DOCUMENT RESUME

ED 214 805 SE 037 097

AUTHOR	Robinson, James T.
TITLE	Evaluation of the BSCS Human Sciences Program.
INSTITUTION	Biological Sciences Curriculum Study, Boulder, Colo.
SPONS AGENCY	National Science Foundation, Washington, D.C.
PUB DATE	Dec 81
GRANT	SED-72-06305-A11
NOTE	213p.
EDRS PRICE	MF01/PC09 Plus Postage.
DESCRIPTORS	Curriculum Evaluation; Elementary School Science; Elementary Secondary Education; Field Tests; Formative Evaluation; Interdisciplinary Approach; *Junior High Schools; *Middle Schools; *Natural Sciences; Program Evaluation; *Science Course Improvement Projects; *Science Curriculum; Science Education; Secondary School Science
IDENTIFIERS	Biological Sciences Curriculum Study; *Human Sciences Program; National Science Foundation

ABSTRACT

Designed for middle and junior high school students, the Human Sciences Program (HSP) represents a three year, interdisciplinary program which focuses on the role of the natural sciences in promoting the cognitive, psycho-social, and moral development of early adolescents. The materials consist of modules of activities designed around major themes with sub-problem areas or clusters. Individual students or small groups select activities of their choice. The evaluation plans, procedures, and results of the project's formative evaluation are presented in this document including the field testing of HSP levels I, II, and III; evaluation of the Level III module, "Knowing"; adult evaluation; and other studies such as readability, attitude, and logical competence. The evaluation indicates, among other findings, that (1) the project was equally effective in different areas of the country and with a variety of teachers and students; (2) that an interdisciplinary program could work in self-contained classrooms, departmentalized schools, and team-teaching contexts; and (3) that student selection of activities was not based upon content, difficulty, or sex of student. (DC)

Evaluation
of the
BSCS Human Sciences Program

by

James T. Robinson

Prepared with support from Grant No. SED 72-06305-All, the National Science Foundation. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author and do not necessarily reflect the views of the NSF.

BSCS
833 W. South Boulder Road
Louisville, CO 80027

December, 1981

Evaluation

of the

BSCS Human Sciences Program

by

James T. Robinson

Prepared with support from Grant No. SED 72-06305-A11, the National Science Foundation. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author and do not necessarily reflect the views of the NSF.

BSCS
833 W. South Boulder Road
Louisville, CO 80027

December, 1981

Preface

The development and formative evaluation of a complex curriculum project such as Human Sciences are documented in many staff-developed evaluation papers. The purpose of this paper is to summarize the results of the formative evaluation and its contribution to the final commercial edition of Human Sciences. Details of findings, data, and data analyses are contained in the staff papers and in the archives of the Human Sciences Project. A bibliography of these papers is included at the end of this publication under the title, "Bibliography of Publications from the Human Sciences Project." Reference to "Staff Papers" with an "SP" designation and "Evaluation Papers" with an "EP" designation will be used throughout this report. These designations refer to papers, published and unpublished, that were produced by the project staff and used during the development and revision of the curriculum materials. Titles will be found in the Copies of these papers are on file in the Evaluation Archives of the Human Sciences Project at the University of Colorado, Norlin Library.

The Human Sciences Project was buffeted by many political forces as the curriculum materials were developed. The freeze on spending that was already authorized in 1973-1974 reduced the budget for the Human Sciences project by 50 percent just as classroom testing in 1973-1974 was starting. Although formative evaluation was to be included in the project from the very beginning, when budgets were cut, development and production of curriculum materials had highest priority on staff time and funds. This summary report captures the essential elements of the relationship of evaluation to development in the preparation, testing, and revision of an interdisciplinary science program designed to meet the needs, concerns, and interests of early adolescents.

A large number of evaluation instruments was developed and used in the formative evaluation of the Human Sciences Program. These instruments have been recorded on microfiche and are available from the ERIC Clearinghouse for Science, Mathematics, and Environmental Education (Robinson, 1981). A set of microfiche cards of these instruments is enclosed in a pocket inside the back cover of this document.

Table of Contents

	Page
CHAPTER 1 HUMAN SCIENCES	1
The Challenge	1
The Need	1
A New Curriculum Model	3
The Innovative Model	3
The Curriculum Structure	4
The Curriculum Development-Evaluation Time Line	5
Summary	5
CHAPTER 2 FORMATIVE EVALUATION OF THE PILOT MODULE	7
The Pilot Materials	7
Design and Instruments	8
Evaluation Design	8
Instruments	8
Test Site Selection	12
Conducting the Field Test	13
Results and Recommendations	15
Summary	19
CHAPTER 3 PREPARING FOR LEVEL I FIELD TESTS	21
Test Site Selection	21
The Evaluation Plan for Level I	22
The Level I Human Sciences Program	24
Preparation of Data for Analysis	24
Looking Forward	25
CHAPTER 4 RESULTS FROM LEVEL I FIELD TEST	29
The Students	32
The Curriculum Materials	35
The Test Group	35
Activity Choice	36
Activity Structure	38
Activity Content	39
Module Structure	39
Evaluation Materials	40
The Program and Program Goals	41
Concluding Comments	42
CHAPTER 5 THE LEVEL II FIELD TEST	45
Evaluation Plan	46
Data Analysis Procedures	47
Results of Level II Field Tests	49
Activity Choice and Usage	49
Review Team Evaluation	51
Essay Problems	52
Objectively-Scored Problems	55
Student Ratings of Essay and Objectively-Scored Problems	58
Teacher Evaluation of the Level II Program	58
Successes and Problems	60

Page

Evaluation of Teachers Guides 60
Guide Organization 61
Add to Teachers Guides 61
Activity Organization 61
Problem Area Organization 62
Module Organization 62
Evaluation of REPRODUCTION 63
Evaluation of Pedagogy 63
Evaluation of Module Evaluation Materials 64
Evaluation of "With Learning in Mind" 65
Science Supplies in HSP Trail Classrooms 65
Summary 65

CHAPTER 6 THE LEVEL III FIELD TEST 69
Evaluation Plan 70
Data Analysis Procedures 72
Results of the Level III Field Test 73
Evaluation of CHANGE 74
What Students Did: Activity Use Patterns 74
What Did Students Accomplish? 76
Did the Activities Make A Difference? 77
Cumulative Effects of Activities and Skill Booklets 78
Evaluation of FEELING FIT 80
Activity Use Patterns 80
What Did Students Accomplish? 81
Did Activities Make a Difference? 83
Skills Development 85
Evaluation of INVENTION 86
What Students Did: Activity Use Patterns 86
What Did Students Accomplish? 93
Did the Activities Make a Difference? 94
Skills Development 94
INVENTION Questionnaire 95
Evaluation of SURROUNDINGS 99
Activity Use Patterns 100
What Did Students Achieve? 102
Work Habits 102
Skills Development 104
Summary 104

CHAPTER 7 EVALUATION OF KNOWING, A LEVEL III MODULE 111
Evaluation Plan 111
Selection of Field Test Sites 112
Instruments Used in the Field Test 112
Results of the KNOWING Field Test 114
Activity Choice Patterns 115
What Did Students Accomplish? 117
Logical Competence 117
Can Students Make Useful Judgments About Curriculum Materials? 117
What Were Students' Attitudes Toward Science Courses? 125
Summary 128

Page

CHAPTER 8 ADULT EVALUATION OF HUMAN SCIENCES.................... 129
Parental Evaluation.................... 129
Evaluating Innovative Science Curricula.................... 133
What Is Your Opinion of Human Sciences?.................... 134
"Goals of Education" Priorities.................... 143
Teacher and Administrator Evaluation.................... 146
Content Review Conferences.................... 154
Summary of Reviewers' Feedback.................... 155
The Public Reviewers' Conference.................... 158
Summary of Reviewers' Feedback.................... 159
Summary.................... 162

CHAPTER 9 ADDITIONAL STUDIES.................... 163
Readability Study.................... 163
Attitude Study.................... 164
Instrument.................... 165
Discussion.................... 166
Logical Competence.................... 167
National Science Foundation Panel Review.................... 172
Follow-up Study.................... 177
Summary.................... 178

CHAPTER 10 HUMAN SCIENCES, A NEW CURRICULUM DESIGN.................... 179
Choice.................... 180
Evaluation Outcomes.................... 180
Student Attitudes.................... 183
Logical Competence.................... 183
Teachers and Human Sciences.................... 183
Activity Structure.................... 184
Summary.................... 187

Bibliography of Publications from the Human Sciences Program.................... 189

Bibliography of Evaluation Papers from the Human Science Project.................... 197

References.................... 203

CHAPTER 1

HUMAN SCIENCES

The Challenge

Members of the BSCS staff conducted needs assessment studies of science programs for early adolescents from 1967 to 1969. These studies resulted in nine guidelines for middle and junior high school science programs (Clark, 1969). The BSCS was funded by the National Science Foundation to initiate the development of a life sciences program for early adolescents in March, 1981. Two important considerations directed the project staff as they worked toward implementing the first steps in the project: formulating a rationale and framework for a three-year interdisciplinary life sciences curriculum, and preparing pilot materials for classroom testing.

The needs assessments emphasized the "overiding point of view that a life sciences program should center on and emanate from the individual student in his (or her) own environment--his (or her) needs, interests, and social responsibilities..." (Clark, 1969, 4).

The request for the grant asked for and received approval to develop an entirely new curriculum unlike any others in existence, using the nine guidelines as a point of departure (Clark, 1969). This opportunity to build an entirely new curriculum rather than to improve the content of existing science courses was a first for the BSCS and for the National Science Foundation.

The Need

Results of the needs assessment (Clark, 1969) and of the work of the staff in the formative stages of the project identified a host of problems in existing science programs and a long agenda of needs not being met by existing science curriculum materials.

A recurrent issue raised by teachers, administrators, and parents was student motivation--"They just aren't interested...," "they won't do their assignments...," "How can I motivate them?," "They don't like science!" Development from childhood to adolescence seemed to be accompanied by a growth in criticism. Compliant children with cooperative

"I'll try it" attitudes seemed to give way, in many instances, to an early adolescent who asked why this should be studied, one who had an increasing number of dislikes and polarized views of school subjects. Students were either "fantastic" or "the pits."

Another recurrent issue was individual variability. Variability seemed to be a pervasive, multidimensional characteristic. Variability from day to day was a problem, but the variabilities that confronted teachers in providing adequate science instruction seemed insurmountable. There were variabilities in basic skills: in reading and comprehending science texts, in mathematics, in simple computational skills, especially in using those skills required for effective laboratory study in science, and in cognitive development--abilities to approach problems systematically, to apply "logic" in reasoning about natural phenomena. There was also tremendous variability in what students remembered from prior instruction, in their abilities to bring past instruction to bear on a new problem, and especially in interrelating subject matter in what appeared to adults to be "obvious." There was also variability in physical development (as much as 24 inches in height and over 100 pounds in weight in a classroom), in physical dexterity and eye-hand coordination, in interpersonal skills, in confidence and the development of positive self-concepts, and in curiosity and a spirit of "I'll try it," "I'll take the risk."

A final and recurrent issue that was critical in devising the curriculum framework was the intensity with which many early adolescents followed their hobbies and special interests. Where many adults expressed concern with "the short attention span" of middle school/junior high students, there were equivalent expressions about how enduring the students' interests were in sports, ham radio, electronics, music, hunting, etc. Early adolescents seemed to get interested in something, pursue it with intensity, drop it perfunctorily, and start something new. This "erratic" behavior was distressing to many adults. In staff discussions with students there was a general finding that students did not relate their out-of-school interests with school subjects. Even students with a home laboratory saw little or no relation of this interest to their science courses.

Existing science curricula and instructional practices--with single texts at a graded reading level, everyone reading the same material, doing the same exercises, following a reading/recitation pattern, and then being sorted on performance relative to others--seemed to put science teachers in a position of fighting the developmental characteristics of early adolescents, and generally loosing the battle. The challenge to a new science curriculum project seemed clear: Could a new science program put the teacher in a more positive role, in facilitating development of these different individuals as he or she teaches science?

Addressing the issues identified in the needs assessment meant creating something new. However, new things, especially curricula, have never been met with open arms by schools. What if a science program could be devised that students liked and from which they learned? Could the adults who formulate school policy accept such a program, especially if it departed dramatically from extant course materials? How far could a new curriculum go in departing from the past?

The needs were clearly established, an entirely new approach to middle school/junior high science would be required to implement the guidelines and to help resolve issues raised by the examination of the existing situation. The evaluation task would be to find out if it would be educational and if people in the schools--teachers, administrators, parents--could cope with it, both in field test situations and as a potential curriculum choice.

A New Curriculum Model

The first eighteen months of the Human Sciences project were devoted to developing a rationale and framework for a three-year interdisciplinary science curriculum for early adolescents, and to design and produce three pilot modules, consistent with the rationale and framework, that would be ready for field testing. A brief description of the entirely new curriculum model that was developed is necessary to provide the context for the formative evaluation tasks of the Human Sciences program.

The Innovative Model

The goal of the Human Sciences project was to produce three years of science curriculum materials for early adolescents. Traditional content-topic curriculum structures were replaced by a model created explicitly to guide the development of the entirely new materials. Four "generic questions" were derived from the needs, concerns, and interests of early adolescents. These generic questions were utilized to focus attention on the contexts that would relate student materials to student lives. Three content themes, the product of several interdisciplinary conferences (BSCS, 1971a,1971b,1971c) were to provide the content sources for the curriculum materials. The three content themes subsumed subject matter from the natural, behavioral, and social sciences. Three aspects of human development that are especially critical in developing science materials for early adolescents provided the third dimension of the curriculum model. Cognitive development as conceptualized by Jean Piaget (Inhelder and Piaget, 1958), psycho-social development as described by Eric Erikson (1959) and Jane Loevinger (1966), and moral development as described by Kohlberg (1969) provided the third dimension of the curriculum model (see Figure 1, page 4).

The model was formulated as a tool for developing curriculum materials that would enhance individual development by placing interdisciplinary content in contexts that would be meaningful to early adolescents. Associated with the model were twenty-four curriculum characteristics (SP 7302-111) that provided the detail needed to produce specific curriculum materials.

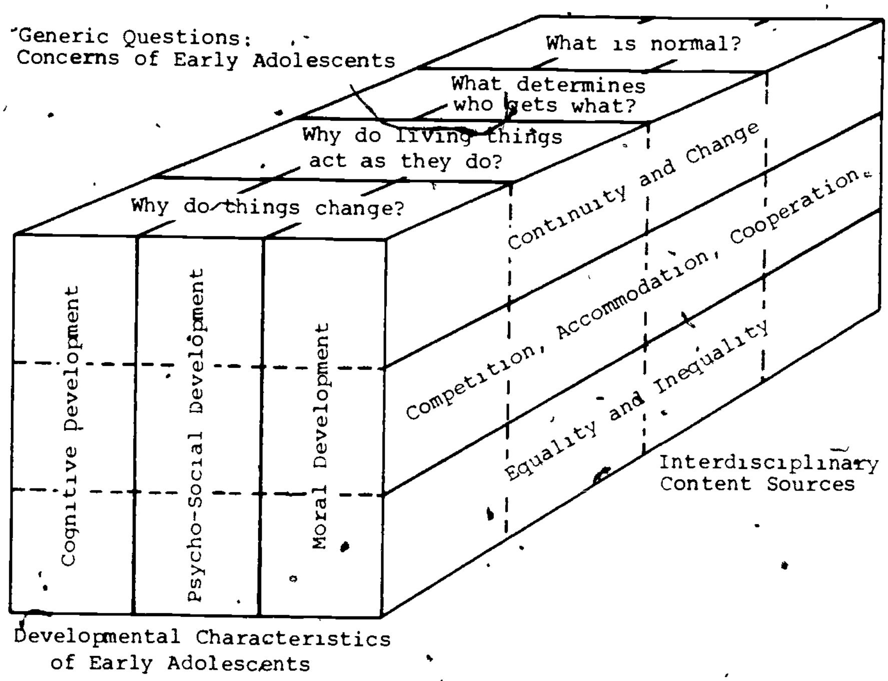


Figure 1. The model framework guiding the developement of the Human Sciences Program.

The Curriculum Structure

A single text for such a varied student population was not considered feasible. Alternatively, several texts, each at a different reading level, were not economically possible. A key decision was made to develop a wide range of materials and to enable students to choose the components they would study. This decision led to the invention of a modular program consisting of large numbers of student activities. The module was conceptualized as a set of related activities developed around a single theme. The module would contain all the materials needed by students to carry out their work with the activities they selected. Choice meant that activities would be designed for individuals or small groups. The activities that comprised modules would need to be heterogeneous in approach to the module theme, in reading level, in cognitive complexity, and in appeal.

The module was to be physically designed to enable a group of students to have access to it in order to choose among the individual components--the activities. With students in a class doing different activities, management of the module and its materials would need to be the students' responsibility. Also, to keep costs down and to provide as great a variety as possible, only a few copies (five to ten) of each activity would be provided. Only single copies of expensive media and equipment would be placed in a module. Students would have to learn (with teacher assistance) to adapt to this overpopulated, under-resourced environment.

The developmental characteristics of early adolescents led the staff to propose that modules at any particular grade level be non-sequential. The importance of novelty and "starting anew" was viewed as more useful than developing modules that would be prerequisites for another. Modules designed for older students would have a central tendency and higher levels of difficulty and/or complexity than modules for younger students. Modules about similar subjects at different grade levels would build on each other, but in such a way that those designed for younger students would not be prerequisites for those designed for older students. By attempting to attain these difficult relationships among modules, a flexible curriculum could result. This flexibility would enable school systems to insert a single module into an existing curriculum, to adopt a single grade level, or to adopt the full three-year program. Flexibility of use was seen as an important structural charcteristic for this innovative program.

The Curriculum Development-Evaluation Time Line

Three pilot modules, designed for exploratory field testing, were developed and produced in the summer of 1972. These materials were tested across the three most common grade levels in middle schools--grades six, seven, and eight. Classroom testing was accomplished in the fall and winter of 1972-1973. Feedback from test sites and formative evaluation data were collected from the test classrooms and used to guide the development of three years of grade level materials.

During the academic year 1972-1973 writers were selected for developing the Level I materials. The materials were developed and produced for field testing in sixth-grade classes during 1973-1974. A similar pattern was followed in developing the Level II and Level III materials. Field tests of Level II were conducted in seventh-grade classrooms in 1974-1975 and Level III in eighth grade classrooms in 1975-1976. The timetable for development and evaluation of the three-year curriculum materials is shown in Figure 2, page 6.

Summary

The Human Sciences Project was funded to create an entirely new, three-year, interdisciplinary program for early adolescents. The typical question of how to simplify science subject matter for a presumably homogeneous age group was replaced with an entirely new and different question: "How can the natural sciences serve the development--cognitive, psycho-social, and moral--of early adolescents?"

Sept '71 Sept. '72 Sept '73 Sept '74 Sept '75 Sept. '76 Sept. '77 Sept. '78

Develop Experimental Modules	Test and Evaluate Experimental Modules	Demonstration Module(s) Available				
	Develop 6th-Grade Modules	Test and Evaluate 6th-Grade Materials	Revision of 6th-Grade Materials	Commercial Materials Available for 6th-Grade		
		Develop 7th-Grade Modules	Test and Evaluate 7th-Grade Materials	Revision of 7th-Grade Materials	Commercial Materials Available for 7th-Grade	
			Develop 8th-Grade Modules	Test and Evaluate 8th-Grade Materials	Revision of 8th-Grade Materials	Commercial Materials Available for 8th-Grade

Figure 2. Timetable for the development of the BSCS Human Sciences materials.

The need for such a new approach to science education at the middle and junior high school level was developed through a needs assessment study funded by the National Science Foundation. Two years after the publication of guidelines, Life Sciences for Middle Schools (Clark, 1969), the Biological Sciences Curriculum Study received a grant from the NSF to develop a curriculum rationale and framework and pilot modules that would exemplify the framework. The result of the initial grant was a unique model for science education for early adolescents.

This brief account is provided to establish the context in which a formative evaluation of Human Sciences was conducted.

CHAPTER 2

FORMATIVE EVALUATION OF THE PILOT MATERIALS

The Pilot Materials

The pilot materials were three non-sequential modules, each designed to be used for approximately six weeks, at the rate of one class period per day. Each module was divided into activities, the student's unit of study. In two modules, ten copies of each activity and related supplies and equipment were presented within small boxes in a module box. A third module had the activities bound into a booklet for each student. Students chose the activities they wished to study, in any order they wanted. Where media or special apparatus were required, only one set was provided. It was anticipated that there would be more activities to choose than any student would have time to do. The module titles and number of activities in each module are shown in Table 1.

TABLE 1
The Pilot Modules

Module Title	Number of Activities
Humanself	12
Developing	11
Learning	14

During the development of the rationale and curriculum framework and planning of the pilot modules, questions that needed empirical data for their resolution were constantly proposed. These questions were used to develop the evaluation design and also influenced the development of the activities in the pilot modules.

At this phase of the project the staff planned to prepare an evaluation design, develop instruments, and gather data that would help to resolve the following issues:

1. What are the appropriate grade placement and cognitive levels for specific content in the materials?
   a. Do different activity designs appeal differentially to students at different grade levels?

b. What proportion of sixth-, seventh-, and eighth-graders will require that the study of a problem be initiated at a concrete level?

c. What proportion of sixth-, seventh-, and eighth-graders will be able to extend studies to the level of formal operations?

2. Can controversial issues be handled in science classrooms by students and by teachers? Will administrators admit such issues into a science curriculum?
3. To what extent do sixth, seventh, and eighth graders vary in interests and in success in studying the social aspects of scientific developments.
4. What issues are raised by sixth-, seventh-, and eighth-graders that suggest additional problems or opportunities that need consideration in developing the experimental grade-level materials.
5. Will the way modules or activities are packaged affect utility and use?
6. Can science teachers handle classrooms in which students will be doing different activities at the same time rather than having the class do experiments and activities as a classroom group?

In addition to these issues, the pilot field test was viewed as an opportunity to test new evaluation instruments and procedures that might be usable in the three-year field test program.

Evaluation Design and Instruments

Evaluation Design

To secure information that would contribute to the resolution of the issues presented above, a Stake (1967) model was utilized to plan the details of the evaluation questions. The questions to be investigated; data to be collected; and data sources for antecedents, transactions, and outcomes are presented in Table 2, pages 10-11.

Implementation of the above plan required:

1. selection of schools and teachers who would test the pilot modules/model materials.
2. selecting and/or designing instruments for data collection.
3. adding personnel to conduct the evaluation study.

Instruments

In negotiating the grant for the period September 1, 1972 to August 31, 1974, the formative evaluation budget was reduced by removing funds for instrument development and reducing the number of classroom observers. As a consequence of this action, staff members on site visits to field test classes served as the primary vehicles for evaluation. Simple data gathering instruments were prepared to answer some of the evaluation design questions.

The instruments used in the pilot module field test included module-specific achievement items; student activity record sheets; teacher record sheets; interview schedules for student, teacher, and principal interviews; and record sheets for use by staff members during site visits.

Achievement items were developed for each module. The items were assigned to be administered individually to students by the teacher, or to be done by the students individually after they had completed several activities in the module being studied. Since the items were to be used for assisting students in choices of additional activities, they were not to be used as end-of-module measures. Teachers were asked to record student responses on audio tape if students lacked writing skills. Three items were developed for the Humanself module, five for Developing, and three for Learning (see EP 7911-54). These achievement items were designed to serve two formative evaluation functions:

1. to provide immediately useful information to the teacher on which to base instructional decisions.
2. to provide data to the staff about the range of explanations that students produced when confronted with the particular problems.

"Activity Record Sheets" were developed for each of the three modules (see EP 7911-54). Test classes were provided with enough "Activity Record Sheets" for every student throughout each module. Students recorded the date they started and the date they completed each activity, circled the word "liked" or "disliked" for each activity they started, and wrote comments about the activity in the space provided. The student "Activity Record Sheets" were designed to gather data directly on the following questions:

1. Which activity was initially chosen by the student?
2. In how many activities did the student participate?
3. How long did the student spend on any one activity?
4. What was the student's reaction to the activity?

During a teacher orientation session, the test teachers asked for classroom record sheets on which they could keep track of the activities the students were doing. The project staff was concerned that such a form would lead to teachers retaining the record-keeping function, a function that was a student responsibility in the curriculum design. The issue was resolved by providing "Classroom Record Sheets" for each module (see EP 7911-54) with strong recommendations that the sheets be posted and that students record the desired information on the sheets.

Structured interview schedules were prepared for use with students, teachers, and principals (see EP 7911-54). The interview schedules were used during staff site visits to field test schools. The student interview schedules were tested by the staff observer in a nearby middle school and were revised several times before being used at the field test site.

