R E P O R T   R E S U M E S

ED 010 998 SE 000 638

CONCEPT DEVELOPMENT MATERIALS FOR GIFTED ELEMENTARY PUPILS--FINAL REPORT OF FIELD TESTING.

BY- SANDS, THEODORE AND OTHERS
ILLINOIS STATE UNIV., NORMAL

PUB DATE 66
EDRS PRICE MF-$0.09 HC-$2.00 50P.

DESCRIPTORS- *CONCEPT FORMATION, *ELEMENTARY SCHOOL SCIENCE, *ABLE STUDENTS, CURRICULUM DEVELOPMENT, GRADE PLACEMENT, INDEPENDENT STUDY, READING ABILITY, RESEARCH REPORTS, ADVANTAGED PROGRAMS, NORMAL

AN ANALYSIS OF A FIELD TEST OF SCIENCE LEARNING MATERIALS FOR ABLE ELEMENTARY STUDENTS IS REPORTED. THE LEARNING MATERIALS FIELD TESTED AND EVALUATED WERE UNGRADED INDEPENDENT STUDY KITS. THEY WERE DESIGNED FOR MAXIMUM SUITABILITY REGARDLESS OF TEACHER OR SCHOOL CIRCUMSTANCES. THE LEARNING MATERIALS INVOLVED CONCEPT FORMATION IN THE TOPICS OF ATOMS, MOLECULES, AND MEASUREMENT. TWO HUNDRED FIFTY-NINE STUDENTS FROM 31 DIFFERENT ILLINOIS SCHOOLS USED THE MATERIALS. STUDENTS IN THE STUDY WERE SELECTED BY THEIR TEACHERS AND ADMINISTRATORS. METHODS OF USING THE MATERIALS WERE ESTABLISHED BY THE TEACHERS. PRE-TESTS AND POST-TESTS WERE ADMINISTERED TO THE STUDENTS. RESULTS ARE REPORTED FOR EACH TEST ITEM AND EACH GRADE LEVEL. THE AUTHORS CONCLUDED THE LEARNING MATERIALS TESTED CAN BE USED EFFECTIVELY. GREATER SUCCESS IN USE IS PREDICTED FOR GRADES 2 AND 3 THAN FOR GRADES 1 AND 4. (RS)

638

CONCEPT DEVELOPMENT MATERIALS FOR GIFTED ELEMENTARY PUPILS:

Final Report of Field Testing

Theodore Sands
Robert E. Rumery
Richard C. Youngs

Illinois State University
Normal, Illinois

1966

This project was supported by a grant from
The Department of Program Development for Gifted Children
Office of Superintendent of Public Instruction, State of Illinois

JAN 1 7 1967

Preface

This report deals with the final stage of a project in which concept development materials for gifted elementary pupils were developed and tested. The materials were originally developed by the Elementary Curriculum Materials Project personnel in 1963-65. Personnel who helped develop these materials are listed on the next page. During the initial phase the emphasis was on translation of a rationale for instruction into concrete procedures and materials which could be subjected to empirical test. In the final phase the emphasis was on refinement of the original products.

We wish to express our appreciation to the many people in the public schools and the University who helped us collect and evaluate the data which are presented in this report. Special mention should be made of Mr. Richard Youngs, associate director, who ably carried on the day to day administration of our work, and Mr. Robert Rumery, evaluation specialist, who was responsible for evaluation strategy and analysis of data. Mr. Fred H. Bradshaw and Mr. John H. Conlin maintained liaison between participating teachers and project personnel. Their sensitivity to the needs of classroom teachers did much to facilitate our work. Mrs. Leda Fahrenkrog, Mrs. Marge Anderson, Sharon Matthews, Connie Mathews, Ruth Johnson, and Kara Knight gave able assistance in the collection of data and the production of teaching materials.

Theodore Sands

ELEMENTARY CURRICULUM MATERIALS PROJECT PERSONNEL

1962-66

Principal Investigator	Theodore Sands, Ph. D. Coordinator of Programs for the Gifted, I.S.U.
Co-Investigator (1962-65)	Charles R. Hicklin, Ed. D. Prof. of Education, I.S.U.
Associate Director (1965-66)	Richard C. Youngs, M. A. Asst. Professor and Supervising Teacher of Science, Metcalf Laboratory School, I.S.U.
Editor and Consultant	Barbara S. Price
Evaluation Specialist	Robert E. Rumery, M. A. Asst. Prof. of Psychology, I.S.U.
Science Specialists	
(1964-66)	Richard C. Youngs, M. A. Asst. Professor and Supervising Teacher of Science, Metcalf Laboratory School, I.S.U.
(1962-64)	Thomas F. Edwards, M. A. Asst. Prof. of Physical Sciences, I.S.U.
Mathematics Specialists	
(1965).	Wilson P. Banks, M. S. Asst. Prof. of Mathematics, I.S.U.
(1964).	John H. Esbin, Jr., M. A. Instructor in Mathematics, I.S.U.
(1962-64).	Kenneth A. Retzer, M. Ed. Asst. Prof. of Mathematics, I.S.U.
Research Assistants	Christine Hooper, M. A. William J. Kritzmire, M. A. Fred H. Bradshaw, M. A. John H. Conlin, M. S.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

CHAPTER I

STATEMENT OF THE PROBLEM

Teaching the gifted can be a rewarding and, occasionally, an exhilarating experience. There are also moments of soul-racking frustration and gnawing feelings of inadequacy. To be sure, these are the joys and sorrows of all teaching; but for those who work with the gifted, all opportunities seem magnified and all failures seem more poignant. In part this situation can be explained by the nature of the gifted: by definition their potential is greater, and as a result most teachers expect more of both their pupils and themselves. This is a part of the problem that we cannot solve. Perhaps, it is just as well, for without high expectations the quest for excellence becomes an exercise in mediocrity.

There are, however, ways to make the efforts of both teachers and pupils more productive. Teachers can be supplied with instructional materials that are appropriate for the instruction of the gifted. In an earlier report, The Development and Testing of Instructional Materials for Gifted Primary Pupils, 1965, Theodore Sands, Charles Hicklin, Richard Youngs, Robert Rumery, Barbara Price, and Kenneth Retzer, we said that

> Regardless of the type of program, the instructor must face the ineluctable question of what to teach, and one of the prime factors in determining this is the availability of appropriate instructional materials. Whether one operates within the context of enrichment or acceleration, homogeneous or heterogeneous grouping, in the end all programs must solve the problem of obtaining suitable instructional materials.

Educators have known for some time that young students, especially the gifted, can learn principles and skills which employ higher thought processes and which have traditionally been taught at an older age. However, before more meaningful teaching can take place, it is necessary to identify the concept in language appropriate to the grade level and to devise experiences which might lead to the development of the concept. The instructional process can be facilitated by providing the teacher with materials designed to accomplish these goals.

Not only is there need for materials which induce the use of higher thought processes in the learner, but equally important these materials should be adaptable to individualized instruction.

