains are not as great as those reported for secondary school science (Kyle, Shymansky, and Alport, 1983). However, fewer resources were devoted to development of elementary school science programs. Also with 100 thousand schools, more than 30 million pupils, and 2.5 million teachers, the tasks of bringing about improvement in elementary school science are indeed immense. The results of this study suggest that it is possible to achieve improvement in areas that are selected for emphasis and energy and resources are devoted to the achievement of the improvement. Perhaps, the results also suggest that ways should be found to nourish the improvements.

SISS--U.S. Summary

Do studies such as the SISS--U.S. have implications for the practice of science education? It may be that some of the findings have profound implications. It may be that we now have some notions as to how science achievement in the U.S. can be raised. The task is immense, but we may have a clearer notion of the directions we can go. For a long time we have been perplexed by the "gender gap" in science achievement between girls and boys. Now some of the results of the science process testing seem to indicate steps that can be taken to reduce and then to eliminate the gap. Also, there is evidence of relationships between laboratory work and certain kinds of teaching styles and achievement in science. Our research supports the findings of others that the unprecedented science curriculum development efforts of the 1960s did make a difference and that progress can be made if we decide on what kind of development we want and provide the ingenuity, human energy, and the resources that are needed to bring about improvement.

BIBLIOGRAPHY

Backman, M.E. (1971). Patterns of mental abilities: Ethnic, socioeconomic, and sex differences. American Educational Research Journal, 9, 1-12.

Baker, S. W. (1980, Autumn). Biological influences on human sex and gender. Journal of Women in Culture and Society, 6, 80-96.

Bornstein, R. (1982). Sexism in education. In M.P. Sadker & D.M. Sadker (Eds.), Sex equity handbook for schools. New York: Longman, Inc.

Comber, L.C., and Keeves, J. P. (1973). Science education in nineteen countries. International studies in evaluation I. New York: John Wiley and Sons.

Dickson, D. (1986). Britain's Royal Society condemns sex bias in mathematics teaching. Science, 233, 618-619.

Haertel, G. D.; Walberg, H. J.; Junker, L.; & Pascarella, E. T. (1981, Fall). Early adolescent sex differences in science learning: Evidence from the National Assessment of Educational Progress. American Educational Research Journal, 18, 329-341.

Harding, J. (1973). What has Nuffield done for girls? Times Education Supplement (30 November).

Jacobson,W.J.; Takemura, S.; Doran, R.L.; Humrich, E.; Kojima, S.; & Miyake, M. (1986). Analyses and comparisons of science curricula in Japan and the United States. New York: Second IEA Science Study---U.S., Teachers College, Columbia University.

Johnson, S., & Murphy, P. (1986). Girls and physics: Reflections on APU survey findings. Occasional paper #4. University of Leeds, Kings, College, London: Assessment of Performance Unit, Department of Education and Science.

Kahle, J.B. (1983). The double dilemma: Minorities and women in science education. West Lafayette, Ind.: Purdue University Press.

Keeves. J.P. (1973). Differences between the sexes in mathematics and science courses. International Review of Education, 19, 47-64.

Keys, W. (1984). Science curriculum case study: England. National Foundation for Educational Research, Slough, Berks. England.

Keys, W. (1986) The Second IEA Science Study: Aspects of science education in English schools. National Foundation for Educational Research, Slough, Berks. England.

Keys, W. (1987) Aspects of science education in English schools. National Foundation for Educational Research - Nelson, Windsor, Berkshire, England.

Kelly, A. (1978). Girls and science. IEA Monograph Studies, #9. Stockholm: Almqvist and Wiksell.

Klein, M. & Rutherford, J. (eds) (1985) Science education in global perspective: Lessons from five countries. American Association for the Advancement of Science. Washington, DC.

Kojima, S. (1974, June) "IEA science study in Japan with special reference to the practical test." Comparative Education Review, 18, 262-267.

Kyle, W.C., Shymansky, J.A., & Alport, J.M. (1982, November). Alphabet soup: A second look at the NSF-funded science curricula. The Science Teacher, 49(8), pp. 49-53.

Lipkin, J., & Sadker, D. (1982). The report card #3: Sex bias in mathematics, computer science and technology. Washington, D.C.: The Mid-Atlantic Center for Sex Equity.

Mattheis, F. E.; Spooner, W. E.; & Coble, C. R. (1985). A study of the logical thinking skills, integrated process skills, and attitudes of junior high school students in North Carolina." Greenville, N.C.: East Carolina University.

Menis, J. (1987): Achievement in chemistry in the United States. New York: Second IEA Science Study--U.S., Teachers College, Columbia University.

Miller, J. (1986): An analysis of science curricula in the United States. New York: Second IEA Science Study--U.S., Teachers College, Columbia University.

Miyake, M. (1985) National report: Japan. Tokyo: National Institute for Educational Research.

Noonan, R. & Engstrom, J. (1986): Science achievement in Swedish schools: Results from the IEA Second International Science Study. Stockholm: Stockholm University.

Rosier, M. (1987): "The Second International Science Study." Comparative Education Review, 31 (1), 106-128.

Sadker, M.P., & Sadker, D.M. (1982). Sex equity handbook for schools. New York: Longman, Inc.

"Science and Mathematics Education in American High Schools: Results for High School and Beyond." (NCES) 84-211b Washington, DC: National Center for Educational Statistics, May 1984.

Steinkamp, M. W., & Maehr, M. L. (1984, Spring). Gender differences in motivational orientations toward achievement in school science: A quantitative synthesis. American Educational Research Journal, 21, 39-59.

Tamir, P. (1985). Some factors which affect science achievement of high school seniors in Israel. Jerusalem: Hebrew University. 17-18.

Vestin, M. (1975). Some information concerning sex role question and programs for development of equality between men and women especially within schools and education. Stockholm: Swedish National Board of Education.

Walker, D.F. & Schaffarzick, J. (winter, 1974). Comparing curricula. Review of Educational Research (RER), 44(1), 83-111.

Wolf, R. M. (1977). Achievement in America. New York: Teachers College Press.

Zerega, M. E.; Haertel, G. D.; Tsai, S.; and Walberg, H. J. (1986). Late adolescent sex differences in science learning. Science Education, 70, 447-460.