TABLE 2

Data Gathering Plan for Formative Evaluation of the Pilot Modules/Model Materials

Questions to be Investigated	Data to be Collected	Sources of Data
ANTECEDENTS: Conditions existing prior to teaching and learning that may be related to outcome.		
1. Did the range of students using the model materials include the diverse school-community situations that make the data, in part or in total, useful in preparing grade-level materials?	General description of ethnic mix in classes General description of socio-economic level of parents of students in test classes General location of school within its community	Classroom observation Teachers and administrators BSCS School Questionnaire
2. What were students' perceptions of school in general and of their science class prior to starting Human Sciences?	Student comments	Student interview protocol
3. Were the activities within modules designed with the simplest cognitive level tasks first, followed by more complex cognitive level tasks?	Reviewer comments Staff evaluation	Reviews of student materials by a variety of specialists Classroom observations
TRANSACTIONS: Encounters of student with teachers, student with materials, student with community and teacher with community that may be related to outcomes.		
4. What were the patterns of activity selection within modules?	Dates when each student began and stopped work on each activity	Student Activity Record Form
5. What reasons did students provide to explain their choices, likes, and, dislikes for particular activities?	Student explanations	Structured (recorded) interview of a sample of students in test classes
6. What sense did students make of the activities they have done in the Human Sciences program?	Student papers, tables, work sheets, etc. Student explanations	Papers and other student products Structured (recorded) interviews Classroom observations
7. What were the patterns of use and non-use of materials and media provided in each module?	Observations Teacher comments	Classroom observations Teacher interview

TABLE 2 (continued)

8. What problems did teachers encounter in using Human Sciences materials?	Classroom observations Teacher comments and observations	Classroom observation by staff Teacher interviews Teacher comments and annotations in Teachers Guides
OUTCOMES: The impact of instruction on students, teachers, administrators, parents and college educators.		
9. What kind of data did students collect for activities and what did they do with data after collection?	Student data sheets and other student products Teacher observations Classroom observations	Student written material Teacher interview tapes Staff observation records
10. How did 6th-, 7th-, and 8th-grade students differ in the depth to which they pursue activities?	Student records, reactions, and comments Teacher observations Classroom observations	Student written materials Student interview tapes Teacher interview tapes Staff observation records Student responses to evaluation activities
11. What differences occurred in the degree to which 6th-, 7th-, and 8th-graders lost interest in each module?	Student records, reactions, and comments Teacher observations Classroom observations	Student written materials Student interview tapes Teacher interview tapes Staff observation records Student responses to evaluation activities
12. What range of responses did 6th-, 7th-, and 8th-graders present in response to evaluation activities?	Student responses to evaluation activities	Tapes and written responses to evaluation activities
13. What problems did teachers encounter in using individually administered evaluation activities?	Teacher reports	Teacher interview tapes and written reports
14. What advantages and disadvantages did teachers identify in using the experimental modules?	Teacher reports	Annotations in Teachers Guides Teacher interview tapes
15. What were the perceptions and reactions of principals to the Human Sciences materials?	Principals' reports	Principal interview tapes

A "Classroom Observation Record" was developed to be used during site visits by BSCS staff. Use of the record was described in the "Instruction Form" (see EP 7911-54).

Since the evaluation study involved schools across the United States, teacher and student feedback letters were seen as an important and immediate means of communication. A student volunteer in each field test class was given the responsibility of mailing comments solicited from classmates to the project staff. Correspondence was acknowledged and replies to questions and comments were returned. Teachers were also asked to make comments periodically and to return their personally annotated teachers guides for each pilot module to the Human Sciences Project staff upon completion of each module. Student products--drawings, graphs, tables, worksheets, etc.--were also collected and sent to the BSCS for review.

Test Site Selection

The minimum number of test classrooms for each test site for each grade level was judged to be six. This number was advocated in order to have an initial group of at least 150 students at each of the three grade levels. The project was funded in September, 1972 to field test three pilot modules/model materials in eighteen classrooms and to add an evaluator and classroom observer to the staff to conduct the field test.

Classes for testing the pilot modules/model materials of the Human Sciences program were selected in the following manner:

1. Letters were sent to all educators who had written to the BSCS expressing interest in testing middle school science materials. These letters were received after publication of <u>Life Sciences in the Middle School</u> (Clark, 1969), the report of the needs assessment study completed by the BSCS in 1969. Over three-hundred letters, including a two-page questionnaire, were mailed to all correspondents.
2. Responses to the letters and completed questionnaires were received from some fifty schools. Because many of the letters were returned unopened due to changes in address, the response was considered adequate.
3. Criteria for selecting the eighteen school test sites were:
   a. expressed interest in testing the pilot modules and potential interest in participating in the grade-level field test.
   b. one or two test sites near the BSCS headquarters
   c. other test sites distributed geographically.
   d. a variety of ethnic groups in test classes.
   e. a variety of school settings--urban, inner city, suburban, rural.
   f. a variety of socioeconomic levels in test classes.
   g. a large enough school to permit both Human Sciences and regular science classes to be offered at the same grade level.
   h. a variety of school organizational structures.

Responses to items on the BSCS "School Questionnaire" (EP 7911-54) were also used in test site evaluation. For example, if it were likely that teachers' strikes would keep schools closed for a long period of time in the fall of 1972, those schools were considered poor risks for pilot module testing. School locations near pilot module writers were also considered if other criteria were met. Arrangements for test sites were made with school administrators at both building and school district levels. Teachers for the field test classes were selected by their principals.

Eighteen school sites were selected initially. A nineteenth site, the University of Wyoming laboratory school--Unversity School--was added after many conversations and visitations of personnel from their Science and Mathematics Teaching Center facility. The Wyoming site was to receive materials only, with the University assisting the teachers and supporting their participation in the teacher orientation conferences for field test teachers. Table 3, page 14, presents descriptive data about the selected test sites.

Conducting the Field Test

The BSCS staff consultants, teacher associates, and a full-time observer were selected as the primary instruments for evaluation. This decision was based on the judgment that existing objective measures were inadequate for the purposes of evaluation, that funds were inadequate for developing and validating new instruments, and that the staff would be interacting and advising writers of the grade-level modules. In such interactions, experience derived from classroom observation was judged to be potentially greatest value.

A full-time observer was assigned to the sixth-grade field test class at John Dewey Middle School in Denver, Colorado. The observer was in the test classroom a full week prior to the initiation of the Human Sciences materials to test observation protocols. Observations were recorded, observation schedules were tested, and student interview protocols were practiced. As part of the training program the observer followed an individual student to all classes for a day. The observer also observed classes in Louisville Middle School, Louisville, Colorado, as often as possible.

All field test classes were visited between the second and sixth weeks of study, as shown in Table 4, page 15. The purpose of this visit is described in Appendix I. Classes were visited while students were using the first experimental module, <u>Humanself</u>. During the two-day visit, classroom observations and interviews were scheduled. On day one, Human Sciences classes were observed, students were interviewd, and the teachers' problems and concerns were discussed. On day two, the same classes were again observed, students were interviewed, and the teacher and principal were interviewed. All interviews were conducted from written interview schedules (see EP 7911-54), were tape recorded, and were then transcribed.

TABLE 3

Test Classes for the Pilot Modules: Humanself, Developing, and Learning, 1972-1973

SCHOOL LOCATION	SCHOOL	DEMOGRAPHIC CLASSIFICATION	GRADE	ETHNIC COMPOSITION[a] (%)				NUMBER OF STUDENTS			TEACHER
				BLACK	ORIENT	CHICANO	ANGLO	GIRLS	BOYS	TOTAL	
Chalmers, Indiana	Frontier Middle	Rural	6th	0	0	0	100	19	14	33	Davis
Denver, Colorado	John Dewey Middle	Suburban	6th	0	1	1	98	11	9	20	Spensieri
Detroit, Michigan	Luddington Middle	Lge city/urban	6th	95	1	0	4	13	15	28	Macinkowicz
Louisville, Colo.	Louisville Middle	Rural	6th	0	0	2	98	10	9	19	McLellan
Mansfield, Ohio	Mansfield South, Elem. & Jr. High	Suburban	6th	2	0	0	98	12	14	26	Beer
Portland, Oregon	Whitaker Middle	Lge city/urban	6th	15	5	0	80	8	12	20	Puckett
Shoreham, N. Y.	Shoreham Middle	Suburban	6th	3	1	1	95	13	12	25	Aykroyd
	TOTALS, GRADE 6			-	-	-	-	86	83	169	
Atlanta, Georgia	Walden Middle	Lge city/urban	7th	100	0	0	0	21	14	35	Smith
Laramie, Wyoming	University Sch.	Urban	7th	1	0	0	99	16	13	29	Alchediak/ Abelson
Lugoff, S. Carolina	Lugoff-Elgin Mid.	Rural	7th	25	1	0	74	16	16	32	Freeland
Madison, Wisconsin	Marquette Middle	Urban	7th	3	0	0	97	17	17	34	Slominski
Oakland, California	Havenscourt J.Hi.	Lge city/urban	7th	90	2	5	3	18	21	39	Reynolds
Philadelphia, Pa.	T. Fitzsimons J.H	Lge city/urban	7th	97	0	0	3	16	10	26	Holloway[b]
	TOTALS, GRADE 7			-	-	-	-	104	91	195	
Baltimore, Maryland	Lansdowne Middle	Suburban	8th	0	0	3	97	17	13	30	Corley
Detroit, Michigan	Cadillac Jr. Hi.	Lge city/urban	8th	84	4	0	12	11	12	23	Ostenfeld
Freeport, Illinois	Carl Sandburg Mid	Urban	8th	10	0	0	90	14	13	27	Larson
Laramie, Wyoming	University School	Urban	8th	1	0	0	99	11	15	26	Abelson/ Alchediak
Los Angeles, Calif.	Belvedere Jr. Hi.	Lge city/urban	8th	6	3	88	3	16	15	31	de Mordaigle
Oakland, Calif.	Havenscourt J.Hi.	Lge city/urban	8th	80	0	5	15	9	30	39	Reynolds
	TOTALS, GRADE 8			-	-	-	-	78	98	176	
	TOTALS, ALL GRADES							268	272	540	

[a] Percentage of ethnic composition is an estimate provided by the BSCS staff from classroom observations.

[b] Philadelphia schools were on strike during the trial period. No data were secured from this site.

JTR:cl

TABLE 4

Project Staff Visits to Field Test Schools, Fall 1972

Location	Grade	Teacher	Observation/Visitation Schedule				
Denver, CO	6	S	Observations daily throughout field test period				
Louisville, CO	6	M	Observations intermittently throughout field test period				
			Oct.	Nov.	Dec.	Jan.	Feb.
Chalmers, IN	6	F		⊢⊣			⊢⊣
Detroit, MI	6	K		⊢⊣			⊢⊣
Mansfield, OH	6	D		⊢⊣			⊢⊣
Portland, OR	6	O	⊢⊣				⊢⊣
Shoreham, NY	6	C		⊢⊣			⊢⊣
Atlanta, GA	7	L		⊢⊣			⊢⊣
Laramie, WY	7	B					⊢⊣
Lugoff, SC	7	H		⊢⊣			
Madison, WI	7	R	⊢⊣			⊢⊣	
Oakland, CA	7	P		⊢⊣			
Philadelphia, PA	7	I	⊢⊣				
Baltimore, MD	8	E		⊢⊣			
Detroit, MI	8	N		⊢⊣			⊢⊣
Freeport, IL	8	J		⊢⊣			
Laramie, WY	8	A					⊢⊣
Los Angeles, CA	8	G		⊢⊣ ⊢⊣			
Oakland, CA	8	P		⊢⊣			

On the first site visit, teachers were asked to have up-to-date Student Activity Record Sheets available to the observer upon arrival, and to recommend five students to interview: the most satisfied, the least satisfied, a high achieving, an average achieving, and a low achieving student. The observer selected four additional students to interview, using the Student Activity Record Sheets as a data source. Students were selected according to these criteria: high number of activities chosen, low number of activities chosen, high number of activities marked "liked," and high number of activities marked "disliked." Observers used the Student Interview Schedule previously described. They interviewed and recorded on audio tape as many of the nine potential student interviewees as possible.

Results and Recommendations

The geographic and community diversity represented by the pilot test sites enabled the evaluator to have confidence in the generalizability of the evaluation results to other middle and junior high schools in which the administrators are interested in change. This latter limitation was imposed because the sites were not randomly selected.

Sites were located in each of the four major census regions of the United States. Inner city, urban, suburban, and rural schools were included in the test sites. Gathering data on the ethnic background of students was not permitted in 1972-1973. Estimates made by staff members on site visits showed approximately 61% white, 32% black, 6% Chicano, and 1% Oriental students in the combined sixth-, seventh-, and eighth-grade classes.

The cognitive complexity of activities was assessed by a developmental psychologist (EP 7211-27), who reviewed the three pilot modules. His review suggested many ways in which activities could be improved. His suggestions were incorporated into the development of the grade-level activities and again in reviewing activities for final revision after classroom testing. His general comment was:

> "Before noting any 'possible' problems in the curriculum, let me say perfectly candidly, that the modules and overall idea of the project indicates one of the most exciting projects I have seen in education. You capture the thrust of the Piagetian-Bruner focus in the student as active learner and perhaps empiricist. What I am 'criticizing' should be seen as work to be done, rather than any rejection of the overall premises or idea of the program". EP 7211-27, p. 1).

Activity selection patterns (see EP 7912-55) indicated:

1. Sixth graders tended to choose activities that focused most heavily on self (that is, "Self-Study Booklet") while eighth graders selected activities about what other people think ("What Do You Think of That?").
2. More activities were chosen and completed by sixth graders than by eighth graders.
3. Activities about food and shelter were least chosen.
4. Activities in which students interviewed adults or other students were among the most highly chosen activities, while activities in which students worked with black and white prints, or which were largely reading, were least chosen.
5. When animals were provided in activities, the activities were highly chosen. Most students liked working with animals, but dislike of some animals--for example, flies and spiders--reduced choice of some activities.
6. Eighth-graders found some activities too easy. In the development of the three-year curriculum materials, most activities from the three pilot modules were included in Level I modules.

Activity choice patterns provided insight into those phenomena and approaches that most interested early adolescents. The specific data were used in two ways. First, they provided information for the kinds of contexts in which the subject matter of activities would be seen as relevant to early adolescents. Second, they showed the work that would need to be done to make important scientific content appeal to and have meaning for early adolescents.

The individually administered test items were difficult for teachers to handle. Students were not used to writing, much less used to providing a written explanation of objects or events. For future development of activities, the following findings were of note:

1. Precision in measurement was not of importance to students. Activities need to be designed to demonstrate the value of precision.
2. Most students could describe the phenomena they studied. A small percentage of students could employ higher cognitive processes such as analysis and inferential thinking. This finding was consistent with the assumed cognitive levels of the target population, confirmed by results of the logic test given at the end of Level I and Level III test years.
3. Graphing problems were difficult for many students. Interpreting simple graphs was done well by most, but when students were asked to construct graphs from data, most could not determine the proper axis for a graph. Recommendations were made to provide a variety of graphing activities in the grade-level materials to give students more practice in preparing and using graphs.
4. Most students were very literal in response to essay questions. They gave brief, generally accurate, but limited responses.
5. The evaluation items were all application items. Teachers reported that the items were too difficult for students and that they did not (students and sometimes teachers) see the relation between the items and activities. For example, metamorphosis of the frog was studied. The evaluation item used the salamander as an example. Student responses were consistent with their literal orientation.
6. Success in application did occur, however. Twenty-three percent of students in a sixth-grade class and 41 percent of students in an eighth-grade class related a series of drawings of an organism they had not studied to growth and development.
7. Teachers did not like evaluation questions with answers graded into different levels of understanding. They preferred questions with right or wrong answers. The view of concept development held by the project staff--that concept development is a process of simple beginnings with development providing multiple connections of concepts to other concepts, objects, and events--was not viewed as useful.

In reviewing teacher comments in Teachers' Guides and from teacher interviews, advantages and disadvantages of the pilot modules were found. Many of the citations below were from single teachers. The most pervasive advantage expressed was the motivation and interest of students in engaging in activities. The most pervasive disadvantage exposed was the openness of activities and "lack of structure."

<u>Advantages</u>

- Discipline problems were held at an absolute minimum.
- Students enjoyed doing most activities and especially liked choosing their own work, working at their own speed, and at times playing the role of teacher.

<u>Disadvantages</u>

- The large posters were impractical.
- Couldn't rate student as to what he or she was learning.
- Few students were motivated to work on activities at home.

- Interest is high, absenteeism is down, and make-up work due to absence is eliminated.
- There were activities for all ability levels except the poorest readers.
- All students, with help, could find activities in which they could succeed.
- Students learned to work independently, decreasing teacher dependency.
- More students are task oriented than in all-class discussions.
- More students were thinking about what they were doing than when they do a regular science experiment or exercise.
- Students didn't keep asking, "Why do we have to study this?"
- Lack of class discussion left many students missing parts of the activities.
- Activities had too little structure to enable many students to learn.
- The classroom was too noisy.
- Average and below-average students felt insecure with choice; some even felt threatened.
- Some students were not capable of making a choice and staying with it.
- Animal activities had a mixed reception. When teachers facilitated, students were very involved.
- In some test classes, teachers did not permit any live animal activities. Vacation periods (Christmas) and weekend building temperatures that were harmful made animal use difficult in cold climate areas.

Interview transcripts[1] were obtained from ten of the test school principals. Tabulations of principals' major responses are listed in Table 5.

TABLE 5
Comments by Principals in Pilot Module Test Schools about Human Sciences, N=10

Comments	Number of Principals
Principal was involved with students	4
Parents' response was favorable	5
No parent responses	7
Valued student learning	4
Valued positive student attitudes	8
Positive about progress	6
Noted positive affect in teachers	5

[1]Interview transcripts were available from ten of the eighteen pilot module test school principals: six sixth-grade, two seventh-grade, and two eighth-grade.

Summary

The pilot modules field test provided valuable data for the development of the grade-level modules. Particular details for the improvement of activity structure and module structure were summarized for use in activities that would be developed later (EP 7912-55).

The field test demonstrated that science teachers could shift instruction from all-class laboratories and exercises to several individual and small-group activities being conducted in the classroom simultaneously.

The use of living materials was not common to the field test teachers and many had had little or no experience with living materials in their classrooms. Students responded enthusiastically to live materials and in many test classes learned to manage and care for the organisms used. However, the field test indicated that many difficulties in teacher preparation, school facilities, and teacher attitudes would be encountered for any curriculum that required the use of live materials.

There were no reports of problems with pregnant animals, mating of mammals, or birth of young. Eight of the eighteen test classes reported criticism of two posters of the human figures that were engraved on the Explorer space vehicle--drawings of a mature male and female figure.

In two classrooms teachers took a laissez-faire approach that was clearly unsatisfactory. The successful teachers were active throughout each class period, talking with students about what they were doing and why. One of the test teachers duplicated activities and assigned them as all-class activities. This mode was not successful, as students needed to secure materials from home for some activities. Without the commitment of choice, they continually failed to bring the needed materials to class.

One eighth-grade student summed up his experiences this way:

The reason I like this science program is because:

1. it's a different way to get to know your fellow classmates better.
2. you can learn more about thjings at the same time.
3. I had time to walk around and learn.

CHAPTER 3

PREPARING FOR LEVEL I FIELD TESTS

The 1972-1973 pilot testing provided new insights into how activities and modules could be more effectively designed and structured. It also emphasized the resources that would need to be allocated to evaluation, if formative evaluation were to make an optimal contribution for improving the curriculum product. This chapter will describe the formative evaluation plans of the experimental grade-level modules.

Testing a three-year curriculum program required school systems that would commit schools to participate for a three-year period. It required a commitment by parents to allow their children to participate in the evaluation study. It also required sixth-, seventh-, and eighth-grade teachers and administrators to agree to the field test with only the pilot modules to show what the program would be like.

Modules to be tested for the first year of Human Sciences were designated Level I and were tested in sixth-grade classes in 1973-1974. Level II modules were tested in seventh-grade classes in 1974-1975. The final year of testing was at the eighth-grade level in 1975-1976.

Test Site Selection

The sites for the grade-level materials field testing were selected primarily from the sites used in the pilot test. Since only seven sites were to be used, the selection was made from schools with high teacher interest in the Human Sciences philosophy and strong administrative support for field testing. Schools were asked to agree to schedule three Human Sciences classes, and at least two teachers for testing the Level I materials. They also were asked to allow students to transfer out of the experimental materials classes if parents desired a transfer. Schools were also asked to provide enough classes each year to accommodate all students who agreed to continue in test classes. New students could be added at any time to maintain the necessary pupil-teacher ratios.

Test sites were also selected for geographical distribution, demographic characteristics, organizational type, agreement to permit observers in classes, and to permit teachers to miss some school days for teacher orientation and debriefing conferences during the field test period.

The seven schools selected for the three-year field test are shown, with initial enrollment data, in Table 6, page 23. Two schools were elementary schools--one with self-contained classrooms, and one with a three-teacher team in an open classroom facility. Students from these two schools transferred to junior high schools for the final two years of testing. The other five test schools were middle schools. Within-school organization varied in different years at many sites; therefore, particular test conditions will be described for each school in each appropriate section.

Enrollment figures provided in Table 6 were compiled from data brought to the teacher orientation conference prior to the initiation of field testing Human Sciences. There were 330 boys and 342 girls in the test classes. Enrollment varied throughout the year with the proportion of boys to girls remaining quite constant.

Each test school used Human Sciences in three classes. Two of the nineteen teachers taught two test classes; the remaining teachers taught one test class each. Class size ranged from fifteen to thirty-nine students with a mean class size equal to thirty-two. There were twenty-one class groups participating in the Level I field test.

The Evaluation Plan for Level I

An elaborate evaluation plan, modified from the field test of 1972-1973 was developed during the spring and summer of 1973. The grant renewal proposal to begin September 1, 1973 included a full-time evaluator for the project. Grant renewal was delayed and on October 1, President Nixon rescinded 50 percent of all federal projects. Evaluation was cut from the 1973-1974 budget, as development of the materials had highest priority.

Part of the evaluation plan was initiated in the spring of 1973 on prior grant funds by holding community seminars at each proposed test site. These seminars included district and building administrators, school board members, parents of fifth-grade students (who would be the parents of sixth-graders in test classes in 1973-1974), science and/or social science educators from a nearby college or university, and sixth-grade teachers who would be field test teachers in 1973-1974.

The purpose of the community seminars was to acquaint participants with the program, to focus on the cognitive developmental and other characteristics of early adolescents, and to establish relationships with each group for purposes of evaluation in the ensuing three-year field test program.* Unfortunately, loss of anticipated funding for this part of the evaluation plan precluded the evaluation effort that was initiated at these seminars. Additionally, two school

TABLE 6

Three-year Test Sites[1] for the Human Sciences Program, 1973-1974

Region	Schools & Grades	Number of Teachers	Number of Classes	Number of Students		
				Girls	Boys	Total
Midwest Large city, Urban	Middle School 4 (6, 7, 8) Detroit, Michigan	2	3	47	45	92
Midwest Suburban	Middle School 6 (6, 7, 8) Madison, Wisconsin	3	3	40	44	84
South Urban	Middle School 1 (6, 7, 8) Columbia, South Carolina	2	3	40	40	80
Mountain States Suburban	Elementary School 2 2 (K-6) Lakewood, Colorado	3	3	60	58	118
Southwest Suburban	Elementary School 3 (K-6) San Jose, California	3	3	55	60	115
Northwest Urban	Middle School 7 (6, 7, 8) Portland, Oregon	3	3	41	47	88
East Coast Suburban	Middle School 4 (6, 7, 8) Baltimore, Maryland	3	3	47	48	95
ALL	7 schools	19	21	330	342	672

[1]Students at School 3 transferred to Junior High School 9 and those at School 2 transferred to Junior High School 8. These two junior high schools replaced the elementary schools in 1974-75 and 1975-76.

sites were changed between April and July, 1973 so that in two of the seven test sites, community seminars were never arranged. A more complete account of the community seminars is presented in a draft paper, "Community Involvement in Curriculum Change," in Appendix K.

The Level I Human Sciences Program

Five modules were developed in the summer of 1973 for testing in the academic year 1973-1974. These modules were, in the order in which they were produced and tested, BEHAVIOR, SURVIVAL, SENSE...OR NON-SENSE?, LEARNING, and GROWING. LEARNING and GROWING were based on the pilot LEARNING and DEVELOPING module. Draft manuscripts for a second HUMANSELF were prepared in the summer Writing Conference of 1973, but was judged by the staff to require more time to bring it to field test level than was available. Each module was designed to provide from five to seven weeks of instruction. Teachers were asked to terminate a module when a significant group of students in a class had chosen and completed all of the activities they wished to do. The purpose of this procedure was to avoid a loss of interest by students in any module.

The five modules were not sequentially related and could have been tested in any order. The major time constraint was for GROWING since it required observation of plant and animal development, best observed in the spring of the year.