The "gifted" are sometimes identified as a homogeneous group, but in reality the range of differences in terms of abilities, skills, interest, and background is usually very large. No two gifted students are "gifted" in the same way. Ideally, the teacher should have available instructional materials which would enable her to vary the curriculum to meet the needs of the individual pupil and give each member of the class an opportunity to develop his talents and skills at a pace and level of learning commensurate with his ability.

Providing appropriate individual learning experiences for gifted students presents problems of a special sort. First, there is the problem of finding materials which are appropriate for the gifted. We would argue that at the primary and middle grades, at least, materials used by the gifted should develop basic concepts and at the same time require the use of higher thought process. Materials designed for use with the gifted should induce in the learner such skills as: analysis, prediction, verification, extrapolation, and at later stages, synthesis.

A second major obstacle facing the teacher is to find materials which will enable her to meet the special needs of gifted students, and at the same time leave her sufficient time to meet her obligations to her other students. In practice this means that the teacher must be able to assign tasks which do not require her constant supervision and free her as the primary source for questions and explanations.

For teachers of the primary and middle grades this presents an almost insuperable problem if they are to rely on existing materials. There is the problem of reading. The capacity of the child is not limited by his ability to read, but his learning ability often is. Much of the material that explains basic concepts uses the written word as the mode of communication. Our previous investigation (Sands, Hicklin, et al., 1965) indicates that many of the concepts in science and mathematics which are currently reserved for the later grades, can be learned by gifted pupils when a "non-reading" means of communication is used. For those gifted pupils who can read, much of current materials is of limited utility. At one extreme, the conceptual demands of materials are beyond the prior experiences and learnings of primary pupils. At the other extreme, language and concepts are understandable, but the content is trivial.

In the previous phase of this project an attempt was made to devise a strategy for creating materials that would satisfy these deficiencies. Specific objectives of the first phase of the project were:

1. To create and test a sequence of instructional experiences which would enable a gifted student while working independently to develop concepts which were considered basic to a discipline, but not usually encountered in the early elementary grades.

2. To develop these concepts in a way which would require the bringing into play of higher thought processes.
3. To identify a strategy of instruction which would enable such materials to be used in the public schools with a minimum of teacher attention and participation, require no special training of the teacher, and be adaptable to whatever patterns of administrative arrangements for instruction of the gifted are current.

The results of the first phase are reported in the Final Report (Sands, Hicklin, et al., 1965). The results were sufficiently encouraging to further field test the materials developed by the project staff.

The phase of the project which this report deals with had as its objectives the revision of the materials and the evaluation of their effectiveness in a variety of school settings.

Specific objectives were as follows:

1. To make the revised instructional material available to all the elementary schools in Illinois.
2. To enable schools to determine whether they wish to make such materials a permanent part of their programs for the gifted.
3. To further test the materials to determine:
   a. The effectiveness of improved format of tests, revision of lessons, and use of phonographs instead of tape recorders.
   b. The effectiveness of the materials with different gifted populations.
   c. Teacher acceptance on a state-wide basis.

CHAPTER II

DESCRIPTION OF MATERIALS

In the summer and fall of 1964, forty-four self-instructional lessons dealing with some fundamental concepts about the structure of atoms, the nature of molecules, and measurement were written. These lessons, organized in three sets or kits of materials, were used by one hundred ten pupils in twenty-one first grade classrooms located in sixteen different elementary school attendance centers in Bloomington, Illinois and McLean County Community Unit District # 5 schools. In the summer of 1965 all lessons were reviewed and revised.

The materials had been written under the following guidelines:

1. Sensory-motor activities. Wherever possible, abstractions and symbols were to be associated with appropriate sensory-motor activities. Each lesson should actively engage the student in manipulation of illustrative materials and overt behavior related to making inferences, solving problems, or predicting.
2. Operational definitions. The student should be given the opportunity to perform an operation and then be told the word or words which denote the behavior.
3. Programming principles. The instructional material should embody the following techniques derived from programming principles: identification and statement of objectives in terms of behavioral outcomes, presentation of information in small steps, careful sequencing, immediate confirmation of adequate responses, self-pacing.

The materials provide a depth of treatment and call for a level of abstraction that goes considerably beyond what is asked of primary pupils in science instruction. The following description from the 1965 Final Report might serve as a helpful illustration:

> In the lessons that deal with atomic theory and molecular structure, the pupil is required to develop a mental construct of the atom and the molecule by means of symbols. The relationships between the symbol and the thing it represents are established by analogy. Once the characteristics of the parts are identified, the child is required to establish the relationship between the parts which form the atom and the molecule. In addition the child is asked to apply theory to explain or predict change.
>
> In the measurement unit measurement is treated as a concept as well as an operation. The child is given the opportunity to identify the elements which are common to all measurement and to apply the concept of measurement to a variety of measuring operations.

These concepts are presented in a way which makes learning largely self-instructional, but the material can be easily adapted to a variety of other teaching strategies. Self-instruction is carried on by means of plastic phonograph records. In addition each child is supplied with a complete set of specially designed materials and a test booklet. The materials provide the child with experiences of an illustrative or problem-solving nature.

It would be useful to ask, what concepts are developed in the units? What higher thought processes are brought into play? These questions can be answered by the following analysis from the 1965 Final Report:

ATOMS

Concept	Activities Involving Higher Thought Processes
The smallest part a whole can normally be divided into is an atom.	The child takes a whole apart, selects a part, which is then treated as a whole and taken apart. After this process is again repeated, the child is asked if the smallest part he sees might be broken down into an even smaller part.
	Child is asked to identify the way in which two lessons which teach a similar concept are alike.
Observations can be made in many ways.	Child is given three samples of unknown liquids and asked to identify the liquids by a variety of observational techniques.
The center of an atom is a place called the nucleus.	The child is given an example of an atom which is constructed incorrectly, asked to identify the error and to correct it.
An electron moves around the nucleus of an atom.	Child is shown a picture of the Bohr atom model and asked how the model might be improved so it would more closely resemble a real atom.
Objects with the same charge repel.	Child is shown pictures of charged atomic particles and asked how the particles will behave towards one another.
Attraction and Repulsion.	Child is asked to compare the ways in which magnets and atomic particles are similar.
Electrons are held in their orbits by the attraction of unlike charges.	Child is shown a model of an atom, which on spinning throws its electrons to the outer limit of the atom model. Electrons of the model are maintained at this limit by wires. The child is asked what the wires stand for in the real atoms (the attractive force between the positive nucleus and the negative electron.)

All atoms are made from the same kind of parts.	Child is given a variety of objects which represent atom parts (electrons, protons, neutrons). He is to arrange these parts so as to construct models of a number of different atoms.
The atoms of an element are all alike.	Child is provided with pictures which represent atoms. He is to select the pictures which represent elements(all the same kind of atoms).