Each module was designed to provide more activities than any student could do within the time period provided. No specific activities were required. Students could choose activities in any order. Ativity cards were provided in quantities of one each. Where expensive equipment was part of an activity, only one set of the necessary equipment was provided. Thus, students had to manage the cycling of activities and equipment through different students in a class.

Preparation of Data for Analysis

As has been described in detail in the archive papers for each module, evaluation data of several types were gathered for each module. Common to all moduels was a form on which students evaluated the activities they chose and indicated whether they had or had not completed the activity. These records were the major source of data for determining the number of students in the field test classes in 1973-1974. The other source for determining both who (as individuals) and how many students participated in the field test was the class list provided by each participating teacher. Table 6 shows that 672 students were reported by name in test classes prior to the introduction of the first module. These students were assigned school, teacher, and individual identification numbers.

Forms for evaluating activities (to be described later) were to be completed by each student when an activity was terminated. These forms were collected by test teachers and mailed to the Human Sciences Project staff. If the form received could be assigned an identification number, the student was counted as having chosen the activity where a title was reported on the form.

In assembling the data file, all students who had any data collected in 1973-1974 and in the subsequent two years were retained in the data file. For students with data for only 1973-1974, the student was dropped if he or she had data from two or less of the five Level I modules and had no end-of-year data. Eighty-three students were dropped, leaving a total enrollment of 589. Two individuals who were dropped had data marked in each of the three years, but entered, dropped, and re-entered classes with more time out of class than in class.

Two hundred and seventy forms for evaluating activities had no student name, school identification, or were otherwise illegible. These forms are not part of the data base. Two of the eighteen test teachers sent in limited numbers of activity forms. These two limitations, plus the self-report nature of the data source for determining the number of activities done, both absolutely and relatively underestimate the actual use of activities in test classes.

Five hundred eighty-nine students form the class enrollment data base for analysis of the Level I modules. Table 7, page 26, shows the reported enrollment by class for students who were assigned identification numbers. It also shows the number of students retained by grade and teacher group. Most of the students eliminated were not enrolled in Human Sciences classes for more than a few months. Those of teacher 3 in test school 4 are the exception. Students in this group tended not to elect Human Sciences in subsequent years, hence the large number eliminated.

In addition to the records for each student, the two hundred and seventy forms previously mentioned had activity titles and were included in the analysis. These non-coded forms are missing data from the standpoint of student records and comprise 2.8 percent of the forms received. A quantity of data is also missing. These data include failure by students to prepare or present forms to their teacher and failure of the teacher to forward forms to the Human Sciences Project staff.

The data from the "Activity Record Forms" were the only data from the Level I modules that were prepared for computer processing. All other data were analyzed by hand.

Looking Forward

The next chapter will present a summary of the results of the field tests of each of the six Level I modules in the order in which they were tested. Additionally, the results of other evaluation activities that were not module specific, but were conducted in the 1973-1974 academic year will be presented. A summarization of the major outcomes from the Level I testing will complete the chapter.

The evaluation materials developed for 1973-1974 were desigend primarily to enable teachers to work with students on self-evaluation.

TABLE 7

Number of Students in Human Sciences by Class, Teacher, School, and Gender; and Number of Level I Activity Record Forms Available for Processing

Teachers/Schools	Reported Enrollment Boys	Reported Enrollment Girls	Reported Enrollment Total	Activity Records Boys	Activity Records Girls	Activity Records Total
Teacher 1 Classes 1&2	31	25	56	31	24	55
Teacher 2 Class 1	9	15	24	8	14	22
School 1 Total	40	40	80	39	38	77
Teacher 2 Classes 1&2	33	29	62	22	28	50
Teacher 1 Class 1	13	13	26	12	13	25
School 5 Total	46	42	88	34	41	75
Teacher 1 Class 1	14	15	29	14	15	29
Teacher 2 Class 1	13	13	26	12	13	25
Teacher 3 Class 1	9	16	25	8	16	24
School 6 Total	36	44	80	34	44	78
Teacher 2 Class 1	17	23	40	11	22	33
Teacher 1 Class 1	23	21	44	17	14	31
Teacher 3 Class 1	26	16	42	20	14	34
School 2 Total	66	60	126	48	50	98
Teacher 1 Class 1	21	16	37	17	16	33
Teacher 2 Class 1	19	23	42	17	23	40
Teacher 3 Class 1	16	23	39	1[illegible]	21	33
School 3 Total	56	62	118	46	60	106
Teacher 1 Class 1	15	13	28	15	13	28
Teacher 2 Class 1	13	15	28	12	15	27
Teacher 3 Class 1	11	17	28	7	15	22
School 7 Total	39	45	84	34	43	77
Teacher 1 Class 1	16	17	33	16	17	33
Teacher 2 Class 1	17	18	35	17	17	34
Teacher 3 Class 1	15	14	29	3	8	11
School 4 Total	48	49	97	36	42	78
Total of All Teachers/Schools	331	342	673	271	318	589

Each instrument and procedure was to serve this end first, and was to be analyzed secondarily for use in formative evaluation. Resources did not permit the development of achievement measures for the activities in modules. This was a most unfortunate circumstance, as innovative programs need correspondingly innovative materials for student evaluation and grading. Staff visits to test classes were again considered necessary for maintaining relations with test school teachers, administrators, and parents, as well as to serve formative evaluation functions.

Partial restoration of funds was made to the project in mid-year and the funds were used for development functions. The funds were received too late to affect module-level evaluation processes, but limited funds were utilized to develop and administer end-of-year instruments to provide limited base-line data for the field test. Funds did not permit employment of an evaluation specialist for the project so the same staff of four conducted all development and evaluation activities, with the aid of two Teacher Associates and a classroom observer. Teacher Associates were middle school or junior high school teachers on leave from their school districts for the academic year and summer to assist the project staff.

CHAPTER 4

RESULTS FROM THE LEVEL I FIELD TEST

Five modules were tested in seven school sites with 19 teachers and 568 students during the academic year 1973-1974. The major staff activities for this academic year were producing the Level I modules, preparing the specific designs of the Level II modules, recruiting writers for the Level II modules, conducting planning conferences, and preparing for Level II development activities to be done in the summer of 1974. Evaluation was fitted in around these development activities. Nevertheless, a large amount of evaluation data was gathered during the 1973-74 school year, and a great deal was learned about the design of modules and activities. These findings were used in the development of Level II materials and became the basis for the revision of Level I materials in two stages. The first stage was through a writing group in the summer of 1974 and the second was through the efforts of the staff in preparing the Level I materials for commercial release.

This chapter will briefly review the major findings that have been documented in evaluation papers (EP 8002-57, EP 8003-58, EP 8005-60, EP 8006-61, EP 8007-62, EP 8008-63). Additionally, further interpretations of the data and remembrances of this writer as these data were reviewed are incorporated into this summary statement.

Six of the nineteen test teachers' experience was in self-contained classrooms. Three of these teachers taught Human Sciences in this context, along with other sixth grade subjects. Three teachers taught Human Sciences in a large open space. They taught other subjects to sixth-graders as part of their teaching schedule. Only one of these six teachers had more than the minimum amount of science in their college preparation. These teachers taught the science part of a four-hour team teaching program. None of these teachers were science specialists. The ten remaining teachers taught science as their major teaching assignment. These teachers were prepared as science teachers. One teacher in the test group was not positively polarized toward the Human Sciences Program. Whether the attitude of this teacher was neutral or negative is not known, but the field test data received were limited and fragmentary. Only four students of the twenty-three who were in this class remained in Human Sciences test classes.

The Teachers' Guides produced for the Level I materials provided limited help for teachers. Even though a five-day teacher orientation workshop was held in the fall of 1973 to review module materials, experience and discuss demonstration teaching, and to discuss appropriate teaching strategies, the need for additional teacher materials was clearly called for as the test year progressed. A few teaching ideas were incorporated into the SENSE...OR NONSENSE? Teachers' Guide, the LEARNING Teachers Guide, and the GROWING Teachers' Guide, but these were clearly not adequate for the needs of teachers.

In addition to this need for additional support materials, there was a need for institutional support within the school for Human Sciences. Although this was not documented, the presence of students outside of classrooms, in the halls, and even off campus was not viewed positively by some teachers within a few test schools. Although the need for administrators and test teachers to inform other teachers within a school building as to what was going on within experimental curricula was known by the BSCS staff, it was not communicated effectively to other teachers at the test school sites. A few schools took measures to involve other teachers in Human Sciences either through faculty meetings, awareness conferences, or invitations by test teachers to visit Human Sciences classes. As a result of this variability, the reception of this "different" curriculum in schools varied considerably.

Teachers found the difficulties of evaluation and grading the most serious problem of Human Sciences. Next to evaluation and grading was the problem of classroom management. There was a great deal of variability in teachers' abilities to evolve from being the sole manager of the classroom environment to involving students, at the level of their competence, in this process. Learning classroom management skills and the acceptance of self-responsibility by students was not a goal that many science teachers felt relevant to "science instruction." At the intellectual level they agreed with the importance of this goal, but at the practical level of day to day operations in the classroom, implementation of the goal--that is taking time to work with students, to explore management problems, to develop plans for improving management in the classroom, and to assess and revise those plans periodically over the year--was not consistently pursued in many test classrooms.

The need for better testing, evaluation, and grading plans and materials that was found during Level I evaluation influenced the resource allocations and plans for developing the Level II modules. It became apparent during the 1973-1974 school year that an innovative curriculum that departed from standard curricula (where every student is studying the same subject matter and is graded in comparison with other students, or in comparison to some "standard" sets by the teacher) presented problems that could not be resolved within the normal curriculum development process. When such an innovative curriculum is planned and funded, a very large proportion of resources, both financial and human, needs to be allocated to the development and testing of student evaluation instruments. To do this effectively, it must be realized that the development and testing task to solve this problem is probably of the same magnitude as that of developing the curriculum and should receive an equal allocation of human and material resources. These resources are in addition to those required for formative and/or summative evaluation.

Reduction of project funds reduced the number of site visits and reduced the amount of data collected from teachers in 1973-1974. There were no funds within the budget to pay teachers for collecting and forwarding the evaluation materials to the project staff. Site visits helped to stimulate the flow of such materials back to the project staff. Reduction in the number of site visits reduced the quantity and quality of materials received from the teachers. It also reduced the human support for teachers, and human support is essential for an innovative project where the teacher cannot look at the materials and obviously see what the pedagogical tasks are.

Holding the first site visits within one month to six weeks of the time school started proved to be a very effective technique for teacher support. If site visits could have been followed by several more visits, at least three more during the school year, there may have been more effective communication to teachers and to principals about the objectives of the program. Teacher support materials could have been provided verbally during these visits.

Teachers found the management of organisms difficult. The energy crisis resulted in the lowering of school thermostats during holidays and weekends. The lowering of thermostats posed severe problems for activities involving the growth and development of plants and animals. Additionally, many teachers at the middle school level had no experience in the care and maintenance of living plants and animals in the classroom. This lack of knowledge, and also the lack of support materials, made it difficult to handle modules that used living things. Students asked for more animals and plants in future activities. Teachers were at best lukewarm toward the idea.

Attitude scales and opinionnaires administered to teachers at the end of the school year indicated a general positive feeling toward the use of Human Sciences in the seventh and/or eighth grades. Of course, for many of the test teachers there was no opportunity for them to continue with the field test classes because either they were in an elementary school, or they were not interested in or could not be assigned to teaching seventh- or eighth-graders.

The results of the first year of testing showed that the field test teachers learned:

- to teach science classes in which many different activities are going on at the same time.
- to treat students as individuals with unique needs and skills and to assist them in selecting the most effective combination of activities for their growth and development.
- that grading and evaluation, even with quantitative data, is a judgment teachers make. Teachers can devise cooperative evaluation and grading programs with students.
- to work with students on solving management problems and to make the management problems associated with a complex multimedia curriculum program a valuable learning experience for students.

During the first year of field testing Human Sciences proved that:

- it is equally effective in different parts of the country with a variety of teachers, and with a wide range of student backgrounds and abilities.
- interdisciplinary studies--selecting content and methodologies from the biological, social, and behavioral sciences--could be accommodated in science departments and in team teaching contexts in middle and junior high schools.
- activity choices of eleven- to fourteen-year-olds were not clustered by content or difficulty, nor were they influenced by the grade levels of the students.
- students could learn to manage an environment (the classroom) that had scarce resources and was overpopulated.
- reading, writing, and arithmetic skills were utilized meaningfully in contexts where students needed them to solve problems of their choice.
- students could improve their skills in self-direction and reduce the need for continuous supervision in a bounded free-choice environment, where they could choose from within a provided curriculum.

The Students

The students in Human Sciences came from a diversity of backgrounds. They were not a random sample of sixth-graders in the United States, but they represented a wide range of that population. As can be seen from the cities in which test classes were located (see Figure 3, page 33) test sites were located in most of the major geographic regions of the United States. The sites on this map included the Dissemination Centers in which Human Sciences materials were used in demonstration teaching, 1974 to 1976, and the field test sites for the Level III KNOWING module, tested in the spring of 1977.

Boys and girls were equally represented in the 569 students from whom data were obtained. The mean age for the group at the end of the sixth grade was twelve years, one month, but a standard deviation of six months and a range of forty-eight months indicated that there was a wide range of chronological ages within the group.

The development of the Level I Human Sciences materials was based on an assumption that early adolescents would be clustered primarily at the concrete operational level of cognitive development. This assumption was translated into several guidelines for activity structure. The results of the test "How Is Your Logic?," administered in May, 1974 EP 7410-03), confirmed this assumption, but showed that many students were not capable of performing concrete operational thought. Teachers' comments and staff observations suggested that there was a group varying between 5 and 15 percent of the students who were still in the preoperational stage of cognitive development.

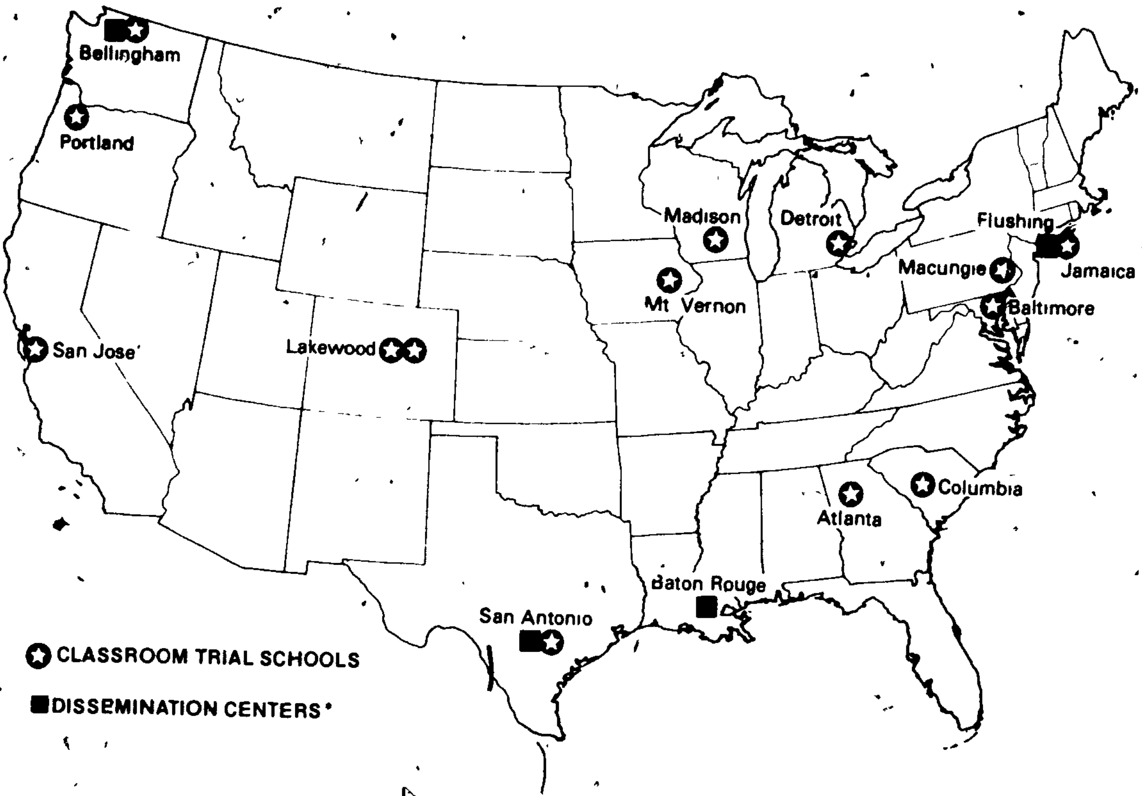


Figure 3. Human Sciences test sites, 1973-1976 and 1977, and Dissemination Centers, 1974-1976.

The logic test showed a hierarchical structure in cognitive development. Success on simple concrete items was necessary for success on more complex concrete items and success on concrete items was necessary for success on formal items. Using the hierarchical structure predictively and referring to Figure 4, a schematic "box and whisker" plot

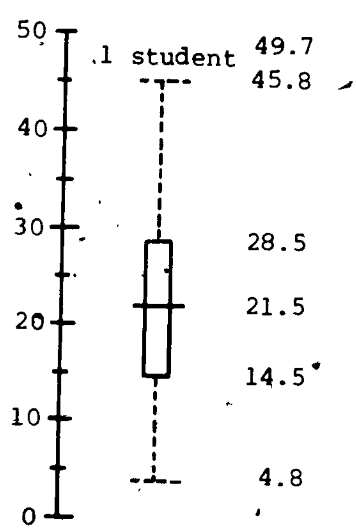


Figure 4. A schematic plot showing the median, hinges, adjacent, and outside values of the weighted scores of students on twenty-three items of "How Is Your Logic?," N=573.

(Tukey, 1977), a general inference can be made about the logical competencies of the test group. There were eight items requiring early concrete operational competencies that were weighted 1 point each. There were six consolidated concrete operational problems weighted 2 points each, six early formal operational problems weighted 3 points each, and three formal operational weighted 4 points each. Assuming the hierarchical structure, a score of 6 points or better would indicate that students could solve early concrete operational problems. (This and subsequent scores assume 75 percent correct for a problem group to indicate competence in the cognitive stage.) A score of 15 points or better would be required to demonstrate concrete operational competence. Half of the test group had scores below 14 points (median, 14.8). Only eight students (assuming complete success on concrete items--a score of 20 points--and 75 percent success on formal items--a score of 13 points) in the test group were capable of solving content-free problems that required formal operational thought. The avoidance of constructing activities that presumed formal operational competence was clearly supported by both field test experiences and by the results of the logic test. Further analysis of the 1974 logic test results will be presented in Chapter 8.

The test group students fell within the predicted theoretical ego development group (Loevinger & Wessler, 1970) of "self-protective," with a small group in the impulsive stage and a very small group (7 percent) in the more advanced conformist stage. The small group of impulsive students (about 18 percent) could present a theoretical problem in a bounded free-choice[1] curriculum environment (see EP 8101-48).

Students' attitudes toward a regular science program were not assessed in the formative evaluation in 1973-1974. For a few students, Human Sciences was the first and only science course they had experienced. Results from the "What's Happening?" instrument, a thirty-eight-item attitude scale, indicated that students concluded their first test year of Human Sciences with positive attitudes on nine of thirteen factors reflecting the structure of the instrument (see EP 7909-44).

At the end of the first year, the Human Sciences program had produced student attitudes that were:

- highly positive toward the science course at the end of one year.
- positive regarding the intellectual challenge of Human Sciences.
- very positive about the self-direction provided in Human Sciences.
- positive about Human Sciences in comparison to other classes.
- supportive of their development of self-confidence.

Level I materials were designed for early adolescents who were believed to be concrete operational thinkers, who could, with the guidance of a facilitating teacher, learn the skills of self-direction and personal responsibility necessary to function in a bounded free-choice

[1]Bounded refers to the condition that student choice was generally limited to the activities within a problem area or module. Free-choice denotes the fact that no activities were prescribed. Students built their own curricula within the boundaries of Human Sciences.

environment. Test data from the end of the school year indicated that many students in test classes [illegible] this expectancy. However, there was an indication that a small portion--from 15 to 20 percent--of the student group did not meet all of these expectancies. These students had developed neither concrete operational competence nor the level of ego development (conformatory) that would enable them to function most effectively in the Human Sciences classroom.

The Curriculum Materials

The Test Group

Data to determine student choices of activities and their evaluation of activities were gathered throughout the year on the five Level I modules that were field tested. The number of students from whom data were obtained varied with each module that was tested. Figure 5 shows the total number of students who were enrolled at anytime in 1973-1974 (total), the number of students who provided data for at least two modules (adjusted), and the number from whom data were obtained for each Level I module that was tested.

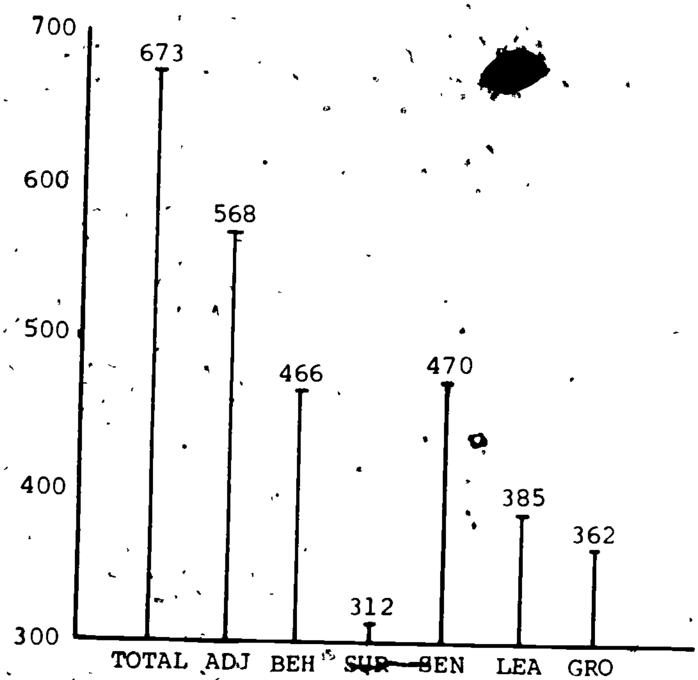


Figure 5. The total number of students enrolled in Human Sciences during 1973-1974, the adjusted total--those who were in test classes for at least two modules--and number of students from whom data were obtained for evaluating each of the five Level I modules.

Activity Choice

Students had the opportunity to choose from about 150 activities in the five modules tested in Level I. Table 8 presents calculations of the mean number of activities completed in each module from the number of activity forms that were turned in by all students divided by the total number of students in the group (569). This calculation includes all students in the data pool regardless of whether or not they turned in any activity evaluation form data for each particular module. The mean activities completed ranged from a high of 4.83 in BEHAVIOR to a low of 1.97 for the GROWING module. In every module, girls turned in more activity evaluation forms than did boys (see Table 8).

TABLE 8

Number of Activity Forms Completed for Each Level I Module (students with no data in some modules are included)

Module, Group		No. of Activity Forms	Mean	S.D.	Variance	N
BEHAVIOR,	All Students	2,749	4.83	4.26	18.15	569
	Boys	1,098	4.26	3.70	13.69	258
	Girls	1,651	5.31	4.63	21.40	311
SURVIVAL,	All Students	1,313	2.31	3.07	9.43	569
	Boys	510	1.98	2.76	7.61	258
	Girls	803	2.58	3.29	10.80	311
SENSE...,	All Students	2,683	4.72	4.21	17.75	569
	Boys	1,195	4.63	4.35	18.90	258
	Girls	1,488	4.78	4.10	16.85	311
LEARNING,	All Students	1,779	3.13	3.34	11.17	569
	Boys	783	3.04	3.67	13.46	258
	Girls	996	3.20	3.05	9.30	311
GROWING,	All Students	1,122	1.97	2.34	5.47	569
	Boys	478	1.85	2.22	4.90	258
	Girls	644	2.07	2.44	5.93	311
TOTAL		9,646	16.95			569
	BOYS	4,064	15.75			258
	GIRLS	5,582	17.95			311

If the students who did not turn in any activity evaluation forms for each particular module are removed, the mean number of activity forms increases to a high of 5.90 for BEHAVIOR and a low of 3.12 for GROWING (see Table 9, page 37). We also note in Table 9 that the number of students missing from each module ranges from 96 to 254. These data indicate that the completion of an "Activity Evaluation Form" by students or the return of such forms to the Human Sciences Project is likely to be incomplete. It is unlikely that as many as 250 students were absent for the full period of a module. It is unlikely even that as many as 96 students were absent for a module; therefore, the interpretation that many students either failed to turn in activity evaluation sheets or that their teachers failed to forward them to the Human Sciences Project, is tenable. It is not known, however, how many of these students actually failed to choose any activities within each module. This unknown factor confounds the interpretation of the mean number of activities chosen in each module. It also confounds the total number of activities studied on the average by students during the school year 1973-1974.

TABLE 9

A Comparison of Activity Use in Level I Modules, based on "You Are the Judge" Forms, with Students without any Forms in any Module Removed from the Calculations

Module	Mean	S.D.	N	N of Students with No Data
BEHAVIOR	5.90	3.98	466	100
SURVIVAL	4.21	3.03	312	254
SENSE...	5.71	3.98	470	96
LEARNING	4.62	3.10	385	181
GROWING	3.12	2.25	359	207
Total	23.55	--	--	--

Although activity choice was not the subject of direct formal investigation, the data presented in this section support the positive reception of activity choice by students, teachers, parents, and administrators.

Activity choice was viewed as a major positive contribution to student motivation and to the observed task-oriented behavior that was so evident in many Human Sciences classes. Choice accommodated the wide range of developmental levels found among early adolescents. It made it possible for slow students to complete several activities successfully, and to feel a sense of accomplishment from their efforts. It also made it possible for other students to choose a great many activities and to accomplish many learning tasks without being held back by the class as a group. Choice also made it possible to include activities that would not be feasible if required of all students.

Choice made it possible to include activities in the modules that enabled a small group of students to carry out a community-based study (a field trip) without requiring teachers to arrange a field trip for the entire class. Planning the field trip, arranging for parental transportation, and arranging the necessary school and community permissions were part of the learning opportunities offered by such out-of-class activities.

Choice had its limitations. When activities were linked so that data from two activities were to be re-examined and compared as the content of a third activity, choice for the third activity was quite low. This pattern of choice cannot be strictly attributed to its building on other activities, for this kind of activity generally required complex cognitive operations. The choice patterns of this activity were similar to other activities with complex cognitive operations.

Choice was a concern to those who felt that all students should have a basic core of activities. This concern was expressed in informal discussions with both educators and parents, but did not represent a majority viewpoint.

With regard to choice, it was found that:

- students learned to use their time constructively in the bounded, free-choice environment of Human Sciences.
- the Human Sciences Program had a major positive influence on student motivation and commitment to learning.
- choice made grading students difficult because activities were of different durations and difficulties.
- teachers found that choice provided effectively for individual differences, enabling the full range of students to be productively engaged in activities that had personal relevance. It was found, however, that some modules in Level I did not provide enough complex activities for the most able students.

Activity Structure

Production of the early modules in Level I was being completed as the data from testing the pilot modules were being prepared for interpretation. Therefore, the findings from major data sources were similar to the outcomes from testing the pilot modules. Unfortunately, these outcomes were not assimilated adequately to make an impact on the first three Level I modules developed and field tested. Additionally, fund reductions prevented the greater use of art work in directions to students and allocation of greater resources to student evaluation materials. These findings were delayed in the implementation until Level II modules were developed.

The statements below summarize the major findings from the Level I field test that provided further guidance to the development of student activities:

- the assumption that activities should be introduced in concrete ways and not require formal operational competence was a correct assumption in terms of student competence in logical thinking.
- activities should be designed for students to complete independently or with a partner. Activities requiring a group of three or more were generally beyond the group skills of early adolescents.
- activities with living organisms were popular and valuable for students unless they required formal logical competence, or well developed conceptual schema for their accomplishment.
- activities need structure and explicit directions for students rather than providing suggestions for what might be done.

- introductory prose should be limited so that students can get actively engaged in something early in an activity. Exposition, clarification, and elaboration, where needed, can follow action.
- activities selected as most valuable by students required continuous thought and action throughout. When thought could be delayed--where the "doing" could be done without thinking--student ratings were lower.
- activities had limited appeal to a limited group of students where reading was the only "doing."

Activity Content

The major content criticism of Level I activities was that too many activities were restricted to a single discipline rather than being interdisciplinary. Secondly, most persons educated in the natural sciences felt that too many activities were based in the social sciences.

The complete exclusion of the physical sciences in Level I was criticized by some students, parents, and teachers. The general science tradition of sixth-grade science made many people uneasy about a science program that had no physical sciences activities.

Content reviewers found the content of most activities in Level I accurate as field tested. There was a group of approximately 18 to 20 of the 150 activities tested that were rated as needing extensive revision or replacement. Some reviewers questioned the place of several social sciences activities in a "science" course (see EP 7704-19).

Public reviewers also found the activities worthwhile and felt that most parents in their communities would approve of these activities. Some questioned several social sciences activities as having no place in a "science" course. From this observation and the similar one for content reviewers, it can be inferred that a small group of adult reviewers did not accept interdisciplinary studies as an option in curriculum design.

A most important finding about content was that when activities were potentially controversial, parents were more often supportive of including them in the school curriculum than were school administrators or science department chairpersons (see EP 7704-18).

Module Structure

Only one variation of the planned module structure was tested in Level I. In all but one module, activities were housed in a module box, with each activity in a separate compartment. Students selected an activity to study, secured the necessary materials, and went to work. In many instances most students never read a significant number of activities that were available in any module. In LEARNING, activities in each of the three problem areas were bound into booklets. Class sets of each of the booklets were provided, and students were issued a booklet when a problem area was opened for study. The evaluation of this format for a module was overwhelmingly negative. Single activities in boxes or other compartments were clearly preferred.

The color coding of activities was found to be a valuable management resource. But at the same time, color, rather than the conceptual design title of problem areas, was the commonly used referent by both students and teachers.

The hypothesis that modules should only be used for five to seven weeks so that students could start anew in a fresh subject was confirmed in the field test. A full year program for sixth-graders, with 45 to 55 minutes per day every day for science, would be best served by six modules in Level I.

Evaluation Materials

During Level I field test a variety of evaluation materials were used for student record keeping, student evaluation of activities, teacher observation records of students, and class records charts. Forms for students to evaluate activities proved to be useful for formative evaluation, as did the <u>Teacher Observation Record</u> and student journals. However, teachers did not find these latter two materials useful. They also wished to have achievement measures for each module or perhaps for each activity. Students were also dissatisfied with evaluation and grading practices.

Test teachers had based their grading practices on normative procedures, in which students were compared and sorted. Activities were of different durations and difficulty levels making the number of activities completed a very limited data source for grading. The other data sources for grading students were activity products-- worksheets, quantitative data, and constructions of various kinds. Since all students in class did not choose any one activity, there was no common ground for the usual normative mode of evaluation or grading. Unfortunately, the developers could not allocate resources to develop specific materials to fill this void. Instead, they suggested ways for teachers to interact with students and how to subjectively assess their growth and development. They left it up to the teachers to translate their own findings into evaluation data and eventually into grades. This placed a burden on teachers that most could not handle effectively. The various suggested materials and procedures were later prepared as a working paper for use in discussing evaluation systems for the revised materials (see SP 7601-46).

In the case of evaluation, testing, and grading, the following generalizations are supported by field test data:

- Students need to learn skills in keeping records and in self-evaluation. These can be learned, but most curricula provide no opportunities for such learning.
- Teachers need specific, well developed, and simple materials for evaluating and grading students. This development effort could not be made within the constraints of the development and testing of

Level I. To accomplish this task, resources at least half as large in magnitude of the development of the curriculum itself would be required in an innovative curriculum project.

The Program and Program Goals

The field test of Level I made it possible to make general statements about the attainment of program goals. These statements are provided within the limitations of the data gathered, and were further evaluated in the field testing of Level II and Level III. The goals are not listed in any particular order.

Goal: to help students develop curiosity about and motivation to study the natural and social worlds about them.

Data from the classroom observations of task-related behavior and from the "What's Happening?" questionnaire support the interpretation that this goal was being attained.

Goals: to enable students to use science process skills and logical thinking.

to enable students to use decision-making skills.

Data from content analysis and from the choice component in the program support attainment of this goal. Additionally, data from parents, teachers, content reviewers, administrators, and students provided valid data that these skills and their development were as important as the attainment of skills such as reading, writing, and arithmetic. The also pointed out the value of choosing and evaluating activities in the development of decision-making skills.

Goals: to enhance students' knowledge and acceptance of themselves--their body, mind, feelings, attitudes, interests, and values.

to enhance students' knowledge and acceptance of and empathy for others--other students, teachers, parents, and those older and younger than themselves.

Both content analysis of the materials and data from the field test indicate that activities in which the human organism was the object of study were considered valuable. Data from parents, teachers, administrators, and students validated the study of human beings as a legitimate object of study in middle school/junior high school science classes. Providing opportunities for students to discuss with, question, interview, observe, and otherwise interact with peers, adults, and younger children were also viewed as legitimate and important.

Goal: to enhance the basic skills of recording and following written directions; communicating orally and in writing; gathering, displaying, and interpreting quantitative data.

Content analysis of activities supported the presence of these opportunities in the activities in Level I. Early in the school year teachers were concerned about students circumventing the reading of directions by doing activities as they had seen others do them. The structure of the program--few all-clas activities--forced students to find alternatives to relying on the teacher to tell them how to do some tasks. In regular science programs, the teacher usually uses a great deal of class time explaining the text and especially interpreting what a laboratory is all about. The learned dependence on the teacher that students had developed was transferred in Human Sciences classes to other students. As the year progressed, most students improved in self-reliance. Teacher assistance, in the form of helping students turn to the activity and not other students for help, was required for many students to progress toward this goal.

Goals: to enhance students' appreciation of science as a way of gaining knowledge about the natural and social worlds.

to enhance students' range of interests about and understandings of the natural and social worlds.

to enhance awareness that there are many modes of learning and sources of knowledge that service a variety of human purposes.

Data from content analysis showed that the program offered many opportunities for development toward these goals. Data from "What's Happening?" also supported attainment of these goals, except for the goal of understanding. No assessment of this part of the second goal statement was included in the formative evaluation of Level I. The last goal was initiated through the diversity of activities in Level I but was not formalized at that time in the program.

Goals: to enhance self-esteem due to personal success in the program.

to enhance responsibility for their own learning.

Data from student interviews, teacher feedback, "What's Happening?," and "What Is Your Opinion of Human Sciences?" support the inference that many students were making improvements toward these two goals. No direct assessment of self-esteem was conducted. However, responses about self in relation to others, as measured by two factors of "What's Happening?," indicated that many students were uncertain about themselves. Attainment of these goals would not be expected at the end of one-third of the program. Also, one would not expect attainment by large numbers of sixth-graders on developmental grounds.

Concluding Comments

For the very small investment in evaluation, the data gathered from the field testing of Level I yielded a great deal of information about modules, activities, evaluation, and the Human Sciences Program in general. Some of this information was fed immediately into the planning and development of the Level II materials. Other data were not assimilated in time, nor in some instances were resources available to implement them in time to influence Level II. Many of these latter results were used in the development of Level III materials.

The field test of Level I raised questions about the inclusion of physical science materials in what was initially conceptualized as a life sciences-oriented program. It also raised questions about a second kind of "balance": How much and what kinds of activities oriented to the social and behavioral science could be tolerated in a program that was designed to be interdisciplinary, but in the practical world of the schools was to fit into the niche occupied by existing courses in the natural sciences?

CHAPTER 5

THE LEVEL II FIELD TEST

Four modules were tested in seven school sites with thirteen teachers and between 310 and 490 students during the 1974-1975 academic year. Five school sites were the same as in 1973-1974 (see Chapter 3), but students in two elementary schools transferred to junior high schools within the same school districts. Table 10 shows the fall, 1974 enrollment in test schools and classrooms with experienced and inexperienced teachers and students. "Experienced" indicates that the individuals had participated in testing Human Sciences in one or more previous years.

TABLE 10
Number of Experienced and Inexperienced Teachers and Students in the Seven Field Test Schools, Fall, 1974

	Teachers			Students			
School	Experienced	Inexperienced	Total	Experienced	New	Total	Percent Retained
1	1	1	2	59	7	66	76.6
4	0	2	2	82	19	101	84.5
5	1	0	1	48	3	51	55.2
6	0	1	1	56	0	56	70.0
7	1	2	3	67	21	88	79.8
8	0	2	2	59	9	67	46.4
9	0	2	2	42	19	61	35.6[1]
ALL	3	10	13	412	78	490	61.9

[1] Half of the students enrolled in Human Sciences were transferred to a second junior high school contrary to a prior agreement with the school district. Funds did not permit extending the number of test sites so these students were excluded from further participation in the field test.

All teachers were provided with an orientation session of two and a half days in Boulder, Colorado, with the Level II materials and with the evaluation procedures that would be used. Four modules--RULES, WHERE DO I FIT?, PERCEPTION, and REPRODUCTION--were to be tested during 1975-1975. The modules were produced, distributed, and field tested in the order presented above. Field testing began the first week in October, 1974.

Evaluation Plan

The major evaluation effort for 1974-1975 was to develop self-evaluation materials and procedures for student and teacher use. The initial plan, to be tested in the RULES module, was a self-evaluation system involving individual evaluation activities at three time periods coupled with group evaluation--"Review Team"--activities.

Each student was provided with a records folder for each module for recording when each activity chosen was started and completed. When each problem area was to be replaced by introducing a new problem area, several evaluation class periods were scheduled. Students were to complete the individual evaluation activities first and then to participate in the Review Team Activities, as explained in the Teachers Guide section, "Facilitating Self-Evaluation," and the guide given to each student, "Evaluating Your Progress." The rationale for evaluation in the Human Sciences Program was presented in an experimental edition of the program teachers guide, <u>With Learning in Mind: A Guide for Human Sciences Teachers</u> (SP 8111-132).

The self-evaluation system was devised to meet the recommendations that were proposed during the field testing of Level I. The system provided materials to enable each student to collect data during the study of each module. These data were to be kept in the records folder for the module and were to include a record of activities chosen, samples of work accomplished (such as completed worksheets, papers prepared, products produced), and evaluation papers. The folder, when completed, would represent a portfolio of work accomplished during the module and would be the central material used to determine the student's grade.

Achievement was to be determined informally by student-teacher interaction and formally by student responses to essay questions. The essay form was selected because of the open-ended nature of most Level II activities. General suggestions were provided to teachers for grading the essay problems, but these were not completed and made available until January, 1974 (see SP 7509-41).

Since students were allowed to choose the activities they studied, essay problems were provided in groups, one for each of the evaluation periods scheduled for each module. Students were to choose one or two essay problems to answer at each evaluation period. The essay problems were answered on NCR® paper so that a copy of the student's response could be sent to the BSCS; the original was kept for the student's portfolio.

Review Team evaluation problems also used NCR® paper for recording the results of review team tasks. Review Team problems were designed to engage students in organizing and synthesizing ideas for the activities studied. The problems were designed to have small groups of students--three to five--work together to discuss the activities; search for commonalities; and reorganize, categorize, or describe reasons for choice, or evaluate activities. For example, the first Review Team problem in RULES asked a team to sort activity cards into two piles, one of the piles being the activities that at least one member of the group had chosen. Then the team was to discuss each of the "chosen" activities, group similar activities into categories, and give the category a name.

The self-evaluation activities for each of the Level II modules were to serve as sources of information for formative evaluation. Site visits during the year, a feedback conference of the Level II teachers in May, 1974, and the employment of an experienced Human Sciences teacher to observe two to three days per week in one test school completed the evaluation plan for Level II.

The initial plan for self-evaluation was used in RULES. Teacher feedback indicated that evaluation activities were too long. The Review Team activities took two to three class periods and were judged to be too difficult for seventh graders. Students complained that they would much rather have time to choose more activities than do the evaluation activities. Teachers were not able to help students understand the purpose or value of the evaluation activities, nor were they valued by the teachers.

With the initial response, the Human Sciences staff decided to eliminate the Review Team activities. The difficulties teachers had with evaluating and grading essay problems caused them to press the staff for the development of multiple-choice and other item forms that could be objectively scored. The REPRODUCTION module, with fewer open-ended activities and more activities with right or wrong outcomes or convergent thinking products, became the first module to include objectively-scored item forms. Table 11, pages 48 and 49, provides a list of the evaluation materials provided for each of the Level II modules.

Data Analysis Procedures

Data from the activity records for RULES were transferred in the Human Sciences office to optical-scan sheets, converted to computer cards, and processed using SPSS, BMP, and BMDP computer programs. Activity data from the other Level II modules were transferred to printed optical-scan sheets by each student.

Essay examinations, administered periodically during a module were the only achievement data collected for RULES, WHERE DO I FIT?, and PERCEPTION. Protocols were developed for scoring those "Choose Your Problems." Processing for RULES was done by hand. Processing for the remaining modules was done by computer.

The REPRODUCTION module achievement instruments included multiple-choice and essay problems, and a separate booklet with self-report problems for skills development and attitudes. Evaluation booklets were returned to the project office, responses were coded onto optical-scan sheets, with the resulting cards processed by computer.

TABLE 11
Evaluation Activities for the Level II Modules

Module	Problem Area	Number of Activities	Individual Evaluation Activities	Group Evaluation Activities
RULES	Is there a rule?	18	Rules Record My Activity Record 2 of 9 essay problems Samples of best work	Review team problems 1 and 2
	What should I do?	14	Rules Record My Activity Record 1 of 5 essay problems Samples of best work	Review team problem 3
	How do rules change?	12	Rules Record My Activity Record 2 of 12 essay problems Samples of best work Grading (optional)	Review team problems 4 and 5
WHERE DO I FIT?	Where do I fit as a person?	22	Where Do I Fit? Record Samples of best work	None
	When do I fit as a person?	11	When Do I Fit? Record Samples of best work	None
	When do I fit in the future?	11	When Do I Fit? Record Samples of best work	None
PERCEPTION	Perceiving	19	Perception Packet My Activity Record 2 of 20 essay problems Samples of best work	None
	Using Perceptions	14	Perception Packet My Activity Record 1 of 10 essay problems Samples of best work	None
	Exchanging Perceptions	14	Perception Packet My Activity Record 2 of 16 essay problems Samples of best work	None

TABLE 11 (continued)

REPRO-DUCTION	What's going on inside?	13	Reproduction Report Sample of best work	None
	What's going on between?	9	Reproduction Report Sample of best work	None
	How does reproduction affect the family?	14	Reproduction Report Sample of best work	None
	Evaluation for module		Evaluation Booklet 1 38 essay and multiple-choice problems	None
			Evaluation Booklet 2 Essay and multiple-choice problems	None
			Evaluation Booklet 3 13 "Skills I developed" 10 "Feelings"	None
			My Activity Record Forms 1 and 2	None

Results of Level II Field Tests

Activity Choice and Usage

The first three modules tested were used from eight to twelve weeks in test classes. WHERE DO I FIT? was used longer as the printing of PERCEPTION was delayed. REPRODUCTION was not tested in two of the seven test sites. One of these schools had a long teacher's strike, the other closed early in May. The remaining five school sites used REPRODUCTION from a minimum of three or four days to three weeks.

Table 12, page 50, shows a comparison of mean activity use patterns for the Level II modules. Students transferred information from their folders to optical-scan sheets. Teachers reported ambiguity in the term "finished," with some teachers having students use the term very rigorously and others quite loosely. The total data for activities chosen is probably the more accurate figure for the mean number of activities studied.

Table 12. Statistical data about activity choice patterns for the four Level II modules.

MODULES	ACTIVITIES CHOSEN						ACTIVITIES NOT CHOSEN									TOTAL NUMBER OF ACTIVITIES IN MODULE	ACTIVITY	
	FINISHED			NOT FINISHED			NOT INTERESTING			NO TIME			DID NOT GET TO LOOK AT				TOTAL CHOSEN	TOTAL NOT CHOSEN
	$\bar{X}$	SD	RANGE	$\bar{X}$	SD	RANGE	$\bar{X}$	SD	RANGE	$\bar{X}$	SD	RANGE	$\bar{X}$	SD	RANGE		$\bar{X}$	$\bar{X}$
RULES	11	6	0-34	4	3	0-14	12	8	0-42	12	8	0-42	Question not asked			45	15	24
WHERE DO I FIT	10	7	0-41	4	4	0-19	9	8	0-43	7	7	0-41	12	10	0-44	45	14	16
PERCEPTION	10	5	0-24	4	3	0-20	10	8	0-43	10	8	0-41	11	9	0-40	47	14	20
REPRODUCTION	6	6	0-30	2	2	0-12	4	7	0-39	12	12	0-38	13	11	0-38	36	8	16

Review Team Evaluation

One of the "Review Team Problems" in RULES asked small student groups to give reasons for "not choosing" the activities they didn't use. The results from one of the three problem areas evaluated is shown in Table 13. Responses from the Review Team Evaluation Sheets were categorized into cognitive, attitudinal, and logistic responses. Cognitive responses were those which expressed experience as contrasted with emotion or feelings (which were coded as attitudinal responses). Logistic responses included materials, locations, or the physical arrangements. A single response may have been coded in all three categories. Cognitive and attitudinal responses were the most common given for not choosing an activity. An unexplained "I didn't like it" was most common attitudinal response and "no time" the most common cognitive response.

TABLE 13
Review Teams Reasons for Not Choosing Activities in the Problem Area "Is There A Rule?" (Problem 2)

Category	Example	Number of Responses	Percent of Responses
Cognitive	No time	1433	41.7
Attitudinal	Didn't like it	1337	38.9
Logistic	Couldn't get materials	408	11.9
No reason		257	7.5
TOTAL		3435	100.5

Another Review Team problem (Problem 5) asked small groups to choose the "best" and "worst" activities. Data were analyzed by tallying responses and rank ordering the "best" five and "worst" five activities by problem area and total for the module. Of the fifteen activities ranked "best" and fifteen "worst," four (26.6%) were common to both lists. These four were "Who's Chicken?," "Rules of the Road," "Powder Horn," and "Loyalty, But to What?," indicating the diversity of likes and dislikes within the student groups.

Teachers reported that the categorization tasks in the Review Team problems were too difficult for some students. They also felt that too much writing was required of students and that students complained of the time involved, especially as they wished to begin the next problem area or module with fresh activities from which to choose.

Responses to the Review Team activities raised a large number of issues for that Human Sciences staff that called for future research. For example, if categorizing activities into groups was too difficult for students, then how could students perform the usual categorization tasks they are asked to perform in science classes? One hypothesis was that in the Review Team task a system of categories had to be invented and justified--thinking, the use of cognitive processes and logical operation, was required for the task to be solved. In categorization tasks typical of science curricula the system is given and the problems students are usually asked to do may be "solved" by memorization.

Essay Problems

Students chose the essay problems to which they would respond. In RULES, and PERCEPTION, students wrote responses to five essay problems. In WHERE DO I FIT? they wrote responses to four problems.

In REPRODUCTION, students were provided two "Evaluation Booklets" with seventy-five problems, sixteen essay and fifteen-nine multiple-choice or completion problems. Students were invited to answer as many problems as they could. They were also asked to rate each problem as important or not important and to make one of six reasons for their ratings.

Essay problems were coded to provide the curriculum developers data about levels of comprehension that students could express when they were required to construct responses to problems. For example, Table 14 shows the coding for one essay problem in the RULES module. A generalized response, one in which the student could go beyond the concrete experience of one or more activities and discuss the problem at a higher level of response, was judged to be the most advanced kind of response to many problems. However, it was not anticipated that many students would be able to construct such a response. Over one-third of the students responded in such a manner to problem 1. These students may have mentioned a specific activity and rule, but went on to generalize. Two-thirds of the students gave correct responses, about half of which were specific to a particular activity they had studied. This was the expected response for most seventh-graders.

TABLE 14
Responses to Essay Problem 1: Write an Explanation of a New Rule You Found in "Is There a Rule?"

Response Type	N	%		
Correct general response	54	34.6		
Correct specific responses	47	30.1		
Correct response for pendulum	5	3.2		
Incorrect response	42	26.9	Percent correct	6.19
Incorrect response for pendulum	8	5.1	Percent Incorrect	32.0
TOTAL	156	99.9	TOTAL	99.9

Responses to "Choose Your Problem A," "Write the title of the most important activity you have done so far in WHERE DO I FIT?," "Why was it important to you?," and "What new ideas did you learn from it?" yielded titles of 35 of the 45 activities in the module. The most highly mentioned activities--there were three--were selected by only 12 percent of the students (N=216) who chose the problem. The very widespread selection of "most important" was consistent with the Human Sciences model of early adolescents that finds high variability within a grade level. The data further support the curriculum model indicating that a curriculum must be diverse if it is to be viewed as interesting and important to early adolescents.

Eighty-three percent of the students responding to question "A" were able to state that they learned a general, a specific, or a combina-combination of specific and general fact or idea. Only 17 percent gave no cognitive reason for their choice of "most important" activity.

When given the statement, "Americans are very much alike," 87 percent of the 211 who chose to discuss the problem disagreed with the statement. Sixty-one percent of the students responding gave a satisfactory reason, presenting differences in beliefs, customs, or both; or differences in ethnic, family, or cultural backgrounds. Twenty-six percent gave reasons that included and distinguished between customs and beliefs, a central distinction in the activities relevant to the essay problem.

Student responses to essay problems in WHERE DO I FIT? are summarized in two tables. Nineteen problems required convergent responses, with superior, satisfactory, and unsatisfactory ratings. Table 15 gives the percentage of student responses to these problems.