MOLECULES

Concept	Activities Involving Higher Thought Processes
Symbols are used to represent things.	Child is given several symbols and told to match them with other symbols which stand for the same objects.
Atoms join to form molecules.	Child is given two symbols which stand for atoms and asked to join these two symbols together. Child is then asked to name the new entity.
Atoms form molecules by sharing electrons.	Child is provided with a manipulative model of a molecule. The atoms of the molecule model are joined with other atoms of the model with mechanical snap fasteners. The child is then asked what the snap fasteners represent. (electrons)
A collection of molecules composed of two or more kinds of atoms which are uniform in arrangement is called a compound.	The child is given several groups of atoms, some of which are uniform in arrangement and composition and others which are varied in their arrangement and composition. The child is then asked which groups represent compounds.
The properties of a substance identify the substance.	The child is given a description of the physical characteristics of an object and asked to identify the object.
All samples of identical compounds have the same kind and arrangement of atoms.	Child is given pictures which represent the atomic structure of a variety of substances. He is to identify which substances have identical properties.

Atoms can be arranged in a variety of ways to form different molecules.	Child manipulates symbols for atoms to form molecules. (The symbolic manipulation is followed by a chemical experiment which confirms their symbolic manipulations.)
The combination or separation of atoms and molecules may result in the release of energy.	Child is provided with mechanical models which release energy when they are separated. Child is asked to indicate the similarity between the mechanical models and chemical reactions.
An increase of energy increases the motion of atoms and molecules.	Child is asked to explain evaporation using this concept.

MEASUREMENT

Concept	Activities Involving Higher Thought Processes
The amount of space an object occupies is called its volume.	Child is shown pictures of solids, liquids, and gases and asked to determine if the concept of volume applies to them.
	Child is asked to determine why water from one container will not fill another container of exactly the same size which contains some marbles.
	Child is asked to determine volume of air in a sealed bottle which has water, air, and marbles in it and compare the volume of air with that of the water and marbles.
The amount of space an object takes up is independent of its orientation.	Child is asked to determine if the volume of a set of blocks in a given position changes as the position of the blocks is changed.
An object displaces a volume of liquid equal to its own volume.	Child makes a measuring cup using a given cube as a standard. Uses cup to discover that differently shaped solids can have the same volume. Uses cup to measure volume of an irregularly shaped object.

A unit of measure can be any convenient and appropriate object.	Child is asked to use three different objects as a unit of measure, and asked to decide if other objects might be used.
Area and shape are separate.	Child is given four square regions to manipulate and asked to determine if the different resultant shapes have the same area.
The most important characteristic of a standard unit is its constancy.	The child is asked to decide what is most important about an inch.
A standard unit is a convenience and a convention.	Child, having measured in various sized units, is asked to evaluate their convenience.
The greater the mass of the nucleus of an atom, the greater is the gravitational attraction.	Child is shown pictures of groups of atoms and is asked to identify which would weigh most.
The standard units of measurement need to be kept at a constant temperature.	Child is asked to identify from several alternatives what would happen if the standard units were kept at varying degrees of temperature.

CHAPTER III

USE OF MATERIALS

One of the objectives of the project was to get data on use of the materials from a state-wide sampling of school districts. This objective was achieved.

In the fall of 1965 a brochure announcing the availability of the materials was sent school districts throughout the state. The brochure explained the nature of the materials and invited participation in the project. To participate, school personnel had to agree to administer and return the pre- and post-tests. They also had to agree to return the kits at the end of the year. A rental fee of twenty-five dollars was levied for each set of three kits.

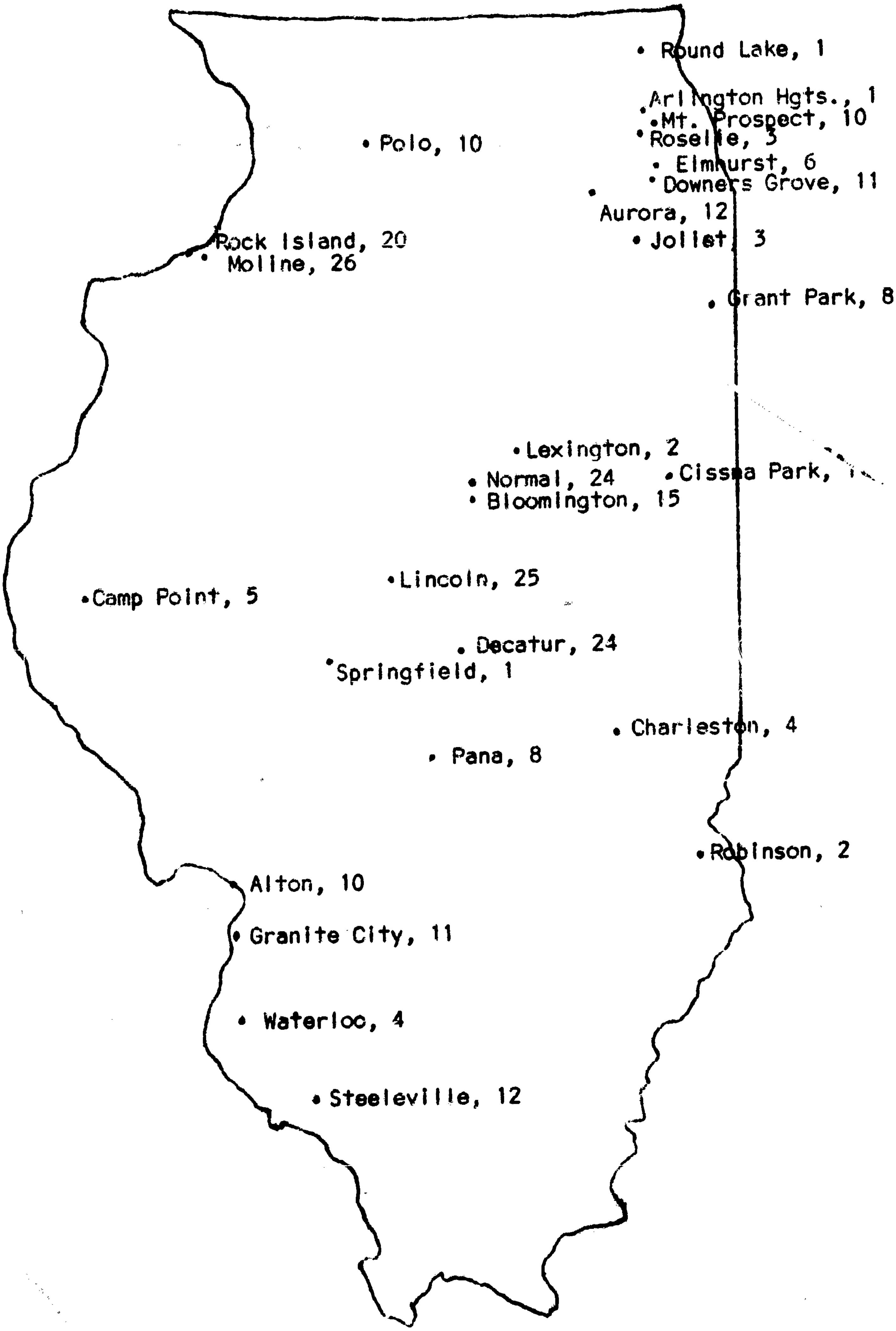
Agreements for use of the kits were reached with thirty-one school districts and two laboratory schools. Within these districts, 288 sets of kits were used in sixty attendance centers. Participating schools were chosen to achieve state-wide distribution.