TABLE 15
Percentage of Students Wesponding in Three Modes--Superior, Satisfactory, Unsatisfactory--to Essay Problems in the WHERE DO I FIT? Module Requiring Convergent Thinking Responses

Problem	Superior	Satisfactory	Unsatisfactory
A01	*	47.0	53.0
A02	*	47.0	53.0
A04	19.9	80.1	0.0
B02	6.8	88.3	4.9
B03	26.0	35.8	38.3
F01	77.5	6.7	15.7
F02A	*	30.3	69.7
F02B	*	18.0	82.0
H01	*	26.8	73.2
H02	*	74.8	25.2
H03	*	59.3	40.6
M01	26.3	21.1	52.7
M02	*	52.6	47.4
N01	26.7	63.4	10.0
O01	0.0	100.0	0.0
P01	70.2	1.6	28.1
Q01	64.3	14.3	21.4
Q02	*	67.9	32.1
R01	86.2	2.1	11.7
R02	*	66.0	34.0

*A superior rating was not appropriate for this problem.

Not all problems called for the distinction between superior and satisfactory, a distinction that placed responses of high generality and accuracy as superior, and a correct, but specific response as satisfactory. Of the nineteen problems, eleven were answered superior or satisfactory by more than 60 percent of the students choosing those problems, a response considered very satisfactory for those application problems.

Table 16 shows the responses to problems requiring divergent thinking. Five of the six problems were responded to satisfactorily by the students choosing them.

TABLE 16
Percentage of Students Responding Satisfactorily and Unsatisfactorily to Essay Problems in the WHERE DO I FIT? Module Requiring Divergent Thinking Responses

Problem	Satisfactory %	Unsatisfactory %
A07	60.0	30.9
D01	71.4	28.6
D02	75.5	24.5
D03	42.9	57.1
K01	76.2	23.8
K02	76.2	23.8

The quality of responses to these essay problems was used as a major criterion for revising both activities and essay problems. In some instances, the problems were judged too difficult for seventh-graders. In all cases, activities were revised to provide more structure, and to clarify and distinguish the key ideas where student responses indicated that distinctions were not clear.

One essay problem in PERCEPTION, chosen by 109 students, asked for an explanation of how they felt about the activities they had selected to study. Responses were categorized into cognitive, attitudinal, and logistic groups. Eighty-seven percent of the students responding included attitudinal remarks within their responses. About fifty-seven percent of the responses were positive, 14 percent negative, and 15 percent were mixed.

There were sixteen essay problems in the seventy-five-problem evaluation booklets used in REPRODUCTION. Responses to those essay problems were rated in the same manner as were responses to the essay problems in WHERE DO I FIT? Table 17, page 55, shows the percentages of students with superior, satisfactory, and unsatisfactory responses to these problems. Sixty-two percent of the responses were judged superior or satisfactory. This response level was judged satisfactory, given the short time that most students had to study REPRODUCTION. Data analysis indicated that many students responded to test items that assessed activities they did not do. This indicated that students felt they knew more about the subject matter of the module than they actually did, based on their responses to the essay problems. Responses to the essay problems gave valuable information for the curriculum developers as they revised the activities in REPRODUCTION.

Objectively-Scored Problems

Multiple-choice and completion problems were given only in the REPRODUCTION module. They were designed to produce means near

TABLE 17
Scores of Students Who Chose to Respond to the Essay Problems, REPRODUCTION Evaluation Booklets 1 and 2

Problem	Number of Students	Superior Sequence N	Superior Sequence %	Satisfactory Sequence N	Satisfactory Sequence %	Unsatisfactory N	Unsatisfactory %
10	117	8	6.8	15	12.9	94	80.3
17A	206	119	57.8	12	11.1	64	31.2
17B	217	*	*	90	87.6	27	12.4
20	83	30	36.1	18	21.7	35	18.4
21	136	21	15.4	0	0.0	15	84.6
22	136	21	15.4	0	0.0	15	84.6
23	133	14	10.5	9	6.8	10	82.7
24	119	1	.8	69	58.0	49	41.2
26	147	23	15.6	55	37.4	69	46.9
37	163	68	41.7	58	35.6	37	22.7
38	63	*	*	42	66.7	21	33.3
45	128	17	13.3	67	52.3	44	34.4
55	226	167	73.9	54	13.9	5	2.2
57	155	64	41.3	62	40.0	29	18.7
60	97	13	13.4	39	40.2	45	46.4
63	139	8	5.8	97	69.7	34	24.5
64	28	*	*	27	96.4	1	3.6
65	40	*	*	26	65.0	14	35.0
66	40	*	*	23	57.5	17	42.5
67	27	*	*	26	92.9	1	3.6
68	29	*	*	25	86.2	4	13.8
69	30	*	*	21	70.0	9	30.0
74	92	16	17.4	44	47.8	32	34.8
75	155	10	6.5	86	55.5	59	38.1
All Problems	2706	600	22.2	1076	39.8	1030	38.1

*A superior rating was not appropriate for this problem.

50 percent to give the best psychometric data about the test. Students were asked to mark the best response to the items. Many items were constructed to assess levels of generality that students could achieve in order to provide maximum information for activity revision.

For example, item 18 in Evaluation Booklet 1 provided a range of more- and less-inclusive choices regarding the term "sexual reproduction." Table 18, page 56, presents the item and item statistics for this problem. A more complete analysis of this problem is provided in EP 7906-40 where the item was analyzed as part of a subscale of three problems dealing with sexual reproduction and fertilization. The item statistics refer to this subscale.

Three-fourths of the students in test classes for REPRODUCTION ventured a choice on this problem. This large choice pattern indicated wide enough familiarity with the term for most students to risk a response.

About one-fourth of the students selected (c), the most general response that was valid. Response (d), restricting the applicability of the term to vertebrates, was the most preferred alternative. Only 13

TABLE 18
Responses to Choices for Problem 18

Item Stem and Response Choices	Item Statistics		
	N	%	r_{bis}
The term "sexual reproduction" can be used correctly for:			
a. all kinds of animals, but not plants	42	13.3	-.16
b. only human beings	30	9.5	-.09
c. most plants and animals	82	26.0	.49
d. only fish, birds, amphibians, reptiles, and mammals	67	21.3	.11
e. all kinds of plants and people,	16	5.1	-.13
f. not chosen	78	24.8	-.41
TOTALS	315	100.0	-

percent of the students excluded plants from inclusion in "sexual reproduction." The negative biserial correlations for all but two distractors indicates inverse relationships of selection of those responses with success on the subtest. The biserial correlation for response (d) shows a very weak relation between this response and the subtest score, as was expected.

Fifty-nine multiple-choice and completion problems were used in Evaluation Booklets 1 and 2 in REPRODUCTION. Mean achievement in these items was 45.7 percent, based on choice of the best response. Again, analysis of activity selection by problem choice indicated that many students responded to problems when they had not studied the activity the problem was designed to assess (see EP 7906-40). Achievement was within the expected range, especially if scores were calculated by crediting the second best choice, usually a correct but not best choice, as a reasonable response.

Student Ratings of Essay and Objectively-Scored Problems

Students were asked to rate the objectively-scored problems important or unimportant and to check the most appropriate reason, using the rating scales shown in Table 19, page 57.

TABLE 19
Response Choices for Student Ratings of Evaluation Problems in REPRODUCTION

Importance and Reason	Importance and Reason Answer Guide
I	Important
NI	Not Important
a.	Everyone my age should know it
b.	I'll need to know it later
c.	I want to learn as much as I can
d.	It's too technical (has special use only)
e.	I don't see any reason for knowing it
f.	None of these

Reasons "a" to "c" were considered as indicating support for the choice "important" and "d" to "f" as reasons to support the choice, "unimportant".

Table 20, page 57, shows the subject matter of the eleven problems rated important by 70 percent or more of the students responding. Note that all of these problems relate to the human organism, even though reproduction in plants, animals, and humans were the objects of study in the module.

TABLE 20
Subject Matter of Problems to Solve Rated Important by 70 Percent or More Students, REPRODUCTION Evaluation Booklets 1 and 2

Item No.	Objective Problems	Item No.	Essay Problems
8	Identical twins	10	Major events of baby's birth
9	Identical twins	26	Breast, bottle feeding of baby
10	Site of sperm production	37	Care of newborn baby
36	Sex determination of baby	38	Birth defects
47	Human reproduction	55	Needs of children
62	Needs of children		

Parts and functions of seeds and flowers had the lowest ratings of importance along with male and female characteristics of animals. However, there was not an exclusive relation between "important" and "not important" using the criteria described above, as the highest rating of "not important" was 61 percent of the students. The items highly rated were consistently those requiring memorized responses.

These student ratings were used to help the revision writers identify problems rated "not important," with the inference that the subject matter or the evaluation item in related activities item needed to be revised to assist students in learning the importance of the content of activities.

Attitudes and Skills

A third evaluation booklet was used in the REPRODUCTION module. This booklet had items in a section titled, "Skills I Developed," and ten items in a section titled, "Feelings." The "Feelings" problems asked students to mark one of four choices, from strongly agree to strongly disagree. Five concepts were planned for the ten items. The item statements, grouped by the conceptual description of the item groups, are shown in Table 21, page 58. The four statements about evaluation were rated in the desired direction, with the means clustering around the "agree" value.

"Working style" yielded scores showing that the central tendency was that students did not work alone on activities, and that they generally worked with the same individual or group. Responses to "Activity Choices" indicated that more time would have been valuable and that the module could have been used longer than it was.

Single items assessed student attitudes toward what they were learning and development in self-direction, both of which yielded positive means. A series of skills development questions in multiple-choice format was provided to determine the degree of science students were having primarily with the line material activities in REPRODUCTION. Some activities, like the one using Medaka fish, were chosen by only 11 of the more than 300 students in classes testing REPRODUCTION. Securing the fish was difficult in some communities and time did not permit the use of mail order forms that were provided with the activity.

Difficulty getting fertile chick eggs was also experienced in some test schools. More than half of the students who started the activity didn't get to finish it (data not shown).

Responses to the skills development problems showed that, in general, students were able to do the tasks required to conduct the activities. However, keeping plants alive and doing the pollination to produce seeds were both accomplished by less than half of the students reporting.

Teacher Evaluation of the Level II Program

This section summarizes the responses to the questionnaires and provides recommendations about Level II teacher and student materials. All seventh-grade teachers using the Level II modules in 1974-1975 were

TABLE 21

Mean Scores of Students' "Feelings" about the REPRODUCTION Module, from a Likert Scale with Four Choices: Strongly Agree (=1), Agree (=2), Disagree (=3), and Strongly Disagree (=4). ("Desired Mean" is the score indicating a positive attitude.)

		N	Mean	S.D.	Desired Mean
	<u>Attitudes toward evaluation</u>				
89	1. The questions I marked important in Evaluation Booklets 1 and 2 are good measures of what I have learned in this module.	271	2.06	.74	2.4
90	2. All of the questions together are a good measure of what can be learned from the activities in REPRODUCTION.	271	1.99	.78	2.4
95	3. This evaluation activity has helped show me how much I have learned from REPRODUCTION.	270	2.13	.92	2.4
96	4. The evaluation activities in REPRODUCTION are more helpful than the evaluation activities in other modules this year.	265	2.12	.93	2.4
	<u>Choice of working style</u>				
92	1. I worked by myself on most of the activities I completed in this module.	271	3.12	.91	open
91	2. I worked with the same kids on most of the activities I completed in this module.	271	1.99	.88	open
	<u>Importance of learning</u>				
93	1. I feel that I learned important facts, ideas, or skills from the activities I completed.	269	1.95	.86	2.4
	<u>Development of Self-Direction</u>				
94	1. I am learning to work more independently now compared to this time last year. My teacher doesn't have to check up on me very often.	262	2.00	.88	2.4
	<u>Activity Choices</u>				
97	1. It was easy for me to find activities I wanted to do in this module.	245	2.16	.90	2.4
98	2. I would have completed more activities in this module if I had time.	248	1.53	.78	2.4

invited to participate in a conference to evaluate the five Level II modules.

The purposes of the conference were:

1. to identify successes and problems in using Level II modules,
2. to evaluate the materials preovided for teachers and students, and
3. to provide information to the project staff to be used in revising Level II materials and for developing Level III.

Nine teachers and the classroom observer for Level II participated in the conference.[1] A series of nine questionnaires for individual and small-group responses were utilized to gather data.

Successes and Problems

The conference opened with a request for participants to identify and list their three major successes and three major problems with HSP during 1974-1975. Successes were provided in two forms, general statements and specific activities. Generally, students seemed more highly motivated in HSP than in other courses. They viewed HSP positively, the course seemed attractive to non-HSP students, students learned new skills, and "problem" students had great success. In general, short duration activities seemed more successful than long duration activities. Specific activities, such as "Class Newspaper," "Hear It from a Judge," and "Help-a-Person" were mentioned as having great impact on students.

One major problem revolved around arranging out-of-class activities. Parental concern and school regulations made such activities difficult to do in some schools. A second problem centered around evaluation. Students had difficulty understanding the relationship between the activities they studied and the evaluation problems. Teachers did not understand how to use the evaluation products. Motivation, getting students involved, and student disruptive behavior were also identified as problems.

Evaluation of Teachers Guides

Each participant completed a questionnaire in which he or she rank-ordered the contents of each Level II guide. Two rank orderings were requested, "most used" by the teacher during the module and "most important" for teachers new to the program.

The activity guides ranked first in both the "most used" and "most important" categories for three of the four guides, and ranked second for one guide. Module introductions generally ranked third or fourth. The packing list generally ranked lowest. Questioning techniques in

---

[1]Closing dates of schools made it impossible for five teachers to attend. They were asked to complete all of the individual questionnaires used at the conference.

PERCEPTION ranked second. Problem area introductions generally ranked near the center of ranking hierarchy.

From comments made during discussion and in response to other questions on this questionnaire, the concensus of the group was that all materials in Level II guides were important. Rather than delete any materials, teachers recommended changes or additions.

Guide Organization

Teachers favored guides with the following characteristics:

1. loose leaf
2. teachers guides for an activity next to the activity
3. everything together for each problem area
4. a special considerations section in the front
5. a more compact list of materials included and not included
6. specific teaching ideas, such as techniques in questioning in PERCEPTION
7. charts giving a quick overview of activities for each problem area (e.g. REPRODUCTION, Teachers Guide, p. 41).
8. questions in lists, not embedded in paragraphs
9. PERCEPTION is a good example of the best guide

Add to Teachers Guides

1. variety of suggestions for measuring student gains
2. supplementary film and book lists
3. supplementary activity suggestions
4. for student use only, a self checklist at the end of each activity (for those who need more direction about "what did I learn?")
5. more suggestions for teachers to help students become problem solvers and classroom managers
   a. setting up the module
   b. getting started
   c. solving management problems
   d. communicating with others
6. suggestions for motivating students
7. easy reference overview charts
8. reinforce ideas for With Learning in Mind
9. evaluation should focus on student progress
   a. use short answer questions
   b. possibly a pretest/posttest for each activity

Activity Organization

Three groups of three teachers each met to discuss and complete a questionnaire relating to activities, problem areas, and module organization and design. The first set of questions was concerned with activity format and organization. Short introductions and art work were important in attracting students to activities. More films, film loops, and tapes would be desirable, either as part of the module or as supplementary materials to be purchased separately from the modules.

Questions embedded in paragraphs were generally ignored. They should be set apart and numbered. Some questions were too difficult for students.

Step-by-step directions were most helpful. Students needed more specific direction in pulling data together and using it.

Three questions were asked regarding duration of activities. There was preference for short activities--those that can be completed in from 1 to 3 class periods. However, variety of activity length is needed in every module. Long-duration activities are:

1. needed for a small number of students
2. difficult to store
3. subject to damage by other students
4. accompanied by loss of interest by many students

Students are reluctant to do a second activity while waiting for seeds to sprout, etc. Long-duration activities need specific suggestions about other activities to do while waiting. Alternatively, a cluster of long-duration and related short-duration activities might be one kind of choice provided.

Participants were asked to provide their reactions, other school staff reactions, and student reactions to activities requiring students to leave the classroom and school. Legal restrictions and red tape were the major problems for teachers, but such activities were viewed positively by trial teachers. Mostly favorable comments and support from other staff members were reported. Student reactions were generally favorable, with reports that students policed each other when out of class, but that some students took advantage of the situation.

Problem Area Organization

Teachers reported that the problem area themes were generally too sophisticated for students. They found the short, concise statements very helpful for themselves. The matrix analysis in REPRODUCTION was useful for both students and teachers. A recommendation was made that matrices should be provided on a form for student use.

Specific comments were asked about strong and weak problem areas in each module. In RULES, "How Do Rules Change?" was considered weak. In WHERE DO I FIT?, "Where Do I Fit in the Future?" and "Where Do I Fit as an Organism?" were both questioned. PERCEPTION was considered to have effective problem organization structure. REPRODUCTION could not be evaluated because of little time in use.

Module Organization

WHERE DO I FIT? was the only module not mentioned by anyone in response to the question "Which modules were most effective in developing the module theme?" REPRODUCTION and RULES wre cited as most effective by two of three groups. WHERE DO I FIT? was judged least effective in making sense to students as a module title.

Module packaging was considered to be much better for Level II modules than for Level I. The major complaint was that most modules were incomplete (materials missing) upon arrival. Where materials were missing, students were disappointed and teachers frustrated.

In RULES, the integrative activity "Selur Island" was not considered effective. In WHERE DO I FIT?, the integrative activity "People" was not considered effective, but, "The Where Do I Fit Library" was found to be valuable and was used. It is too early to judge the three integrative activities in REPRODUCTION, but one teacher reported that students were starting to use them.

Inclusion of integrative activities was recommended for Level III, at lease one per module. Games were not viewed positively and small group integrative activities were suggested.

Evaluation of REPRODUCTION

Participants were asked to list activities which were removed from REPRODUCTION. Seventeen classes are using REPRODUCTION. Thirteen classes were reported in the following section.

"Putting It All Together" and "Birth of a Baby" were removed in two classes. "Nursing" and "Am I Regular" were removed in one class. Removal was not a school matter, as the practice varied with teachers and students (and their parents). Need was expressed for an activity on venereal disease and, from one of three groups, for an activity on birth control. Participants were asked to cite activities whose use was restricted and how those were used. At three (of five) schools, parental permission slips were required to do REPRODUCTION activities. Within classes, some students were doing activities that others were not permitted to do. This variability was very successful from the viewpoint of teachers, students, and parents.

Evaluation of Pedagogy

Each participant completed a questionnaire concerned with three aspects of pedagogy: student choice, timing, and student involvement. Participants met in groups of three to discuss additions to program materials that would assist teachers in terminating problem areas and modules and in helping a larger teacher population to be comfortable with the materials. The first question asks for teacher and student reactions to activity choice.

Teacher Reactions: Participants were unanimous in support of choice as an essential part of HSP. Enthusiasm, interest, meeting the diversity of student competencies, and the importance to students of learning to choose were some reasons stated in support of choice.

Teachers found that student competence in decision-making varied a great deal. Several teachers provided structures beyond those suggested in the Teachers' Guides, such as weekly plans prepared by each student or carefully structured introductions to each problem area. One teacher commented on the growth of student competence in decision-making during the year.

Student Reactions. For the most part, students were reported to respond positively and effectively to choice of activities. Some students had difficulty choosing, and one teacher reported that students became lazy and chose nothing. One solution for helping students with limited competence in decision-making was to have each student prepare a weekly plan of what they wanted to accomplish. Plans were considered to be plans, not blueprints, with the option of changing plans during the week. This option was a very important aspect of planning.

Evaluation of Module Evaluation Materials

A questionnaire asking for three comments about each component of each module was completed by pairs of seventh-grade test teachers. The three comments were: "your reactions," "students' reactions," and "suggestions for improvement." A brief summary regarding each evaluation item is provided below.

Folders. Folders with pockets were used and were useful for both students and teachers. Those folders without pockets were not as useful. Folders could be improved by replacing the comment section with a question for students to answer. This could be an activity-specific question, a question for the problem area, or a generic question. It would provide a statement that the teacher could use at a glance to see one main idea the student learned from the activity.

My Activity Record. Students preferred the Optical Scan format. Teachers would like to have a simple summary from data on this sheet. Students had trouble defining completion of an activity. There needs to be space for student-developed activities and additions to activities. A student profile for all modules for the year was suggested.

Choose Your Problem. Teachers did not understand or use the output. They felt students picked the easiest or shortest problems. Illustrated and simplified items (see PERCEPTION compared with RULES) were more acceptable, but students still found the evaluation activities distasteful. Teachers need help in using the varied student responses. There was some support for multiple-choice questions.

Review Teams (RULES only). Students could not work together in large student-led groups. The problems were too complex for them to solve. This was judged to be an unsuitable activity for seventh graders.

Student Evaluation Booklets. The booklets were improved from module to module. Art work, simplified directions, color, and simple questions all contributed to more effective booklets. Evaluation materials for REPRODUCTION had not been used and were not evaluated at the conference.

Teachers provided brief accounts of how grades were determined in their classes. Several ideas to encourage and promote self-evaluation skills were suggested.

Evaluation of "With Learning in Mind"

Participants rated each item in the contents of With Learning in Mind on frequency of use and usefulness. This publication was a first draft of a Human Sciences Program guide for teachers. With few exceptions, the contents were used once and were rated useful. Some suggestions were provided for revising particular sections of the publication. Judgments were divided on whether the following four items were useful or should be omitted: "What Does Piaget Say to the Teacher?" "For Parents," "Modes of Learning," and "Guidelines for the uses of Animals in School Science Behavior Projects."

There was unanimous agreement on the need for an over-all program teachers guide. The major thrust of suggestions for helping teachers was to provide an interesting, attractive, stimulating, well-illustrated, "fun" way to learn about the program.

Science Supplies in HSP Trial Classrooms

A list of common science materials was provided to seventh-grade teachers at the Level II Evaluation Conference. This materials list included such items as beakers, test tubes, and common chemicals. Data concerning the list were obtained from schools 4, 7, 8, and 9.

School 7 had no equipment with the exception of one meter stick per classroom. All of the other schools had all of the apparatus, equipment, and glassware. Most chemicals listed were available, but some respondents couldn't remember if particular barium or potassium compounds were available. Balances and microscopes were present in numbers ranging from 7 to 18.

Summary

Four Level II modules--RULES, WHERE DO I FIT?, PERCEPTION, and REPRODUCTION--were field tested in 1974-75 in the same seven schools that tested Level I. Thirteen teachers, 10 of whom were new to teaching Human Sciences, and 490 students, 62 percent of whom had participated in the Level I field test, were involved in testing Level II.

Three modules were adequately tested, but the fourth, REPRODUCTION, was tested for only three weeks (or less) and with only four of the seven schools participating.

Activity choice showed the wide range of use, likes, and dislikes that were found in Level I. The longer module use time resulted in higher use of activities within modules. As with Level I, the mean use was about one-third of the number of activities provided in a module. Animal and plant activities in these modules were not as highly chosen as people-oriented activities. Activities involving interviews with adults were highly chosen. In classroom visitations, students affirmed this interest, indicating that they had very limited interactions with adults, except for parents and/or teachers. Interview activities gave both a purpose and a subject to talk about, providing the means by which a strong desire to talk with a wider range of adults was achieved.

Self-evaluation, supported by students' preparing portfolios of best work was advocated and supported by records folders, work and date sheets, and by essay test items that were written on NCR® paper so that each complete student response was available for inclusion in the portfolio of "best work" and for the Human Sciences formative evaluation. A section on evaluation was included in each module Teachers' Guide. An evaluation booklet, with explanations of what self-evaluation was and how the student was to be involved in it, were also provided with each module. Several teachers developed materials for students to summarize their work, to grade themselves on this work, and to justify their grade. These teachers found that the student grades were the same as their grades, except for a small percentage of students who consistently over- or under-graded their work from the teacher's perspective.

Despite these successes, both students and teachers were dissatisfied with the evaluation materials and procedures. There was support, but not unanimous support, for objectively scorable problems to be included with the Level III materials.

Assessment of achievement, as measured by essay and objectively-scored problems, was a part of each Level II module. Essay problems for each module yielded satisfactory responses by about 60 percent of the students responding, and superior responses by about 30 percent.

Objectively-scored problems were first introduced in the REPRODUCTION module. Mean scores for the objective problems were 45 percent for the twenty-eight problems in Evaluation Booklet 1; 44 percent for the thirty-one problems in Evaluation Booklet 2. Reliabilities for the objectively-scored problems (Hoyt analysis of variance) were 0.79 and 0.82 respectively, with ranges from 0 to 75 percent and 0 to 71 percent respectively. The short time for study and the invitation to students to answer any question they thought they could, without penalty, yielded these results. The critical distinctions among the choices on the objectively scored tests provided valuable information for activity revision. For example, if the objective in the revised materials would be to enable more students to generalize the meaning of "sexual reproduction," all relevant activities could include the term and provide examples of inclusion and exclusion.