The thirty-one participating districts encompassed urban and rural communities, suburban and inner-city schools, culturally deprived and culturally advantaged populations. The geographic distribution of participating schools is shown in figure 3.1.

Who were the users of the materials and how were they chosen? Since the question we posed was what results we could expect if these materials were put on the market for general use, minimal restrictions were imposed on manner of use of the materials. Teachers and administrators were free to select any pupil who in their judgment was "gifted" or who they had reason to believe might benefit from working with the

FIGURE 3.1

LOCATION AND NUMBER OF PARTICIPATING PUPILS



kits. The rationale behind this method of selection of pupils was that in the previous phase of the project data had been obtained on how pupils within a specified range of I.Q. scores performed with the materials. The objective this time was to see what would be done with the materials by pupils chosen as they would most likely be chosen under normal, non-controlled conditions.

Teachers and administrators selected 259 pupils on whom pre-test data was returned. These students had an I.Q. range of 86 to 171. We have reason to believe that a considerable number of additional pupils used the materials.

Where and how were the materials used? Again no restrictions were placed on teachers and administrators. They were free to use the materials at any grade level they thought appropriate and in any way they found useful. The result was that the materials were used with pupils in the kindergarten through the fifth grade. The number of pupils in each grade and their location is given in Table 3.1.

The heaviest concentration was in grade one with thirty-one schools using it at this level. Use in grades two, three and four was fairly evenly distributed; eighteen schools used the materials in grade two, fifteen in grade three, and ten in grade four. Three schools used the materials in grade five and one in kindergarten.

A special situation existed in the Washington school, Decatur. This school has a well-established program for culturally deprived pupils. The staff in Washington school feels that I.Q. tests do not accurately measure the intelligence of their pupils who come from culturally-deprived environments. Using teacher judgment and performance in class, twenty-four pupils were chosen to use the materials.

TABLE 3.1

USE OF MATERIALS

Location	Attendance Centers	Grade	Pupils per Grade
Alton	Gilson Brown (Godfrey)	2	10
Arlington Heights	Windsor	2	1
Aurora	Frank Hill	1	1
	Freeman	1	3
	Greenman	1	1
	McCleery	1	2
	Nancy Hill	1	1
	Smith	1	2
	Mary A. Todd	1	2
Bloomington	Centennial	1	5
	Lincoln	1	1
	Oakland	1	3
	Irving	1	2
	Washington	1	4
Camp Point	Clayton	4	1
	Coatsburg	4	1
	Golden	4	1
	LaPrairie	4	1
	Maplewood	5	1
Charleston	Buzzard Laboratory School	1	1
	"	2	1
	"	3	2
Cissna Park	Unit 6	1	1
Decatur	Washington	1	3
	"	2	3
	"	3	4
	"	4	10
	"	5	4
Downers Grove	Fairmont	2	3
	Kingsley	1	2
	Lester	3	3
	Washington	1	3
Elmhurst	Conrad Fisher	3	1
	"	4	2
	Cornille	2	1
	"	3	1
	"	5	1
Granite City	Frohardt	1	2
	Logan	2	2
	Maryville	2	2
	Mitchell	1	2
	Wilson	3	2
	(Special)	K	1

TABLE 3.1 continued

Location	Attendance Centers	Grade	Pupils per Grade
Grant Park	Grant Park Elementary	2	4
	"	3	4
Joliet	Forest Park	4	3
Lexington	Lexington Elementary	1	2
Lincoln	Jefferson	1	3
	"	2	3
	"	3	2
	"	4	5
	Northwest	1	2
	"	2	2
	"	3	8
Moline	Williard	2	9
	"	3	17
Mt. Prospect	Lincoln	1	3
	"	2	3
	"	3	3
	Lion's Park	2	1
Normal	Hudson	1	4
	"	2	4
	Fairview	1	8
	Oakdale	3	5
	Metcalf Laboratory School	1	3
Pana	Lowell	4	8
Polo	Centennial	1	10
Robinson	Lincoln	1	2
Rock Island	Eugene Fields	2	10
	"	3	10
Roselle	Campanelli	3	1
	Hillcrest	1	1
	Lakeview	1	1
Round Lake	Round Lake Park	2	1
Springfield	Lindsay	2	1
Steeleville	Steeleville Elementary	4	12
Waterloo	Unit 5	3	4

How were the materials used? Data on modes of instruction was obtained from forty of the sixty participating schools. Twenty-eight teachers chose to use the materials in their original self-instructional format. In this mode of instruction the child worked independently with the phonograph, records, and kit materials, and received only incidental and occasional assistance from the teacher. Most teachers allowed pupils to work during regular class time that otherwise would have been free. One teacher scheduled use of the materials during the lunch hour, and one pupil used the kits before the school day started.

The second most commonly used mode was teacher-led small group instruction. Nine schools reported using the materials as a basis for small group instruction. In this mode the records were played to the group, the teacher controlling the record player and leading discussion or asking questions about the concepts being developed in the lesson.

Three schools used the materials for large group or class instruction. In one school two sets of kits were used to instruct twenty-seven pupils by using an opaque projector and demonstrating the kit materials rather than allowing each child to perform the experiments and illustrative activities. In the other two schools materials were supplied to each child, but the entire class worked on them at the same time.

What was the reaction of teachers, parents and pupils to the materials? On the basis of data collected through visitation and informal conversations with teachers and administrators, it seems appropriate to say that the materials were well received. A member of the project staff visited each school in which the materials were used. Most teachers saw the kits as useful teaching aids. At only five of the sixty attendance centers using the materials was teacher response negative. At one school the materials were examined but not used. A number of suggestions were made for improving the kits; many were concerned with errors in production rather than content deficiencies. A frequent suggestion was that additional enrichment materials be included.

The overall favorable response of teachers undoubtedly represents some element of the Hawthorne effect. If there are elements in these materials which would elicit a negative reaction among teachers, we have not been able to identify them.

The response among students was favorable. At one end of the spectrum we received unsolicited fan mail from students telling us how much they enjoyed the materials. The following are representative samples--though something of the flavor is lost when the smudges and labored lines are eliminated by the typewriter:

> I liked the storys you told. I learned many new things about science. I hope I see you in secornd grade. Love.
>
> Mike D....

> I thought your lessons were very nice. Now I know all about science. I liked the storys that you told. We told some kindergartens al' about the science. Love,
>
> Rita H....

> I love the record. And the stoff we did. Love.
>
> Richard B.....

On the other hand, 41 of 237 pupils who completed the atoms unit, failed to complete the molecules unit. Most of this attrition was among first grade pupils. However, teachers reported that their pupils seemed to be interested and highly motivated by the materials.

A number of teachers remarked on positive responses from parents of children using the kits.

CHAPTER IV

EVALUATION

Introduction

Evaluation goals for this project were less general but more detailed than evaluation goals in the previous phase of this project (Sands, et al., 1965). In the earlier phase, evaluation centered on three questions.

> First, did the instructional material ... enhance the learning of scientific concepts considered to be fundamental to further study of science? Second, were the cognitive objectives achieved with minimal inter-classroom variations? Third, were the cognitive objectives achieved with minimal undesirable incidental effects?

In the earlier project, substantial weight was given to validity of inferences about the effect of experimental treatments in producing gain. Because of the importance of inferring causal connections between gains and treatments, the earlier study was evaluated within the framework of a quasi-experimental design--a modified nonequivalent control group design (Campbell and Stanley, 1963).

In the present study, no causal inferences were to be made; rather a causal connection was assumed on the basis of the results of the earlier study. As a result, in the present study, no safeguards against external influences on validity of inferences were applied. Evaluation of the present study involved six questions. First, were students better able to make correct responses after instruction than they were before instruction? Second, is the instructional material difficult for students to master? Third, is there any difference in the level of difficulty of the three units used? Fourth, do students have prior knowledge of some of the concepts included in the instruction? Fifth, are there grade differences in the ability to master the material? Sixth, do the materials have sufficient holding power that students could be expected to complete all three units in a school year?

Most of the information relevant to these questions was obtained by a detailed analysis of the results of a set of revised achievement tests. Revisions were made on achievement tests used in the previous project on the basis of data obtained from test performance and from teachers' comments. The nature and extent of these revisions is discussed in the next section.

Revision of Achievement Tests

Analysis of data and examination of teachers' comments from the previous study indicated the need for several modifications in the achievement tests used for evaluation. Among the difficulties in test administration reported by teachers were: (1) Difficulty in manipulating test booklets and instructions reproduced in separate booklets;(2) Difficulty in maintaining student attention throughout the testing period; (3) inability of students to remember accurately all the alternatives read to them. In addition, summary statistical data indicated that some items did not change in difficulty level as a result of instruction, some did not adequately discriminate high scorers from low scorers, and some were outside a desirable range of difficulty.

Three kinds of revision resulted. First, instructions and response pages were integrated into a single booklet; second, the number of alternative responses to each item was reduced to three with a fourth "don't know" response category; third, items which had inadequate indices of discrimination, were outside a desirable difficulty range, or were judged to be ambiguous were either revised or eliminated. Elimination of items was selective so that the sampling proportions of domain categories described in the previous report would be disturbed as little as possible.

The structures of the tests on atoms and molecules are given in Tables 4.1 and 4.2 in terms of categories of cognitive or logical operations and content.

TABLE 4.1

COMPOSITION OF TEST ON ATOMS

	Composition	Structure	Models & Symbols
Knowledge, specifics	2	2	2
Knowledge, generalization		2	2
Comprehension	2		
Application	2	4	4

TABLE 4.2

COMPOSITION OF TEST ON MOLECULES

	Composition	Structure	Models & Symbols
Designation	2	2	2
Description	2	2	2
Explanation	2	4	
Conditional inference		4	

The composition of the measurement test is described in terms of the properties measured and concept categories in Table 4.3.

TABLE 4.3

COMPOSITION OF MEASUREMENT TEST

	Conservation	Conventions	Application
Volume	3	2	
Area	2	2	1
Length	2	2	1
Weight	2		3
Temperature	2	2	

As a result of these revisions, all three tests were shortened; the atoms test from 32 to 22 items; the molecules test from 30 items to 22 items, and the measurement test from 30 items to 24 items. This shortening implied a theoretical reduction in reliability, but this loss was considered to be more than offset by other gains. The difference in observed reliabilities between the earlier tests and the revised tests is summarized in the next chapter.

CHAPTER V

ANALYSIS OF DATA

Introduction

Two kinds of statistical analysis are reported in this chapter: analysis of individual item responses and analysis of total test scores. The analysis of item responses identifies those items for which it can be said that ability to respond correctly increased. Further, it helps to identify categories of items which may be especially difficult or not sufficiently challenging; it provides us with information about the feasibility of self-instructional use of the material; and it provides us with a means to identify differences between units. Analysis of total test scores provides a means for describing the overall level of performance and for making comparisons between pupils in different grades.

Description of Tables

Summary statistics describing test performance for pooled groups before and after treatment are presented in Tables 5.1 and 5.2. These statistics include means, standard deviations, and reliability estimates.

TABLE 5.1

SUMMARY STATISTICS (POOLED GROUPS)

Unit	Mean	S.D.	N	R(KR-20)	R'(S-B)
Atoms Pretest	9.838	3.251	235	.5937	.6800
" Posttest	15.016	3.464	237	.7061	.7775
Molecules Pretest	8.775	3.422	223	.6242	.7073
" Posttest	12.530	3.577	196	.6653	.7452
Measurement Pretest	14.701	3.417	184	.6430	.7267
" Posttest	18.263	3.043	171	.6547	.7033

The summary statistics in Table 5.1 do not include a group of culturally disadvantaged children in Decatur. Their performances are summarized in Table 5.2.

TABLE 5.2

SUMMARY STATISTICS - DECATUR GROUP

Unit	Mean	S. D.	N	R(KR-20)	R'(S-B)
Atoms Pretest	10.920	2.855	25	.4428	.5361
" Posttest	12.608	3.240	23	.6315	.7138
Molecules Pretest	8.444	2.910	18	.5283	.6196
" Posttest	8.266	2.669	15	.4232	.5162
Measurement Pretest	12.608	3.117	23	.5493	.6037
" Posttest	15.652	3.016	23	.5412	.5960

In this table, differences in the number of test scores between pre-treatment and post-treatment tests represent students absent on test days. Differences from one unit to the next represent attrition. Reliability estimates were computed using Kuder-Richardson Formula 20. The corrected reliability estimates are Spearman-Brown corrections to compensate for the shortening of the revised tests. The corrected reliabilities estimate the reliability of revised tests if they had been the same length as the original tests. The differences between the corrected reliabilities of the revised tests and the reliabilities of the tests used on the earlier project (Sands et al., 1965) are no larger than could be accounted for by the decrease in the number of alternative responses.

The lower test reliabilities observed for the Decatur group are undoubtedly due to the limited range of performance imposed by the small size of the group.

Item performance is summarized in Tables 5.3 through 5.8. Tables 5.3 through 5.5 describe performances of pooled groups. The item performance of the Decatur group is summarized in Tables 5.6 through 5.8. For the pooled groups the statistical significance of gains in the proportions of pupils responding correctly to items was evaluated using the unit normal curve as an approximation to the binomial distribution curve. Because gain rather than unsigned difference was of interest, the test was one-tailed. For the Decatur group, the sample size was too small to use the normal curve approximation. In lieu of this approximation, a graphic approximation of confidence interval for the appropriate sample size is used.

Those items for which the proportion responding correctly in the posttest is outside the .95 confidence interval for the proportion responding correctly on the first test are marked with an asterisk. These items can be considered as showing significant gains.

Differences between pupils in different grade levels were evaluated by the use of analysis of covariance with pretest scores as the covariate. Summaries of analysis of covariance are presented in Tables 5.9, 5.11, and 5.13. Tables 5.10, 5.12, and 5.14 describe the adjustments and adjusted gains resulting from covariation between pretest and posttest scores.