Essay problems were difficult for teachers to evaluate and grade. Keys for scoring both objective and essay problems were provided in REPRODUCTION. Additionally, the first parts of a teachers guide, <u>Development of Self-Evaluation for the Human Sciences Program</u>, (SP 7601-46), were drafted and distributed late in the school year.

The teachers guides were considered an improvement over those provided in Level I. Guides for Level III were to incorporate additional teacher aids, such as "Questioning", in the PERCEPTION Teachers Guide, which was commended as an example of what was needed.

With ten new teachers of thirteen in the formative evaluation in 1974-1975 assimilated into the field testing program, the orientation sessions and teacher materials were shown to have capitalized on the results of early evaluation studies. The Level II modules were considered to be improved over Level I. The evaluation materials provided in REPRODUCTION were judged the best in Level II, and to serve as the point of departure for new evaluation efforts for the Level III materials.

CHAPTER 6

THE LEVEL III FIELD TEST

Four modules were tested at seven school sites with ten teachers and approximately 335 eighth-grade students during the 1975-1976 academic year. All school sites, most teachers, and most students had previous experience with Human Sciences at Levels I and II (Table 22). The original evaluation plan--to begin with three classes at the beginning of field testing in order to have at least one class at each school--was met, with three schools having more than one class (schools 1, 4, and 7).

TABLE 22
Numbers of Experienced and Inexperienced Teachers and Students in Seven Field Test Schools, Fall 1975

School Number	Teachers			Students		
	Experienced	Inexperienced	Total	Experienced	New	Total
[illegible]	1		1	36	22	58
[illegible]	1	1	2	78	2	80
5		1	1	29	0	29
6	1		1	23	5	28
7	2	1	3	74	11	85
8	1		1	23	0	23
9	1		1	20	12	32
ALL	7	3	10	283	52	335

All but one teacher were provided with a two-and-one-half day orientation session in Boulder, Colorado, with the Level III materials and the evaluation materials and procedures to be used in the Level III field test program. Four modules--CHANGE, FEELING FIT, INVENTION, and SURROUNDINGS--were to be tested during the year. The modules were produced, distributed, and tested in the order presented above. Field testing began in the first week in October, 1975 and was completed when each school closed, May and June, 1976.

Evaluation Plan

The major goal for evaluation of Level III was the development and testing of simple, marketable, self-evaluation materials. These evaluation materials were developed to meet several critical criteria:

1. Flexibility in accommodating unequal numbers of activities completed by students.
2. Flexibility in accommodating any pattern of activity choice.
3. Providing for success for the full range of variability in student cognitive development.
4. Providing simple scoring procedures that yield numerical data.
5. Providing data that could be used for grading individual students.
6. Providing data that could be used in evaluating the effectiveness of the curriculum materials.
7. Maintaining consistency with the goals of the Human Sciences Program.

Recognition, knowledge, and intellectual skills were evaluated through objectively-scored problems. Constructed knowledge and intellectual skills were assessed through essay problems. Laboratory skills, student attitudes, and affective dimensions of learning were assessed through checklists, self-rating scales, and similar devices.

Students chose the problems they wanted to answer. There were no requirements as to how many problems a student should choose. Students built portfolios of their best work, including the evaluation documents they produced as they studied activities in the module. A _Mapping My Progress in Level III, Human Sciences_ booklet provided for graphs of activity and objectively-scored problem data for the Level III modules. It also provided for student interpretations of their achievement at the end of each module. An addition to _Mapping My Progress in Human Sciences_ was provided with FEELING FIT and subsequent modules to map and summarize work habits and skills development in Level III. The student evaluation activities for the Level III modules are briefly presented in Table 23, page 71.

SURROUNDINGS was a very different module in terms of its structure. It became obvious, as field testing proceeded, that there would be little time left in the school year for the field test of SURROUNDINGS. Discussions with teachers showed that they and their students felt that too much time was being required for evaluation activities. All of these forces and ideas led to a SURROUNDINGS module with only two problem areas and twenty activities.

The evaluation design was changed to provide a specific set of evaluation problems for each activity. Students would not have test booklets for any specific evaluation period. When they completed an activity they would get the "Problems to Solve" sheet for that activity from the teacher, complete it, and turn it in for grading.

The use of modules in test classes also varied considerably. CHANGE was started in test classes the first or second week in October. The module was used from twelve to fifteen weeks. FEELING FIT was used

TABLE 23
Evaluation Activities for Level III

Module(s)	Number of Evaluation Periods/ Module	Student Evaluation Activities
CHANGE FEELING FIT INVENTION	Three each module	Module RECORD Record activities started completed daily Problems to Solve Evaluation Periods 1 to [illegible] 1 Objective Problems Fifteen each period 45 total per module 2 Choose Your Problem Four to six each period Answer any one problem, each evaluation period total three per module 3 Self-Rating Problems Seven Work Habits problems each period answer all Six to eight Skills Development problems each period answer all Data Transfer 1 Activity records from Change Record to "My Activity Record" optical scan form each evaluation period 2 Objective Problem responses from Problems to Solve booklet to optical scan forms each evaluation period Mapping My Progress All modules 1 Calculate percentage of activities chosen each period 2 Graph percentage of activities chosen each period 3 Calculate percentage of problems chosen each period 4 Graph percentage of problems chosen each period
		Mapping My Progress beginning with FEELING FIT 5 Interpret meaning of changes in activity and objective problems graph lines each evaluation period 6 Calculate average Work Habits score and Skills Development score 7 Plot average scores for Work Habits change in Work Habits and Skills Development on graphs added to Mapping My Progress 8 Answer problems about Work Habits and Skills Development
INVENTION only		Invention Questionnaire Fifteen agree-disagree statements about the module
SURROUNDINGS	Two	SURROUNDINGS Folder Record activities started completed, daily Problems to Solve An activity specific quiz to be completed after each activity studied was completed A [illegible] of objectively-scored and essay problems for each activity All problems to be answered if activity was studied Mapping My Progress each Evaluation Period 1 Calculate and graph percentage of activities chosen 2 Calculate and graph the total number of test items answered 3 Integrate activity and test item graphs Self-Rating Problems last Evaluation Period Respond to all Work Habits and Skills Development problems Activity Evaluation Form An activity specific form with eight Likert-type items and two essay problems to give students the opportunity to evaluate each activity studied Mapping My Progress last Evaluation Period Mark graph and interpret Work Habits and Skills Development data Data Transfer end of module Transfer activity records from SURROUNDINGS Folder to "My Activity Record" optical scan forms

from eight to nine weeks, INVENTION from five to nine weeks, and SURROUNDINGS from zero to three weeks.

These differences in modules made it more convenient to treat each module as a unit, and describe the evaluation results for that module. Commonalities and differences in the results will be discussed in the "Summary."

The objectively-scored problems developed for use in Level III were constructed to determine "levels of discrimination" that students could make among plausible distractors. No problems were designed for mastery, but rather to provide normative data and to have mean scores near the 50-percent level. This was done to provide variance in responses.

Two end-of-program evaluation instruments were prepared for use in the second week of May, 1976, in all test schools. "How Is Your Logic?," 1976 edition, a revision of the 1974 version, was administered to determine logical competencies of students for selected concrete and formal operational tasks. The "Science Questionnaire" was administered to all Human Sciences students and to an equal number of eighth-graders in each school who had not been in Human Sciences test classes. The "Science Questionnaire" included a request for students to rank order their eighth-grade course in order of "best course" to "next best course," etc.

A conference of the Level III field test teachers was held in Boulder, Colorado in June and July, 1976. Eight of the ten test teachers participated in the conference. Written evaluations and taped group discussions of modules and activities were obtained and analyzed as part of the formative evaluation. In addition, some activities were reviewed by parents and academic scholars in public and content review conferences of the Human Sciences Program held in April, 1977.

Data Analysis Procedures

The optical-scan sheets used for activity records, responses to objectively-scored problems, and self-rating problems were converted to data cards. Listings and frequencies (SPSS) were used for detecting coding errors.

An important evaluation question considered for Level III was the question: "Did finishing an activity make a difference?" This question may be expressed in other ways, such as: "Are the specific evaluation items chosen for answering dependent on which activities are started or completed?" "Are the specific items chosen and answered <u>correctly</u> dependent on which activities are started or completed?" "Which items are chosen for answering most often and why?" "What is the relationship between total number of activities started and/or completed and total number of items chosen for answering and/or items answered correctly?" These questions were answered for some activities in each module by analysis of responses to activity-related problems. Cross-tabulations and correlation analysis of data are shown below in Table 24. These data were completed for specific activity objectively-scored problem pairs.

TABLE 24

Design for Testing the Hypothesis, "Success on Problem Y is Dependent on Completing Activity X (numbers are call numbers for reference in the test)

Problem or Problems	Activity or Activities			
	Completed	Not Chosen	Data Missing	Totals
Correct	1	2	-	3
Incorrect	3	4	-	7
Not Chosen	5	6	-	11
Missing Data	-	-	-	-
Totals	9	12	-	21

This analysis made possible testing the hypothesis, "Success on problem Y is dependent upon completing activity X." A higher proportion of responses in cell 1 (correct, completed) than cell 2 (correct, not chosen) would support this hypothesis. Higher proportions of responses in cell 4 (incorrect, not chosen) than cell 3 (incorrect, completed) and in cell 6 (not chosen, not chosen) than 5 (not chosen, completed) would provide additional support for the hypothesis.

If proportions in cells 1 and 2 were equal and high, the hypothesis that the problem assessed common knowledge that is activity independent would be tenable. Further confirmation of this hypothesis would be indicated by equivalent proportions in cells 3 and 4, and in cells 5 and 6. Chi-square tests of significance of differences between cells were applied to crosstabular analyses. One-tailed tests of significance were applied to the correlation analyses.

A second important evaluation question for Level III was the range of student achievement. Answering the question, "What did students accomplish?" utilizes data from Problems to Solve, both objectively-scored, and essay problems, and the activity selection data. The thrust of these analyses will be to display the range of student outcomes as determined by the module evaluation instruments. Other statistical analyses were employed as the need for them developed in examining the data obtained.

Results of the Level III Field Test

Detailed analysis of the evaluation data for each of the 148 activities that were developed and tested in the four Level III modules to be discussed here were provided in activity folders for use in revision. A complete written analysis of the module CHANGE (EP 7706-16) was prepared as an example of the kind of module evaluation report initally planned for each module. Staff time was not available to carry out this task for each Level III module. The decision was made that evaluation data in folders was adequate for revision and that written reports would not be as valuable as devoting staff time to the actual revision of modules and activities.

The following parts of this chapter present a brief summary of the findings from the Level III field test. CHANGE and FEELING FIT were similar in structure, with three problem areas each, and with forty-six and fifty-three activities, respectively. INVENTION, as was previously discussed, requires a different kind of analysis, as does SURROUNDINGS. These two modules will be described separately.

Evaluation of CHANGE

The CHANGE module, with forty-six activities and three "skills" booklets, was introduced into field test classes during the first or second week in October, 1974. This summary will present brief accounts of activity use patterns, what students accomplished, and whether the activities made a difference in test achievement.

What Students Did: Activity Use Patterns

Data for activity selections were recorded by students in their CHANGE Record. At the end of each problem area one or more evaluation periods were scheduled. Students were asked to indicate if they wished to be accountable for each activity according to the criteria shown in Table 25. In the subsequent analyses, students were said to have "done" an activity if they responded to a "Yes" category and "not to have done" an activity if they responded to one of the "No" categories (exceptions will be specified).

TABLE 25
Student Indication of Their Experience with Each Activity in CHANGE

if	YES	Mark 1 if you completed at least one part of the activity. Mark 2 if you completed all parts of the activity. Mark 3 if you learned by observing another's activity.
if	NO	Mark 4 if you haven't looked at it. Mark 5 if you haven't had time to do it. Mark 6 if you haven't wanted to do it.

Mean number of activities "done" for the total group, for boys, and girls are shown in Table 26, page 75. A mean of 14.9 represents 32.4 percent of the forty-six activities. Although there was a significant difference between the means for boys and girls ($p=.03$), determined by one-way analysis of variance, the mean number of activities done by boys and girls differed by only one activity. It is noteworthy that the variation in number of activities completed by boys was greater than that of the girls.

TABLE 26
Mean Number of Activities Done in All Problem Areas In CHANGE

Students	Number of Students	Number of Activities "Done"	
		Mean	Standard Deviation
Boys	165	14.6[1]	7.08
Girls	175	15.3[1]	5.24
Total	340	14.9	6.25

[1]Significant difference p=.03

The content emphasis of the forty-six activities in CHANGE was used to construct Table 27. Categorizing the content emphasis in CHANGE activities was somewhat arbitrary since many activities included content across the disciplines. Within this constraint, Table 27 shows the proportion of activities having a biological, physical, or social sciences emphasis, and the choice patterns of students. One-way analysis of variance was computed to determine if the number of activities chosen by boys and girls was different. Boys chose activities with biological and physical science more than girls. Girls chose activities with social sciences emphases more than did boys (p=.003). From these data it appears that the physical sciences activities had more appeal to boys, the social sciences to girls, and the biological sciences activities, though they were selected by boys to a greater extent, were more evenly chosen by boys and girls.

TABLE 27
Content Emphasis in CHANGE as a Function of Activities Chosen by Boys and Girls

Content Emphasis	Percent of Activities in Module	Percent of Activities Chosen		
		Boys	Girls	All Students
Biological Science	40.8	41.4*	40.1*	41.0
Physical Science	30.6	31.1**	29.5**	30.2
Social Science	28.6	25.1*	29.5***	27.3
Totals	100.0	97.6	100.0	98.6

*Significant difference p=.04
**Significant difference p=.03
***Significant difference p=.00

Three skills booklets, "Working with Fruit Flies," "Making and Using Graphs," and "Seeing Small Things," were included with the CHANGE module. These booklets were designed to be used most heavily in the problem area, Change in Non-human Organisms.

For the first time in a Human Sciences module, four integrative activities were included in CHANGE. Unlike other activities in CHANGE, these were teacher directed. Use data were obtained from six of the ten test teachers who attended the Level III teacher feedback conference. All six teachers used the integrative activity, "Earthwatch," but none used the activity, "What Does it Take?"

What Did Students Accomplish?

Data for this section were obtained from Objective Problems and Choose Your Problems (essay questions) from the Problems to Solve booklets.

There were forty-five objectively-scored problems, fifteen for each problem area, administered at three different time-periods during the CHANGE module. Students could elect not to answer any problem by selecting the response "I do not choose this problem." There was no identification to enable students to relate objective problems to an activity.

Student achievement, as reflected by mean scores on the Objective Problems tests, was 53.6 percent correct (Table 28). Mean scores and test reliabilities for each problem area are also shown in Table 28. The tests had satisfactory reliabilities. Each problem on each test was referenced to one or more activities. Since these problems were untested, as were the activities, student achievement is judged to be satisfactory, but lower than will be desirable for use by students and teachers in the revised materials.

TABLE 28
Mean Number of Objectively Scored Problems Chosen, Problems Correct, and Test Reliability Scores for Each Problem Area in CHANGE

Problem Areas	No. of Problems	Problems Chosen				Problems Correct				Reliabilities[3]
		Mean	%[1]	Range	S.D.	Mean	%[2]	Range	S.D.	
Change in Non-Human Organisms	15	6.2	41.3	0-15	3.98	3.0	49.2	0-14	2.58	.80
Change in Humans	15	6.1	40.7	0.15	4.22	3.5	56.4	0-14	3.11	.83
Change in Non-living Things	15	5.9	39.3	0-15	4.21	3.2	54.2	0-14	2.74	.84
CHANGE Module	45	18.1	40.0	0-44	9.84	9.7	53.6	0-33	6.89	.89

[1]Percent of objective problems in the problem area test
[2]Percent of problems chosen [3]Cronbach's alpha

Essay problems were provided at each evaluation period. In the CHANGE module, a total of fifteen problems were included in the three Problems to Solve booklets. Students were asked to select one Choose Your Problem to respond to at each evaluation period. Responses to their selections were returned to the Human Sciences Project for coding and scoring.

Student achievement varied from just below to a little above the 50-percent level (Table 29). Achievement on these essay problems was at the same general level of performance as was achievement on the objectively-scored problems.

TABLE 29

Mean Achievement of Students on the One Choose Your Problem They Answered for Each Problem Area and for the Sum of Three Choose Your Problems

Problem Area	Mean	%
Change in Non-human Organisms	1.6	53.3
Change in Humans	1.8	60.0
Change in Non-living Things	1.4	46.7
CHANGE module	4.8	53.3

Did the Activities Make a Difference?

Data to use in explaining the effects of activities on success with objective problems were obtained by crosstabulations of students "doing" and "not doing" activities, by whether they answered a related objective problem or not, and if they answered the problem correctly. Most objectively-scored problems were specific to single activities, but several problems related to more than one activity. There were twelve activities that had no objective problems related to them. In all, sixty-three comparisons were made between one objective problem and one activity. The results of these comparisons are shown in Figure 6. For forty-two (66.7 percent) comparisons, students doing the activity chose to answer the related problem in greater proportions than those not doing the activity. There were no significant differences between students "doing" versus students "not doing" activities, as determined by the chi-square statistic. Chi-square was pre-set at the 0.05 level.

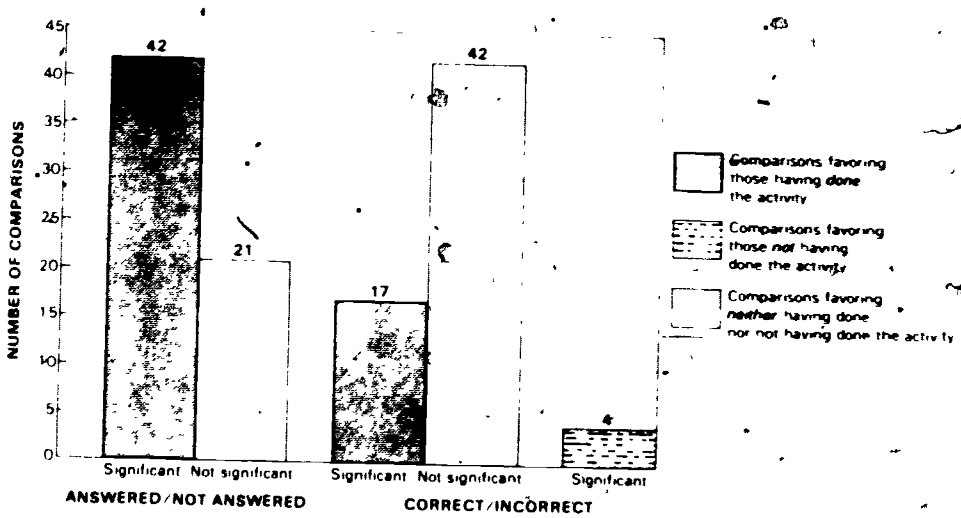


Figure 6. "Objective Problems by Activity" comparisons computed for CHANGE. There were sixty-three comparisons.

In this report significant difference means $p<0.05$, unless specified otherwise.

Having done the activity made a significant difference in getting the problem correct for 27 percent of the comparisons. In 6.3 percent of the comparisons, students who did not do the activity answered the problem correctly in greater numbers than did students who did the activity. For two-thirds of the comparisons there was no significant difference between those doing and those not doing the activities.

In terms of correctly answering most objective problems, doing the related activity did not appear to be advantageous. It must be remembered, however, that students were advised to be selective and to do activities for ideas and problems they didn't know about, but to answer every test problem they wished. The assumption can be made that many students who did not choose a particular activity already had some knowledge about it. The option to select "I did not choose this problem" seemed to provide students the opportunity to avoid problems they could not answer and to reduce guessing. This observation has minor evidential support from one objective problem that was printed with this option omitted. All students answered that problem.

Cummulative Effects of Activities and Skill Booklets

The possible influence of the cummulative effects of several activities and skill booklets requiring similar skills or cognitive learning, as measured by objectively-scored problems, was investigated for four activity/skill booklets (see Table 30). Cumulative effects were found

TABLE 30
Activities, Skill Booklets, and Objective Problems Examined to Determine Cummulative Effects

Group Number	Activity/Skills Booklet Titles	Related Objective Problems	Common Idea	Results of Analysis
1	The More the Better Making and Using	8	Interpreting Graphs	-
2	Yeast It! Change the Recipe	1 6 9	Yeasts are organisms Cause of fermentation Cause of fermentation	? + +
3	Change and Change Again Of Time and Temperature Working with Fruit Flies	4 10	Metamorphosis Life cycle of fruit fly	- -
4	Change and Change Again Of Time and Temperature Microbes in Milk Working with Fruit Flies	11	Incubator function	-

in three of seven investigations. Explanations of two of these investigations will illustrate these findings.

The common idea for Group 2 and objectively-scored problem 1 was that yeasts are organisms, not lifeless chemicals. The question investigated was, "Will the proportion of students who do both Group 2 activities perform better on objectively-scored problem 1 than the proportion doing more or only one of the Group 2 activities?" Figure 7 displays the positive cummulative effects of doing both activities on performance, both for choosing to answer the problem and in selecting the correct response. Similar effects were found for the Group 2 activities and for objective problems 6 and 9.

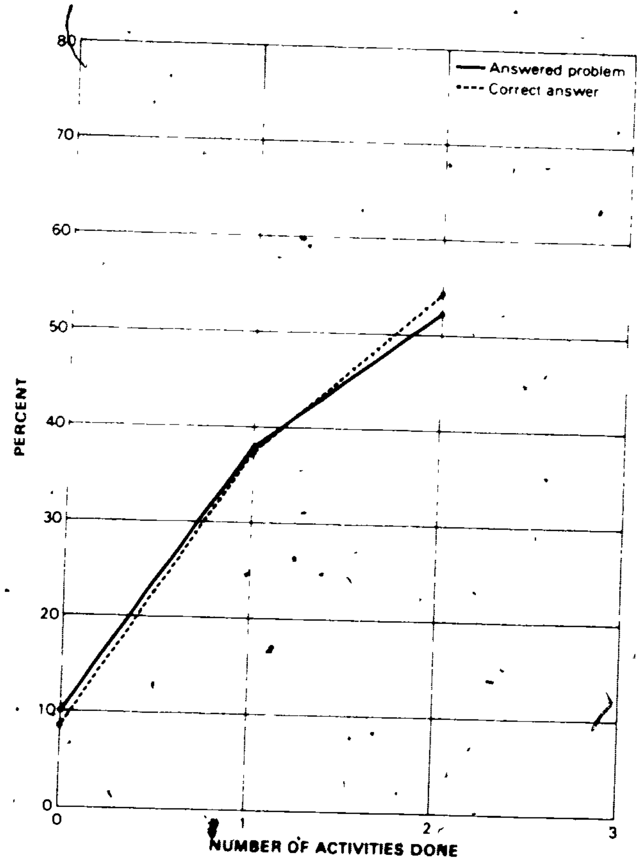


Figure 7. The effects of doing zero, one, or two activities on answering and selecting the correct response to related objective problem 1. The activities are "Yeast It" and "Change the Recipe."

Negative results were found in four of the seven investigations. Investigating the same problem proposed above, but for Group 3 activities and objective problem 10, doing only one of three activities resulted in increased performance over none of the three. Doing either one additional related activity and/or using the related skill booklet did not improve performance (Figure 8).

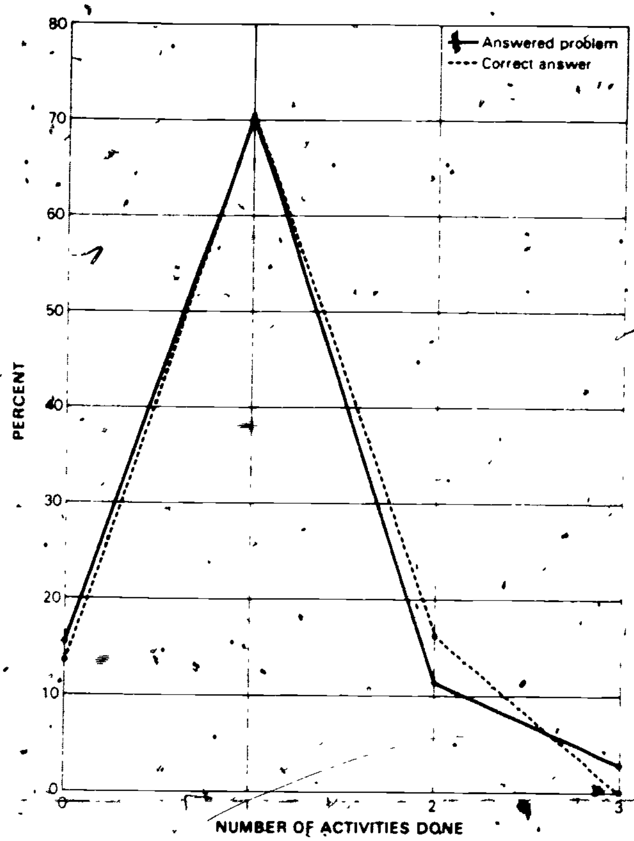


Figure 8. The effects of doing zero, one, two, or three activities on answering and selecting the correct response to related objective problem 10. The activities are "Change and Change Again," "Of Time and Temperature," and the skills booklet, "Working with Fruit Flies".