TABLE 5.3

ITEM GAINS: ATOMS - POOLED GROUPS

	Proportion Correct Responses Pretest	Posttest	Gain	$Z(P_2 - P_1)$	
1.	.3361	.8270	.4909	15.95	**
2.	.4425	.5232	.0807	2.43	**
3.	.8680	.9071	.0391	1.77	*
4.	.6340	.8987	.2647	8.43	**
5.	.1829	.6371	.4542	18.00	**
6.	.2765	.7763	.4998	17.10	**
7.	.4680	.8312	.3632	11.18	**
8.	.1957	.0337	-.1660	--	--
9.	.4595	.7932	.3337	15.01	**
10.	.2723	.6666	.3943	14.02	**
11.	.2936	.5949	.3013	10.35	**
12.	.4425	.7341	.2916	9.00	**
13.	.7617	.7341	-.0276	--	--
14.	.7787	.8312	.0525	1.94	*
15.	.3531	.7510	.3979	13.10	**
16.	.3872	.6582	.2710	8.55	**
17.	.2765	.7890	.5125	17.55	**
18.	.4170	.6455	.2285	6.98	**
19.	.4000	.5907	.1907	5.96	**
20.	.2382	.6160	.3778	11.30	**
21.	.6255	.3037	-.3218	--	--
22.	.7276	.8734	.1458	5.01	**

* Significant at 5 percent level.
** Significant at 1 percent level.

TABLE 5.4

ITEM GAINS: MOLECULES - POOLED GROUPS

	Proportion Correct Responses Pretest	Posttest	Gain	$Z(P_2 - P_1)$	
1.	.3228	.7653	.4425	14.13	**
2.	.1748	.2295	.0547	2.15	*
3.	.5381	.7091	.1710	5.12	**
4.	.4529	.5357	.0828	2.48	**
5.	.3049	.3724	.0675	2.19	*
6.	.9282	.9591	.0309	1.79	*
7.	.4215	.6377	.2162	6.54	**
8.	.4394	.6734	.2340	7.04	**
9.	.3094	.4081	.0987	3.18	**
10.	.3542	.5765	.2223	6.70	**
11.	.2959	.5969	.3010	9.84	**
12.	.6322	.8367	.2045	6.33	**
13.	.3766	.7500	.3734	11.52	**
14.	.2511	.4693	.2182	8.08	**
15.	.2780	.4540	.1760	5.86	**
16.	.3587	.4030	.0443	1.38	NS
17.	.3991	.6632	.2641	8.05	**
18.	.3049	.3112	.0063	.20	NS
19.	.3497	.3877	.0380	1.19	NS
20.	.2600	.4234	.1634	5.56	**
21.	.4843	.6377	.1534	4.56	**
22.	.5381	.7295	.1914	5.65	**

* Significant at 5 percent level
** Significant at 1 percent level

TABLE 5.5

ITEM GAINS: MEASUREMENT - POOLED GROUPS

	Proportion Correct Responses Pretest	Posttest	Gain	$Z_{(P_2 - P_1)}$	
1.	.8532	.9707	.1175	4.342	**
2.	.7119	.8888	.1769	5.108	**
3.	.7717	.8654	.0937	2.919	**
4.	.7173	.8538	.1365	3.963	**
5.	.4782	.9181	.4399	11.515	**
6.	.6739	.8713	.1974	5.506	**
7.	.5217	.7602	.2385	6.243	**
8.	.3750	.7368	.3718	10.043	**
9.	.3586	.4093	.0507	1.382	NS
10.	.3369	.5438	.2069	5.924	**
11.	.7282	.7836	.0104	.305	NS
12.	.6358	.8421	.2063	5.605	**
13.	.8804	.9298	.0494	1.990	*
14.	.7934	.8070	.0136	.439	NS
15.	.6304	.5204	-.1100	--	--
16.	.4619	.6081	.1462	3.834	**
17.	.6630	.8011	.1381	3.820	**
18.	.8967	.9298	.0331	1.422	NS
19.	.5978	.7017	.1039	2.770	**
20.	.8967	.9473	.0506	2.174	*
21.	.1684	.6374	.4690	16.387	**
22.	.5597	.8304	.2707	7.131	**
23.	.9130	.9532	.0402	1.905	*
24.	.0760	.1520	.0760	3.751	**

* Significant at 5 percent level.
** Significant at 1 percent level.

TABLE 5.6

ITEM GAINS: ATOMS - DECATUR GROUP

	Proportion Correct Responses Pretest	Posttest	Gain	$P_2 > P_1$
1.	.5200	.7391	.2391	*
2.	.6000	.5652	-.0348	
3.	.8800	1.0000	.1200	*
4.	.6000	.9130	.3130	*
5.	.2800	.4782	.1982	*
6.	.4000	.5652	.1652	*
7.	.3600	.6956	.3356	*
8.	.1600	.0434	-.1066	
9.	.6400	.6956	.0556	
10.	.4400	.3043	-.1357	
11.	.4400	.6086	.1686	*
12.	.6000	.6956	.0956	
13.	.5600	.4782	-.0818	
14.	.7600	.8260	.0660	
15.	.0800	.3913	.3113	*
16.	.3600	.4782	.1182	
17.	.4400	.8260	.3860	*
18.	.6800	.6521	-.0279	
19.	.6800	.3478	-.3322	
20.	.4000	.3043	-.0952	
21.	.3600	.2173	-.1427	
22.	.6800	.7826	.1026	

* Outside .95 confidence interval.

TABLE 5.7

ITEM GAINS: MOLECULES - DECATUR GROUP

	Proportion Correct Responses Pretest	Posttest	Gain	$P_2 > P_1$
1.	.1111	.4666	.3555	*
2.	.1111	.0666	-.0445	
3.	.8333	.4000	-.4333	
4.	.5555	.1333	-.4222	
5.	.2777	.2666	-.0111	
6.	.8333	.8000	-.0333	
7.	.5555	.4666	-.0889	
8.	.5000	.4666	-.3334	
9.	.2777	.2000	-.0777	
10.	.1111	.3333	.2222	
11.	.1666	.2000	.0334	
12.	.6111	.6666	.0555	
13.	.3888	.7333	.3445	*
14.	.2777	.3333	.0555	
15.	.4444	.3333	-.1111	
16.	.1111	.1333	.0222	
17.	.1666	.6666	.5000	*
18.	.3888	.3333	-.0666	
19.	.3333	.1333	-.2000	
20.	.3333	.2000	-.1333	
21.	.5555	.4000	-.1555	
22.	.5000	.5333	.0333	

* Outside .95 confidence interval.

TABLE 5.8

ITEM GAINS: MEASUREMENT - DECATUR GROUP

	Proportion Correct Responses Pretest	Posttest	Gain	$P_2 > P_1$
1.	.8260	.7826	-.0434	
2.	.8260	.8260	.0000	
3.	.4782	.7391	.2609	*
4.	.5652	.8695	.3042	*
5.	.3043	.7826	.4783	*
6.	.6956	.7826	.0870	
7.	.3043	.5217	.2174	*
8.	.4782	.5217	.0435	
9.	.1739	.3478	.1739	
10.	.3478	.3043	-.0435	
11.	.6086	.6521	.0435	
12.	.5217	.9130	.3913	*
13.	.8260	.9565	.1305	*
14.	.6956	.6521	-.0435	
15.	.4347	.5217	.0870	
16.	.2173	.3478	.1305	*
17.	.5652	.5652	.0000	
18.	.8260	.8695	.0435	
19.	.5652	.6086	.0434	
20.	.9130	.8695	-.0435	
21.	.0434	.3478	.3044	*
22.	.3043	.6086	.3043	*
23.	.8260	.9565	.1305	*
24.	.2608	.3043	.0435	

* Outside .95 confidence interval.

Item Results for Pooled Groups

The increases in proportions of correct responses to items in the test on atoms were significant at the five percent level for 19 of the 22 items. The increases in 17 of the 22 items were significant at the one percent level. For most items, the value of z was beyond the limit of available z-tables which indicates probabilities less than 10^{-7}. Two of the three items (Items 8 and 21) which failed to show significant gains involved distinguishing between models and symbols representing atoms. The third item which failed to show significant gain (Item 13) was concer d with recognition of the fact that scientific knowledge of atomic structure rests on indirect evidence. About 76% of the pupils answered this item correctly in the pretest and about 73% in the posttest. The failure of this item to show gain reflects slight emphasis given to this idea in instruction.

The gains in proportions of correct responses to the test on molecules were significant at the five percent level for 19 of the 22 items. The gains in 17 of the 22 items were significant at the one percent level. All three of the items which did not show significant gains could be considered as disjunctive questions; that is, deficiencies in comprehension of any of several concepts could result in failing to pass the item. For two of the items (Items 16 and 18) the pupil was expected to know that absence of atoms implies absence of matter and vice versa, that energy is distinct from matter in that sense, and that these items of information are relevant to the question. For the third (Item 19) the pupil was required to know that all molecules are alike and recognize that this information was relevant. The item was made more difficult by the fact that one of the foils was true bur irrelevant.

Gains in proportions of correct responses to the test on measurement were significant at the five percent level for 19 of the 24 items. Gains significant at the one percent level were obtained for 16 items. Three of the five items which did not produce significant gains can be considered to involve application of measurement concepts to a concrete situation. One of the items involves measurement of area (Item 9) and although the item was apparently not unusually difficult, the usage of convention might have been confusing, both in the instruction and in the item. Both of the other two items (Items 15 and 18) were sufficiently easy on the pretest that statistically significant gain was difficult to attain. Two of the items (Items 11 and 14) probably failed to show significant gains partly because of their relative ease in the pretest (.7282 and .7934 respectively) and because they dealt with length as an abstraction.

Item Performance for Decatur Group

Although direct comparisons of the Decatur group with the pooled groups is not appropriate, some general observations can be made. Apparently, for the Decatur group nearly all items on tests for all three units were more difficult than they were for the pooled groups--at least on the posttest; fewer items showed statistically significant gains and when such gains did appear, they appeared to be smaller than for the pooled groups. In the unit on atoms (Table 5.6), 9 of 22 items showed significant gains in the proportion of pupils making correct responses. Furthermore, in the posttest, 14 of 22 items were answered correctly by more than half of pupils in the Decatur group as compared with 20 in the pooled groups. In the unit on molecules (Table 5.7), only three of 22 items showed significant gains in the proportion of students making correct responses. In the pretest, seven items were answered correctly and in the posttest only four. Similar information is revealed in Table 5.2. The pretest mean for this group is 8.444; the posttest mean 8.266. In the unit on measurement, the performance of the Decatur group appears to be more nearly comparable to the performance of the pooled groups. Significant gains in proportion of correct responses appear in 10 of 24 items (Table 5.8). In this unit 13 items were correctly answered by more than half the students on the pretest; 19 on the posttest.

Again, it should be emphasized that substantial inferences are not possible from these results, largely due to the great difference in sample size. Nevertheless, comparisons of the descriptive data are highly suggestive of differences between the culturally disadvantaged group and the pooled groups.

TABLE 5.9

SUMMARY OF ANALYSIS OF COVARIANCE: ATOMS

Source	Sums of squared deviations	Degrees of freedom
Group regression coefficients about common regression coefficient	79.923	3
Individual scores about group regression lines	1511.277	197
Group means (posttest) about regression line based on means	282.975	2
Difference between regression coefficient based on means and common regression coefficient between groups	4.396	1
Individual scores about regression lines with slope b_w	1591.200	200
Group means about regression line with slope b_w	287.371	3
Individual scores about regression line for total group	1878.571	203

$F_1 = 4.455 \quad P < .01$

$F_2 = 12.040 \quad P < .01$

$F_3 = 17.784 \quad P < .01$

TABLE 5.10

COMPARISON OF ADJUSTED MEANS FOR FOUR GROUPS: ATOMS

Group	Observed Means Pretest	Observed Means Posttest	Adjustment	Adjusted Posttest Means	Adjusted Gain
I	8.364	13.449	-.206	13.243	4.879
II	10.629	15.750	.086	15.836	5.207
III	10.109	16.543	.053	16.596	6.407
IV	11.559	15.765	.260	16.025	4.466
Combined	9.757	15.156			

TABLE 5.11

SUMMARY OF ANALYSIS OF COVARIANCE: MOLECULES

Source	Sums of squared deviations	Degrees of freedom
Group regression coefficients about common regression coefficient	16.714	3
Individual scores about group regression lines	1736.008	171
Group means (posttest) about regression line based on means	188.709	2
Difference between regression coefficient based on means and common regression coefficient between groups	18.495	1
Individual scores about regression lines with slope b_w	1752.722	174
Group means about regression line with slope b_w	207.204	3
Individual scores about regression line for total group	1960.426	177

$F_1 = .549$ NS

$F_2 = 6.86$ $p < .01$

$F_3 = 9.37$ $p < .01$

TABLE 5.12

COMPARISON OF ADJUSTED MEANS FOR FOUR GROUPS: MOLECULES

Group	Observed Means Pretest	Observed Means Posttest	Adjustment	Adjusted Posttest Means	Adjusted Gain
I	8.674	11.442	.031	11.411	2.737
II	8.071	12.125	-.215	12.430	4.359
III	8.780	14.380	.074	14.306	5.526
IV	9.167	13.000	.233	12.767	3.602
Combined	8.598	12.737			

TABLE 5.13

SUMMARY OF ANALYSIS OF COVARIANCE: MEASUREMENT

Source	Sums of squared deviations	Degrees of freedom
Group regression coefficients about common regression coefficient	11.654	3
Individual scores about group regression lines	845.532	154
Group means (posttest) about regression line based on means	12.139	2
Difference between regression coefficient based on means and common regression coefficient between groups	29.359	1
Individual scores about regression lines with slope b_w	857.186	157
Group means about regression line with slope b_w	41.498	3
Individual scores about regression line for total group	898.684	160

$F_1 = .708$ NS

$F_2 = 3.700$ $.05 > p > .01$

$F_3 = 1.112$ NS

TABLE 5.14

COMPARISON OF ADJUSTED MEANS FOR FOUR GROUPS: MEASUREMENT

Group	Observed Means Pretest	Observed Means Posttest	Adjustment	Adjusted Posttest Means	Adjusted Gain
I	13.265	16.824	-.822	17.646	3.559
II	15.020	18.163	.105	18.058	3.143
III	15.583	19.438	.402	19.000	3.870
IV	15.032	18.484	.111	18.373	3.452
Combined	14.821	18.321			

Comparison of Performance by Grades

The analysis of covariance was used to compare posttest performances of the several grade levels represented in the pooled groups. Analysis provides for comparison of posttest means adjusted for differences between groups on the basis of covariation between pretest and posttest scores. The numbers of cases on which the covariance analyses are based are smaller than those on which summary statistics are based because the analyses require cases for which both pretest and posttest scores are available. When attention is directed to these cases, a clear picture of attrition patterns can be observed in addition to information provided about differences in attainment between groups.