These seven investigations showed that when students chose activities that could be related to each other (Group 2), their performance on objectively-scored problems was increased. When potentially related activities were not written so as to complement one another, cumulative effects did not result. The negative results in the analysis are consistent with the low usage of the skills booklets and point to potential activity interrelations that were strengthened in activity revision.

Evaluation of FEELING FIT

Activity Use Patterns

This module contained fifty-three activities grouped into three problem areas. An evaluation period (or periods) was scheduled before each teacher decided to introduce another problem area.

Activity choice records were kept in the FEELING FIT folder, with students transferring the data to optical scan sheets with the same six choices used in CHANGE (Table 29). The mean number of activities done for boys, girls, and the total group is shown in Table 31. There was no significant difference between the number of activities done by boys and girls (one-way analysis of variance, data not shown). The mean of about fifteen activities represented 28.3% of the fifty-three activities in the module. Although the percentage is lower than that of CHANGE, the total number of activities done was about the same, even though the use of FEELING FIT was shorter than CHANGE by two to four weeks. Analysis of activities by problem areas showed that fewer of the community-oriented activities in the last problem area were chosen by students than were activities in the first two problem areas. The standard deviations showed a greater variance of number of activities done by boys than by girls.

TABLE 31
Mean Number of Activities Done in All Problem Areas of FEELING FIT

Students	Number of Students	Number of Activities Done		
		Mean	Standard Deviation	Range
Boys	151	15.0	8.54	-
Girls	168	15.7	6.59	-
TOTAL	321	15.4	7.57	0-40

The content emphasis of the fifty-three activities in FEELING FIT showed a pattern similar to that found in CHANGE, with significant differences (one-way analysis of variance) between boy and girl choice patterns in activities with physical science and social science emphasis. There was no difference between choice patterns of boys and girls in biologically-oriented activities (Table 32).

TABLE 32
Content Emphasis in FEELING FIT as a Function of Activities Chosen by Boys and Girls

Content Emphasis	Percent of Activities in Module	Percent of Activities Chosen		
		Boys	Girls	All Students
Biological Sciences	57.7	54.1	51.1	52.5
Physical Sciences	3.8	4.8*	3.7*	4.2
Social Sciences	26.9	27.0**	32.3**	29.9
Interdisciplinary	11.5	11.8	10.6	11.1
TOTALS	100.0	97.7	97.7	97.70

*$p = .05$ **$p < .001$

What Did Students Accomplish?

There were forty-five objectively-scored problems in FEELING FIT, fifteen for each problem area. Each problem was identified as to the activity or activities it was designed to assess. Students could elect to answer or not answer each problem, again selecting "I do not choose this problem" as one alternative response. Student achievement was 49.5

percent of the problems selected, as shown in Table 33.

TABLE 33

Mean Number of Objectively-scored Problems Chosen, Problems Correct, and Test Reliability Scored for Each Problem Area of FEELING FIT

Problem Areas	No. of Problems	Problems Chosen				Problems Correct				Reliabilities[1]
		Mean	%	Range	S.D.	Mean	%	Range	S.D.	
What Makes Me Healthy?	15	7.2	48.0	0-15	3.95	3.5	48.6	0-10	2.39	.79
Does My Health Depend on Others?	15	6.3	42.0	0-15	3.68	3.0	47.6	0-10	2.07	.77
How Does My Community Affect My Health?	15	7.2	48.0	0-15	3.76	3.8	52.8	0-12	2.61	.76
FEELING FIT Module	45	20.8	46.2	0-45	10.06	10.3	49.5	0-30	5.94	.90

[1]Cronbach's alpha

Referencing of the test items to activities was done to assist students in choosing problems in relation to the activities they had studied. The anticipated increase in mean achievement in FEELING FIT and CHANGE was not accomplished by this procedure as achievement was lower, but was still within the expected range of approximately 50 percent correct.

Essay problems were used only in the evaluation period at the end of the module. Seven problems cut across activities and were labeled, "Any activity." Seven problems were referenced to specific activities, and four problems were designed to be used with student-developed activities. Students were asked to choose and write responses to any two of these eighteen problems.

Most of the problems were not achievement problems and will not be reported here. Two achievement problems will be presented to serve as examples of student responses.

Problem F presented blood pressure data (systolic pressure over diastolic pressure) for ten persons--24 hours coffee-free, and then 30 and 60 minutes after drinking one cup of strong coffee. Averages for each time period were also presented: Caffeine free, $\bar{x}$=124.3/82.2; 30 minutes after coffee, $\bar{x}$=124.9/82.4; and 60 minutes after coffee, $\bar{x}$=125.1/83.0. Students were asked their conclusions about the effect of of coffee on blood pressure, based on the control data, and to give reasons for their response.

Student responses were codeable into four categories, as shown in Table 34. The modal tendency was to read the data literally, that small differences in blood pressure might not be a "real" difference. The fact that nearly 29 percent of the students responded in a way that they recognized either of these possibilities is rather surprising. Similarly surprising is the approximately 24 percent of the students who could not give a statement or read the data in reverse. These data were very useful in restructuring the activities dealing with blood pressure measurement. It was necessary to make explicit the problem of error and how large the differences in mean measurements need to be if they are to be interpreted as "different."

TABLE 34
Student Responses to Choose Your Problem F1

Response Categories	N	%
Blood pressure goes up	41	47.1
Blood pressure goes down	16	18.4
Blood pressure stays the same	25	28.7
No statement of effect	5	5.7
TOTAL	87	99.9

Did Activities Make a Difference?

Forty-five objectively-scored test items were matched with one or more of the fifty-three activities in FEELING FIT, resulting in fifty-four item-by-activity pairings. Thirty activities were evaluated by objective problems. Figure 9, page 84, shows the outcomes of each of two comparisons. The first comparison shows that forty-five of the fifty-four comparisons for answering or not answering the problems were significant for those having chosen the related activity. Answering the problem correctly showed a significant difference for those having chosen the activity for six comparisons, with one comparison favoring those not choosing the activity.

Choosing and studying the activity related to particular objective problems made a difference, in most cases, on whether students chose to answer a particular objectively-scored problem. Students who hadn't done the related activity, but did choose to answer problems, responded as expected. They knew the material and generally were as able to select the correct response as well as students who chose the activity.

A comparison of a relatively easy and relatively difficult problem further illuminates the difference activities make. Table 35, page 84, shows that 81.5 percent of the students answered the problem (256/314) with 88.3 percent answering correctly (226/256). A little less than half of the the students who did the problem answered correctly and a few more than half who had not done the activity answered correctly. Those who had done the activity and those who hadn't were about equal in numbers. This comparison pair was in the "not significant" groups in both parts of Figure 9.

Table 36, page 85, shows a "significant" (Figure 9) comparison pair for a more difficult item. Only 39.9 percent of the students answered the question and only 36.8 percent of those answering the question answered correctly. The comparison further shows that 84.8 percent of the students who did the activity answered correctly as compared with 15.2 percent of those who didn't.

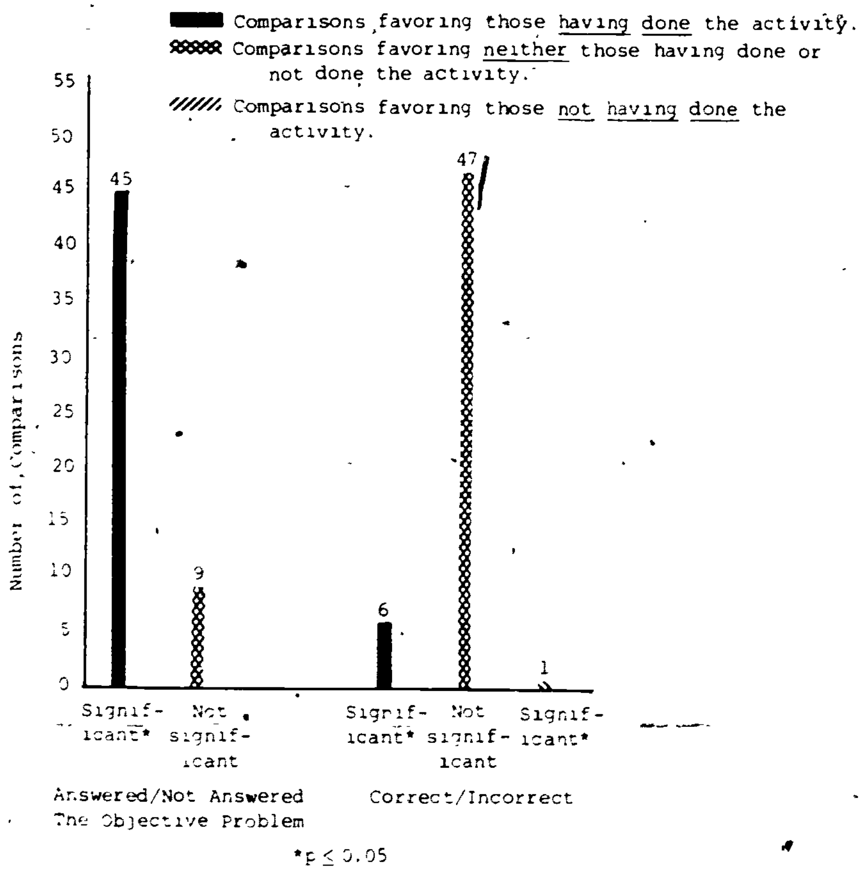


Figure 9. Objective problem-by-activity comparisons for FEELING FIT, computed separately for those who answered or did not choose to answer the objective problem and for those answering the problem correctly or incorrectly.

Even though there were only six comparison pairs in which differences in correct response to objectively-scored problem response favored those doing the related activity (Figure 9), twenty-four of the fifty-four comparisons favored those who did the activity. The opportunity for students to answer questions, whether or not they had done related study, is parallel with the option offered in regular science classes.

TABLE 35
The Relationship of Having Done the Activity, "Having Investigations," and Responses to Problem 25 on Acne

	All students		Answered		Correct answer	
	N	%	N	%	N	%
Did activity	139	44.3	123	48.0	106	46.9
Did not do activity	175	55.7	133	52.0	120	53.1
Totals	314	100.0	256	100.0	226	100.0

TABLE 36
The Relationship of Having Done the Activity, "Venereal Disease," and Responses to Problem 31, on Syphilis

	All students		Answering		Correct Answer	
	N	%	N	%	N	%
Did activity	150	47.9	89	71.2	39	84.8
Did not do activity	163	52.1	36	28.8	7	15.2
TOTAL	313	100.0	125	100.0	46	100.0

The unique design of the evaluations study made it possible to separate those who claimed to have studied relevant material, an option not usually found in current science instruction. No data are generally gathered as to whether the student read the relevant text material or participated in related work in class.

These data show that activities made a difference in achievement in some instances. It also showed that students learned informally in or out of school and that this learning was useful in science achievement.

Skills Development

During the FEELING FIT evaluation period, students were asked to give their perceptions of what they had accomplished by doing the FEELING FIT module. The items and response frequencies are shown in Table 37, page 86. Each student was asked to respond to nine statements by marking "strongly agree," "agree," "uncertain," "disagree," or "strongly disagree." The first three statements referred to the use of a thermometer, sphygmomanometer, and microscope. Sixty percent of the students felt that they could read a thermometer and over 70% felt they they could use a microscope. As for the sphygmomanometer, only 30% felt that they could use one to measure blood pressure while 50% were at best uncertain about its use.

In response to other statements, over 50% of the students felt that they were writing more understandable answers to "Choose Your Problem," that they had learned new things about careers, and that they had tried to change some of their habits because of activities in FEELING FIT. The mean responses to statements about reading books and articles because of FEELING FIT activities done, and to making arrangements with adults to do activities in the community, were 2.8 and 2.7 respectively. About half of the students either were uncertain about these two statements or definitely disagreed.

The ninth statement in the Skills Development Questionnaire was open-ended. Did the students feel that they had improved in any skill important to them? The mean response of 2.2 indicates that they agreed that they had. However, only 21.6% of the students responded to this statement, whereas approximately 88% responded to the other statements. Such a low response level makes it difficult to interpret the data on this item.

TABLE 37
Skills Development, Results of a Self-report Questionnaire

	Strongly Agree 1	Agree 2	Un-certain 3	Disagree 4	Strongly Disagree 5	Mean Response
I can read a thermometer accurately.	31.1	29.0	19.2	6.4	2.7	2.1
I can take blood pressure accurately with a sphygmomanometer.	9.8	20.7	27.7	19.8	10.7	2.7
I can use a microscope to see small things.	45.4	28.7	8.5	4.3	1.5	1.7
I am improving my skill in writing understandable answers to "Choose Your Problem."	21.0	39.3	17.1	7.3	3.7	2.2
I read parts of books and articles, because of activities I did in FEELING FIT.	16.8	23.2	21.6	18.9	7.9	2.8
I am learning how to make arrangements with adults to do activities in the community.	15.5	22.6	25.0	18.6	6.1	2.7
I learned new things about careers in FEELING FIT that I didn't know before.	19.8	34.8	21.3	6.4	5.5	2.3
I tried to change some of my habits that affect my health because of activities in FEELING FIT.	21.3	29.9	21.0	9.8	6.4	2.4
Overall responses on means	20.1	25.4	17.9	10.2	4.9	2.1
Any skills you improved upon, during the module, that were important to you.	6.1	6.7	5.2	1.5	1.2	2.2

Evaluation of INVENTION

The INVENTION module was designed to test the utility of long activities that considered five questions about each of twenty-nine inventions. Each activity was prepared in a series of parts with (generally) each part considering a single question, or sometimes two questions about each invention. The questions proposed for each activity are:

- How does it (the invention) work?
- How is it important to me?
- How is it important to others?
- What has been its past?
- What will be its future?

What Students Did: Activity Use Patterns

Teachers were asked to encourage students to choose an activity (invention) of interest to them and to do at least two parts of the activity before selecting a different invention to study. In practice some teachers in test schools emphasized this suggestion, others did not. There was no formal requirement that all parts, or even more than one part of an activity be attempted. In most instances the first part of each activity involved making something, or performing some other "hands on" experience (including interviewing). Other parts of each activity included analysis of data or of how the invention worked, and reading and/or writing about the invention.

Questions proposed for this part of the study were:

- Which activity parts were most highly used?
- What did students do in these activity parts?
- What use was made of the other parts of these same activities?
- What did students do in these parts?
- Are there any patterns of choice that emerge from a comparison of the highly-used parts of activities and the other parts of the same activities?

The twenty-nine activities in INVENTION consisted of as few as two to as many as thirteen parts (Table 38). In all, students could choose from 154 different activity parts.

TABLE 38
The Frequency of Number of Parts in the Twenty-nine Activities in INVENTION

Number of Parts	f
2	1
3	5
4	7
5	5
6	3
7	5
8	0
9	2
10	0
11	0
12	0
13	1
Total	29

Data for activity use patterns were recorded by each student voluntarily on her or his INVENTION Folder. At the end of the module these data were transferred to an optical scan record sheet. The data reported here were obtained from the optical scan sheets.

The INVENTION folder listed each activity by title. Two columns were provided for each activity listed. The first column directed, "Circle the parts of the activity you did." The second column directed, "Circle the parts of the activity you want to be accountable for."

The optical scan activity record instructed students to refer to their INVENTION folder and for each part of each activity:

- Mark 1 if you circled the part in both columns.
- Mark 2 if you circled the part in the first column, but not the second.
- Mark 3 if you circled the part in the second column, but not the first.

o Mark 4 if you didn't circle the part in either column.

If instructions were followed, data (a mark in spaces 1-4) should have been obtained for every student. Missing data for the 154 parts ranged from 6.6 percent to 12.2 percent of the 288 students from whom data were obtained.

INVENTION was used in six of seven test classes and for a range of nineteen days to fifty-eight days with a mean of thirty-two days. The seventh test school used INVENTION for only seven days as teachers were on strike for a long period and could not complete Level III testing.

A review of the data from the optical scan sheets showed that twelve activity parts were used by 11 percent or more of the students--as indicated by responses 1, 2, or 3. Seven activity parts were used by 10.0 to 10.9 percent of the students. The arbitrary use of 11 percent or more as "highly" used activities was selected for this analysis. Tables of use patterns of the twelve selected parts and of the unselected parts of these same activities were constructed. The analysis below is based on the data displayed in these tables.

Two generalizations are warranted, based on the data from Table 39, page 89.

- o All twelve of the highly used parts involved students in making the invention, operating it, sorting it, sorting pictures of it, interviewing about it, or viewing a film about it.
- o Seven of the twelve highly used parts required no written work. In two activities, recording data was required from picture sorting or from interviews. In three activities, from three to fourteen short answer essay questions were required, but for two of these three parts, the activity did not direct students to complete this part of the activity.

The percentages of students who completed the various parts of these "highly selected" activities were calculated in two ways on the six graphs, Figures 10-15, pages 90-91. First, students were counted as having completed the activity only if they marked the activity record "done and accountable." This gives the most conservative view of activity completion. Second, students were considered to have completed the activity if they marked any one of the three responses, "done and accountable," "done only," or "accountable only." This computation gives a more liberal accounting of the number "completing" an activity part and is probably the more accurate figure.

Two generalizations are made based on the data displayed in these graphs.

- o After the initial "action" part of the activity, approximately two-thirds to one-half of the students went on to complete most or all of the remaining parts of the activity (see Figures 10-15).
- o The pattern of usage of the lesser chosen parts of highly chosen activities indicates that students deliberately chose these parts. A separate computer record search found <u>no</u> students who did only <u>one</u> part of an activity. In the majority of activities, however, the last parts of activities were the least chosen.

TABLE 39

The Twelve Most Highly Used[1] Activity Parts in INVENTION, a Level III Human Sciences Module, 1975-1976, N=288 Eighth-Graders

Activity		1 Done & Account	2 Done Only	3 1 & 2	4 Account Only	5 1 & 2 & 4	What Do Students Do?	Paper Work	N Schls
Printing	IA	15.3	5.2	20.5	6.6	27.1	Make a block print using rubber bands, bits of fabric, etc.	None	5
	IB	13.2	5.9	19.1	6.9	26.0	Make a block print using linoleum.	None	7
Money, Money, Money	IA	12.8	5.9	18.7	4.5	23.2	Make coins by striking from dies.	3 essay questions, not referred to in activity	6
	IB	12.5	2.8	15.3	6.3	21.6	View film "Of Art and Minting."	5 essay questions, not referred to in activity	5
The Camera	I	15.3	5.9	21.2	9.0	30.2	Make a pinhole camera.	None	5
	II	11.5	5.9	17.4	6.9	24.3	Take photos with their pinhole camera.	None	6
The Telephone	IA	19.1	6.6	25.7	3.1	28.8	Interview at least 5 people, some older, some younger, analyze.	Complete interview sheet and 9 essay problems	5
Shoes	I	17.0	7.3	24.3	13.9	38.2	Sort and categorize pictures of shoes.	Record data on record sheet	6
	IIA	13.2	7.6	20.8	14.6	35.4	Make plaster cast of foot.	None	6
	IIB	9.7	6.9	16.6	13.9	30.5	Make shoes from wood and leather.	None	6
	IIC	6.9	5.9	12.8	13.5	26.3	Test shoes for 1-2 days.	None	[illegible]
Machine Shop	I	11.1	3.8	14.9	7.3	22.2	Build a mechanical toy.	14 essay questions	4

[1] The criteria for selection of "most highly used" were that 11.0 percent or more students reported to have done the activity, wished to be "accountable" for the activity, or both.

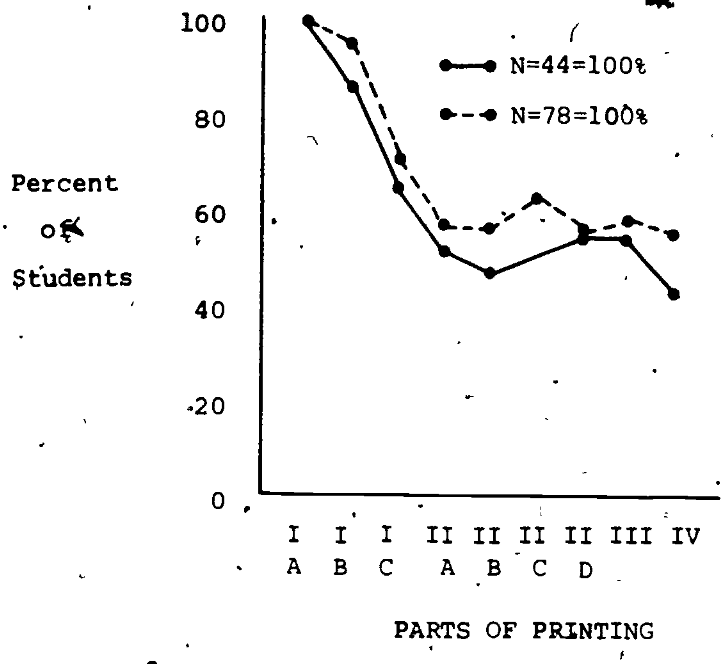


Figure 10. Percentage of students who participated in the activity, "Printing."[1]

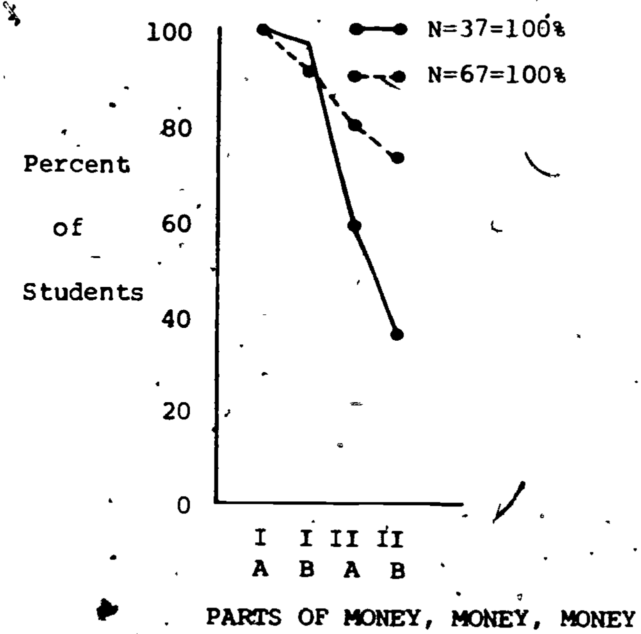


Figure 11. Percentage of students who prticipated in the activity, "Money, Money, Money."[1]

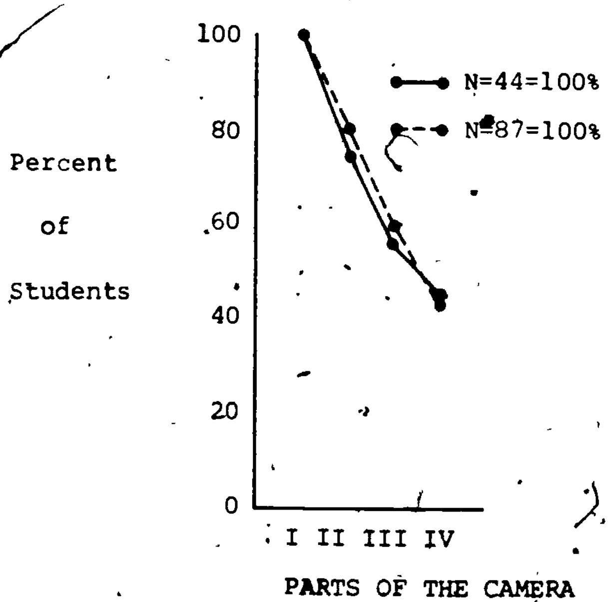


Figure 12. Percentage of students who participated in the activity, "The Camera."[1]

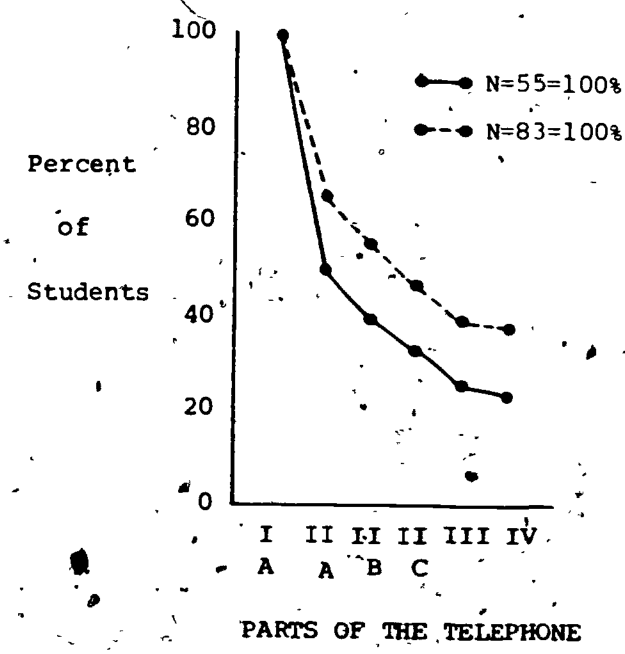


Figure 13. Percentage of students who participated in the activity, "The Telephone."[1]

---

[1] ●——● indicates students who marked the activity "done and accountable" only. ●- - -● indicates students who marked the activity "done and accountable," "done only," or "accountable only." In both instances the number doing the first part of the activity equals 100 percent.