The summaries of analysis of covariance for the three units are presented in Tables 5.9, 5.11, and 5.13. To determine whether or not the differences between means are large enough to be attributed to differences in effectiveness of instruction rather than to random fluctuations in means, the variance of group means about a common regression line is compared to the variance of individual scores about group regression lines having the same slope as the common regression line for pooled scores. This ratio is the familiar F-ratio of between groups variance to within groups variance. This ratio is represented by F_2 in the three summary tables. For two of the three units, atoms and molecules, the probability that obtained variance ratios of the observed size would be obtained by random sampling is less than .01. For the measurement unit, the probability is less than .05.

It is a well-known fac in the measurement of gain that gain scores are negatively correlated with pretest scores. The possibility exists that posttest scores are also negatively correlated with pretest scores. This is not serious unless there are systematic differences between groups in these correlations. Whether this is the case can be determined by comparing the variance of group regression coefficients about a common regression coefficient to the variance of individual scores about group regression lines. This variance ratio is represented in the three summary tables by F_1. Variance ratios represented by F_1 were non-significant for the units on molecules and measurement. For the unit on atoms, the ratio was significant beyond the one percent level. Because most of the spurious variance in between groups variance on this unit involved the second grade, F_2 was retained for this group. Even with this spurious covariation, a greater proportion of posttest variance is accounted for by treatment effects.

Another variance ratio is of interest even though it has no bearing on the validity of the between groups to within groups ratio. This ratio is the ratio of the variance of group means about a regression line based on posttest means to the variance of the difference between a regression coefficient based on means and one based on variance between groups. This ratio determines whether or not the several means are linearly related to pretest scores. For the atoms unit the hypothesis of a linear trend in group means is rejected; for the other two units it is retained.

The analysis of covariance makes it possible to adjust group posttest means for group differences in pretest means on the basis of covariation between pretest and posttest. Comparison of adjusted posttest means and resultant adjusted gains is found in Tables 5.10, 5.12, and 5.14. From these tables, it can be seen that for the atoms unit, the gain was greatest for the third grade, next greatest for the second grade, next greatest for the first grade and least for the fourth grade. The low gain in the fourth grade undoubtedly reflects the rather large size of the pretest mean for this group. For the molecules unit, gains were in the order third grade, second grade, fourth grade, first grade. For the measurement unit, the order was third grade, first grade, fourth grade, second. In this unit, it is doubtful that the difference in gain between adjacent classes in the order can be attributed to differences in effectiveness of the instruction. In general, it seems reasonable to conclude that the materials were most effective for third grade pupils.

Although no statistical comparison was made, it appears that there was a substantial grade effect in attrition of cases. Table 5.15 indicates the number of students at each grade level who took both pretest and posttest in each grade. Examination of this table suggests that most of the attrition occurred in the first grade and that much of this attrition was associated with the unit on molecules. Attrition data combined with results of the analysis of covariance suggest that the materials presented special difficulties for first grade pupils. Special difficulties were not apparent in other grades.

TABLE 5.15

PUPILS COMPLETING UNITS BY GRADES

(Pupils taking Pretest and Posttest)

Unit	Grade I	Grade II	Grade III	Grade IV	Total
Atoms	69	56	46	34	205
Molecules	43	56	50	30	179
Measurement	34	49	48	31	162

CHAPTER VI

CONCLUSIONS

The data obtained from a state-wide trial of the curriculum materials developed in this project indicate that these materials can be used profitably by primary grade pupils who are considered academically superior by their teachers. The data confirm our earlier findings (Sands, Hicklin, et al., 1965) that academically superior pupils in the lower grades can learn content which requires use of cognitive operations typically encountered at more advanced stages of cognitive development. Instruction by means of a self-instructional format was found not to be a negative factor in teacher acceptance; on the contrary, among teachers who felt themselves to be weak in the content area, it seems to have been an important factor in gaining acceptance of the materials. The data suggest that while the materials can be used effectively in the first grade, more effective results can be obtained in the second and third grades. These materials were apparently not effective with a culturally disadvantaged group.

Are the kits effective teaching devices? On the basis of gains registered on pre- and posttest scores, we conclude that the materials are effective in developing concepts associated with the nature of the atom, molecular structure, and measurement. In a previous study, significant gains were made by an experimental group as compared to those made by a control group. In this study, controls were not used. The gains, however, were larger than those made by the experimental group of the previous study. This encourages us to believe that the gains can be attributed to use of the materials.

Were the materials too difficult for students in classes one through four? For the pooled groups the answer appears to be no. In all three posttests considerably more than half the items were answered correctly by more than half the students. This generalization does not apply to the culturally-disadvantaged group. For the intermediate grades, especially grade four, there is a strong indication that the measurement unit is not sufficiently challenging. Pupils in the fourth grade seem to have extensive prior knowledge of the material covered in the unit. For the first grade, extremely small gains and substantial selective attrition suggest that the content of the molecules unit may be at a higher level of abstraction than pupils in this grade can assimilate.

How did pupils in various grades compare in terms of gain in pre- and posttest scores? The patterns of gains for the various grades suggest that the materials are more appropriate for use in the second and third grades than in the first or fourth, and probably most appropriate for third grade use. However, with the exceptions noted, gains made in all grades were sufficiently strong to warrant use of the materials in all four grades.

Are there differences in effectiveness among the three units? Gains in the posttest for atoms were greater than those on the molecules unit. Pretest scores in the molecules unit were higher than in the other two units, resulting in generally smaller gains. Of the three units, the molecule unit seems to be the most difficult.

What is the holding power of the three units? With the exception noted--first grade pupils in the unit on molecules--attrition during the course of the three units was small enough that it could very easily be attributed to extraneous non-instructional sources. This attrition situation suggests that fourth grade, third grade and probably second grade students could successfully complete all three units in a school year. For first grade, the molecule unit would be a defensible omission. A suggested strategy would be to begin instruction with the measurement unit, following it with the atoms unit. The molecules unit could then be started; then terminated if it appeared to be frustrating to pupils.

What was teacher and pupil reaction to the kits? The majority of teachers and pupils seem to have reacted positively to the materials. Sufficient interest has been shown to warrant marketing these kits for continued use.

The data obtained in this phase of the project encourage us to believe that the kits in their present stage of development might be profitably used by academically superior students in the primary grades.

The production of these instructional materials has been turned over to the Illinois State University Foundation. The Foundation will supervise future distribution in the public schools.