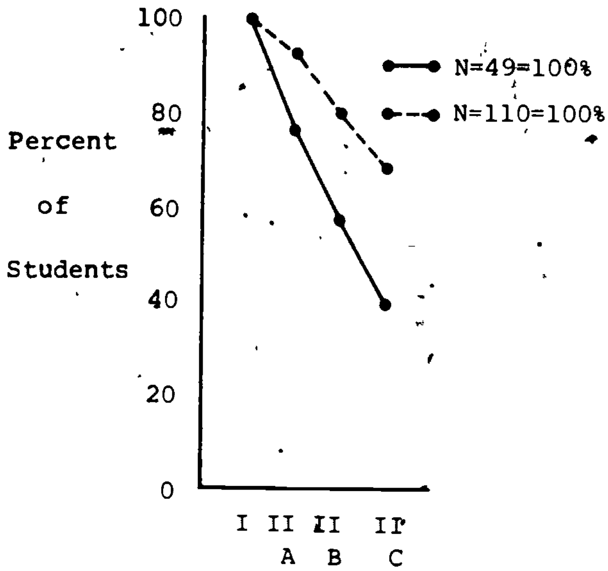


Figure 14. Percentage of students who participated in the activity, "Shoes."[1]

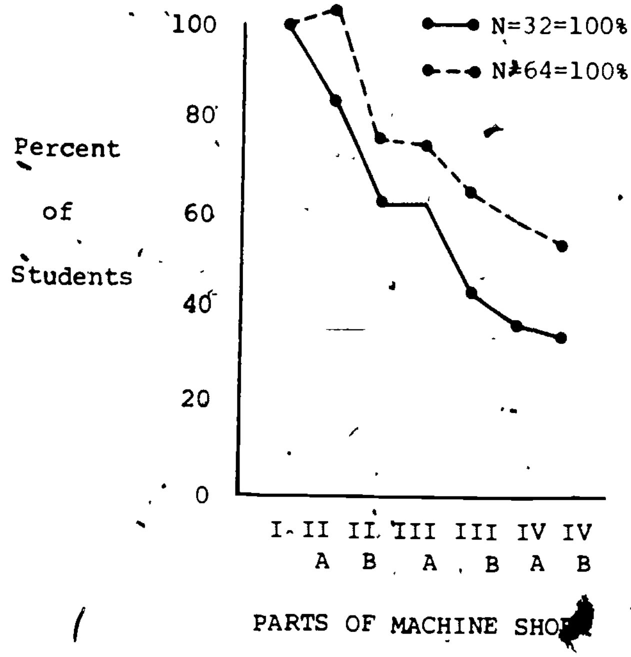


Figure 15. Percentage of students who participated in the activity, "Machine Shop."[1]

Data for the lesser used parts of the most highly used activities in INVENTION are shown in Table 40, page 92. The first generalization below is supported by data from this table. Four additional generalizations are based on data displayed in both Tables 39 and 40.

- o The lesser chosen parts of activities are difficult to generalize about in terms of what students did. For example, it is probable that in "Shoes" and "Machine Shop" the lack of material resources reduced participation in the lesser used parts. Readings that were in reference books rather than part of activities were not well used. Use of this reference in the revision of the module is questionable.
- o In general, where analysis and interpretation were separated from the action part of an activity (as in "Machine Shop," for example) more students chose to do the action part than the corresponding analysis part.
- o A comparison of required written work associated with activity parts indicates that written work did not appear to influence choice patterns.

1 ●——● indicates students who marked the activity "done and accountable" only. ●– – –● indicates students who marked the activity "done and accountable," "done only," or "accountable only." In both instances the number doing the first part of the activity equals 100 percent.

TABLE 40

The Lesser Used Parts of the Most Highly Used Activities in INVENTION, 1975-1976, N=288 Eighth-Graders

Activity		1 Done & Account	2 Done Only	3 1 & 2	4 Account Only	5 1 & 2 & 4	What Do Students Do?	Paper Work	N Schls
Printing	IC	10.1	4.2	14.3	5.2	19.5	Make other block prints--essentially an extension of IA, B.	None	4
	IIA	8.0	4.2	12.2	3.5	15.7	Making and printing with a paper stencil.	None	4
	IIB	7.3	3.5	10.8	4.9	15.7	Printing through a screen.	None	4
	IIC	7.6	4.2	11.8	5.2	17.0	Printing with screen plus paper stencil.	None	4
	IID	8.3	3.8	12.1	3.5	15.6	Printing with screen plus glue stencil.	None	5
	III	8.3	2.4	10.7	5.2	15.9	Investigate mimeo, ditto, and copy machines in school.	None	5
	IV	6.6	3.1		5.2		Investigating printing in newspapers, print shops.	None	5
Money, Money, Money	IIA	7.6	3.5	11.1	7.6	18.7	Bartering.	3 essay questions and chart	4
	IIB	4.9	3.5	8.4	8.7	17.1	Inventing a medium of exchange.	4 essay questions	4
The Camera	III	8.7	3.8		5.9		Read about how cameras work, history of camera.	None	3
	IV	6.9	4.2		5.6		Make own photograph paper and print.	None	4
The Telephone	IIA	9.7	4.5		4.9		Select a communications task and determine when phone or face to face is better.	None	4
	IIB	7.6	6.3		2.4		Try doing without phone, or have parent or business person do without.	None	5
	IIC	6.6	4.2		3.1		Learn about information in phone book; emergencies, etc.	None	4
	III	5.2	4.2		2.4		Read about how phone works.	Make a diagram	4
	IV	4.9	3.8		2.8		Read about history of telephone.	None	4
Machine Shop	IIA	9.4	8.3		5.2		Operate a model steam engine.	None	6
	IIB	6.9	4.5		5.6		Analyze engine; RPM, high speed.	Record RPM's, 3 essay questions	5
	IIIA	6.9	3.5		6.3		Operate a model gasoline engine.	None	5
	IIIB	4.9	2.8		6.9		Analyze engine: needle valves, exhaust parts, dissemble engine.	3 essay problems, function of parts	5
	IVA	4.2	4.2		4.9		Assemble an electric motor; make it run.	None	5
	IVB	3.8	2.8		5.6		Analyze electric motor: draw, use magnet explain how motor works.	5 essay questions	4

- Highly used parts were not dependent upon being used in all seven test sites, but were dependent upon being used by four or more test sites. The lesser used parts of highly selected activities were accompanied by a reduction in the number of sites at which these parts were chosen.

The INVENTION module provided a unique opportunity to test with eighth-grade students the utility of designing long-range activities broken into smaller parts. In the activity design, students were urged to do at least two parts of an activity before deciding to choose to change to another or to choose among the remaining parts of the activity chosen. Data from seven test schools (N=288) were analyzed. No students were found who chose only one part of an activity. However, choice of the second part to do was not limited to the logical second part, although most choices followed this pattern. Data from the six activities (twelve parts) included in this study indicate that from two-thirds to one-half of the students chose to continue with the activity to complete all parts. A consequence of this choice pattern for INVENTION was that the remaining twenty-three activities received less use than was typical for other Level III modules.

What Did Students Accomplish?

Forty-five objectively-scored problems were used to determine student achievement in INVENTION. Because a questionnaire about INVENTION was judged to be important, the questionnaire was used in place of essay problems. The objectively-scored problems were divided into three groups of 15, to be administered at the completion of each problem area during the module. Precise data were not obtained about usage, but in most schools evaluation was delayed to the end of the module and students were asked to answer all problems they felt they could answer correctly. As in FEELING FIT, all problems were identified as to the activity they were designed to evaluate. The range of parts within activities made it impossible to develop problems that would assess understandings of most activities in total. In most instances, problems related to one part, or at most two parts, of activities. In some instances, problems compared several activities.

Student achievement, as reflected by mean scores on each of the three fifteen-item subtests and the full, forty-five item objectively-scored problems, is shown in Table 41, page 94.

Achievement was lower than in other Level III modules in all three problem areas. In this module students chose to answer problems in a proportion of over three times the number who chose the related activity. Apparently, students felt they could handle the problems from their common knowledge, but scores showed that they could not solve problems about the particular invention. About two-thirds of the problems in this problem area were at intellectual levels of application and analysis, with the lowest level being comprehension.

TABLE 41

Mean Number of Objectively-scored Problems Chosen and Answered Correctly, with Subtest and Test Reliabilities

Problem Area	Number of Problems	Problems Chosen		Problems Correct		r^1
		Mean	%	Mean	%	
How has this invention brought us closer together?	15	8.2	54.5	5.8	38.4	.92
How has this invention made life easier?	15	5.9	39.0	6.1	40.4	.94
How has this invention affected length of life?	15	8.3	55.2	6.2	41.6	.95
Totals	45	22.2	49.4	18.0	40.1	.98

[1]Cronbach's alpha

Did the Activities Make a Difference?

Cross tabulations between activity parts done in relation to answering a related objectively-scored problem, and between activity part done and whether the problem was answered correctly or not were used to determine if the activity made a difference in achievement. Figure 16, page 95, shows that in twenty-nine (64 percent) of the comparisons, doing the activity positively affected answering the problem. Once having chosen to answer a problem there were no problems that were significantly affected by having done the related activity. In five comparisons, there was a signifcant difference in answering the problem correctly by those not doing the activity. For four of the five items, the number answering the problem correctly was about four times greater for those not having done the activity. In forty of the forty-five comparisons there was no significant difference in answering the problem correctly between those who did and those who did not do the activity.

Activity choice made a positive difference in whether students chose to answer a problem in INVENTION, but it did not make a difference in getting the problem correct, once chosen. This result led to an examination of the five activities to which the problems were related and to a revision of the activities. Changes were also made in the evaluation items for the activity to make the items more closely related to the instruction in the activity.

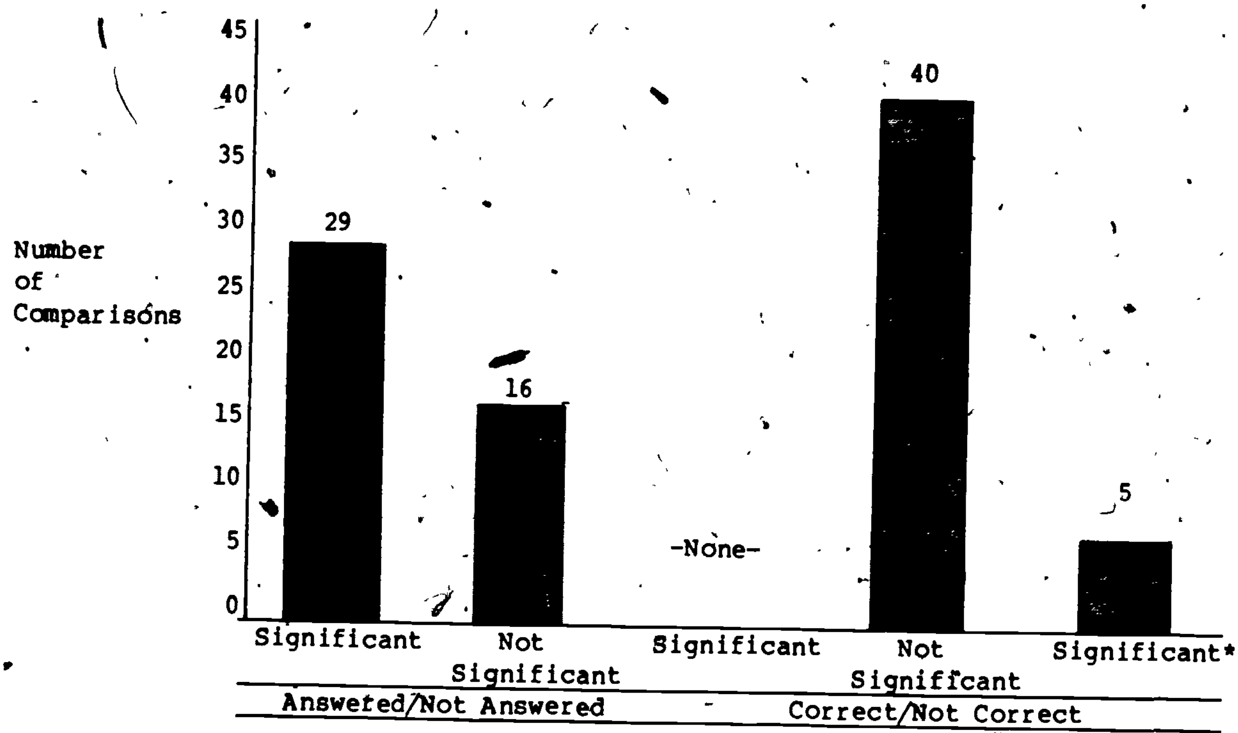
Skills Development

Data from seven, five-response (agree-disagree) statements were obtained. The central tendency for the statements was to the agree-side of the midpoint, uncertain. Three findings from the Skills Development rating scale are:

- o Thirty-eight to 42 percent of the students agreed with statements about skills making models of inventions, making them work, and in being able to explain how they work.
- o Reading the module references, The Way Things Work and How It Works, when the activity suggested it, had the lowest agree

percentages (30 percent each) and highest disagree percentages (21 percent each).

o Learning to make new products and use new tools had the strongest agreement of the seven statements (about 50 percent.)



*The problem was answered correctly by those who had not done the activity. The test for significance was $\chi^2 \leq .05$.

Figure 16. Objective problem by activity comparisons computed for INVENTION. There were 45 comparisons.

INVENTION Questionnaire

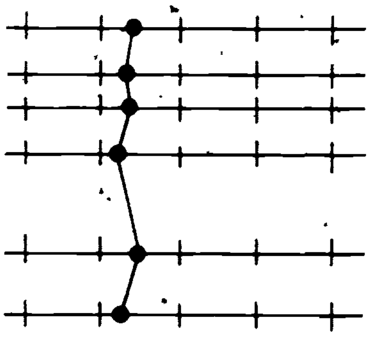
Data from sixteen, five-choice (agree-disagree) statements were obtained to secure answers to three questions. The first question was "Do students like long activities with many parts that characterized the INVENTION module? Six statements were designed to answer this question. Mean item responses (scale 1-5, 3.0 is "uncertain") ranged from 2.2 to 2.6 with a group mean of 2.4. Strongest agreement (mean of 2.2) was in support of the activity many-part structure as an aid to learning (see Table 42, page 96).

A second way to analyze these data is to combine the "strongly agree" and "agree" responses, the "disagree" with "strongly disagree" and to display these data along with the "uncertain" responses. These data support two response groups of three items each. For example, the agreement among items 4, 7, and 15 supports the position favoring short activities. The agreement among items 1, 10, and 13 supports the position favoring the advantages of long activities (Table 43, page 96).

TABLE 42

Mean Responses for Six Agree-disagree Statements Used to Answer the Question "Do Students Like Long Activities with Many Parts that Characterize INVENTION?", N=255

Statement	SA 1	A 2	U 3	D 4	SD 5
1. I like INVENTION because I didn't have to make as many choices as in other modules.					
4. I liked activities that had one part best.					
7. Many activities in INVENTION were too long.					
10. I liked the activities in INVENTION because doing the parts helped me learn about one thing.					
13. I liked INVENTION because I could work for a long time on the same thing.					
15. I would rather have shorter activites than the ones in INVENTION.					



This grouping of six items is supported on logical grounds, but item intercorrelations (Table 44, page 97) also support two subscales with item intercorrelations within each subscale (1, 10, 13 and 4, 7, 15) that are almost twice as high as item intercorrelations between subscales.

TABLE 43

Percentage of Students Agreeing, Disagreeing, or Uncertain for Six Items about the Structure of INVENTION Activities, N=255

Item No.	Paraphrases of Statement	Desired Response	Agree (+)	Disagree (-)	Uncertain
1	I liked INVENTION because I didn't have to make many choices.	(+)	41.6	28.2	30.2
4	I like activities that have one part best.	(-)	47.9	25.8	26.2
7	Many activities in INVENTION were too long.	(-)	44.7	27.7	25.6
10	Like activities, doing the parts helped me learn more.	(+)	55.7	17.8	26.5
13	Like activities to work for a long time on the same thing.	(+)	44.7	29.0	26.3
15	Would rather have shorter activities than the ones in INVENTION	(-)	42.2	27.8	30.0

Mean responses for each subscale (items 1, 10, 13 and 4, 7, and 15) are 2.46 and 2.48, respectively. Students agree to the same degree for both of these subscales. Subscale reliabilities are .68 and .70 respectively. The subscales are weakly correlated (r=.30).

What can be said about student attitudes toward activity structure, as exemplified by INVENTION activities? Students supported the values of the longer activities that characterized INVENTION, but they also

TABLE 44

Intercorrelations between Items Related to Activity Structures, N=255

	Item Numbers				
	10	13	4	7	15
1	.38	.42	.24	.24	.18
10	-	.45	.12	.08	.31
13		-	.17	.10	.10
4			-	.41	.42
7				-	.44

preferred shorter activities than those in INVENTION. Since half of the activities in the module had 5 or less parts, but nearly one-third had seven or more parts, the responses make sense.

Did students think the activities were written clearly and simply? Four statements were designed to answer this question. Mean item responses ranged from 2.2 to 3.5 with a mean for the four items of 2.9. Strongest agreement (mean 2.2) was for the statement in Item 8 (Table 45). A pair of items, 2 and 8, expressed the attitude that reading the activities was difficult. Items 8 and 11 were statements that expressed attitudes that activities were easy to understand. Table 45 shows that students were consistent in disagreeing with the first subscale and agreeing with the latter subscale. As can be seen in Table 46 (page 98), the items in one subscale, items 2 and 5, are more highly correlated with each other than with items in in the contrasting subscale, Items 8 and 11. These data support the interpretation that students were in agreement that the activities in INVENTION were clearly and simply written.

TABLE 45

Percentage of Students Agreeing, Disagreeing, and Uncertain for Four Items about the Readability of INVENTION Activities, N=255

No.	Paraphrases of Statements	Agree	Disagree	Uncertain
2	Many directions hard to follow	28.7	49.4	21.8
5	Many words I didn't understand	17.9	53.1	29.0
8	I could read and understand [illegible]ons	69.1	14.1	16.8
11	[illegible]ons of how inventions [illegible]re easy to understand	55.8	14.9	28.7

Did students like the objective problems? Six statements were included in the questionnaire to seek an answer to this question. Mean scores for the statments ranged from 1.8 to [illegible].3, with the mean for the six items at 2.7. The strongest positive response was in being able to choose the objective problems (Item 3, Table [illegible]).

TABLE 46

Intercorrelations between Items Related to Activity Readability

	Item Numbers		
	5	8	11
2	.64	.17	.24
5	-	.15	.22
8		-	.47

Another way to look at the answer to this question is to combine the percentage of the two "agree" responses and the two "disagree" responses and display these along with the "uncertain" responses (Table 47). It is rather clear that students like to choose objective problems (Item 3) although about one-third think they should be required to do objective problems for activities they do (Item 12). A better

TABLE 47

Percentage of Students Agreeing, Disagreeing, or Uncertain for Six Items Related to Objective Problems in the INVENTION Module

No.	Statements	Agree	Disagree	Uncertain
3	I like to be able to choose the objective problems I want to do.	77.7	4.9	17.4
6	I understand how the objective problems are related to activities.	45.3	15.4	39.2
9	My answers to objective problems show my understanding of activities pretty well.	51.5	13.7	35.2
12	We should be required to solve objective problems for activites we have done.	35.7	31.2	33.1
14	I don't see how I could have learned the answers to objective problems from the activity.	20.1	38.8	41.1
16	Objective problems are hard to understand.	24.0	39.2	36.9

understanding of the way students responded to the items can be made by examining item intercorrelations shown in Table 48.

TABLE 48

Intercorrelations between Items Related to Objective Problems

	Item Numbers				
	6	9	12	14	16
3	.37	.32	.25	.09	.17
6	-	.41	.41	.20	.19
9		-	.47	.23	.34
12			-	.36	.35
14				-	.63

Items 14 and 16 correlate most highly with each other and have much lower correlations with other items. Through their responses to this subscale, students agree that the objective problems seem unrelated to the activities, but that they are not hard to understand.

The other four items--3, 6, 9, and 12--make a subscale, with intercorrelations ranging from .32 to .47. This subscale reflects positive views toward problem choice, toward the relation of answers to student understanding, and more weakly, to the relationship of problems to activities. Items 6 and 12 complement each other in support of students' understanding of the relationship of problems to activities, although there was a large "uncertain" group for these problems. The response to item 9 further strengthens this interpretation.

Evaluation of SURROUNDINGS

The last module to be developed and tested during the three-year test period, 1973 to 1976, was SURROUNDINGS. Only three of the seven test schools used any part of this module. The evaluation materials were changed from the three previous modules to test a design that was felt to be the best design for the module revision.

This new evaluation design included the following elements:

1. A SURROUNDINGS folder for students to record starting and terminating data for each activity chosen. Students were also asked to record the data they completed on an "Activity Evaluation Form" and on a "Problems to Solve" paper for each activity they studied.
2. An "Activity Evaluation Form," printed on an optical scan sheet to be completed by a statement when an activity was completed. This generalized form had eight Likert-type items and two open-ended essay questions.
3. A one-page, activity specific quiz, "Problems to Solve," was developed for each activity in the module. Students were asked to secure a "Problems to Solve" sheet for an activity when they had completed the "Activity Evaluation Form" for that activity.
4. A "My Activity Record Form," used to cross-check the activities completed by each student. This form was completed at the end of the module, with students using data from their SURROUNDINGS folder.

This new evaluation design made it impossible to further explore the question "Does choosing an activity make a difference in achievement?" as students were now required to take a quiz on each activity they studied, and were not given the option to respond to questions in activities they had not studied.

The change in design was a further shift away from formative evaluation of the curriculum, to a formative evaluation of the evaluation materials for student and teacher use. Unfortunately, this new model was not adequately tested in SURROUNDINGS. It was effectively tested in the field test of KNOWING, as will be explained in Chapter 7.

The SURROUNDINGS module was not used at all in Schools 1, 4, 6, or 9. Using dates secured from machine-scored forms for INVENTION evaluation and for final SURROUNDINGS evaluation, the following maximum days for use of SURROUNDINGS was deduced:

School 5, Teacher 3 - 11 days
School 7, Teacher 3 - 11 days
School 7, Teacher 6 - 13 days
School 7, Teacher 7 - 10 days
School 8, Teacher 1 - no data

SURROUNDINGS consisted of twenty activities. Table 49, page 101, shows the data collected by school and teacher identification number.

Activity Use Patterns

There were three potential sources of information about activity selection patterns: "My Activity Record," "Activity Evaluation Form," and "Problems to Solve." If students and teachers had complied with the procedures suggested, a count of each of the forms would have yielded the same number, since an "Activity Evaluation Form" and a "Problems to Solve" sheet were to be completed for each activity a student did. The number of forms and sheets should have been equal and equal to the total number of activities marked as being done on "My Activity Record."

The actual numbers did not reflect the expected consistency among these records. From "My Activity Record" responses, 166 "Activity Evaluation Forms" should have been received; only 80 were received. Even fewer (47) "Problems to Solve" sheets were received, from an expected 173 as reported on "My Activity Record." In a check with teachers, some reported that the module was too rushed so they did not forward the forms. Teachers took time for end-of-year evaluation activities in the final school-closing period. Others indicated that they did not work with students to observe the evaluation procedures as reflected in the instructions. The data show, therefore, only relative, not absolute, use of the activities in this module.

Sixty-six students (of about ninety-five in test classes) turned in one "Activity Evaluation Form," twenty-one turned in two, and ten turned in three. From these data, only 22 percent of the students reporting did two activities; about 9 percent did three. These data produce approximately 1.4 as the mean number of activities completed by a student. Data from "My Activity Record" indicate a mean completion of about 2.5 activities per students.

The most highly chosen activities (those chosen by 20 percent or more students reporting having chosen the activity in "My Activity Record") were: "How Well Do Others Know You?," "Electronic Surroundings," "Wet Pets," and the "Relationships Game."

All activities were reported as being chosen by at least four students, but "Problems to Solve" data were not received for five of the twenty activities (see below). Activity use, for the time period, was comparable to that of other Level III modules, as reported on "My Activity Record."

TABLE 49

SURROUNDINGS Data Collected

My Activity Record		Incomplete
TCHR 53	N=33	0
73	N=25	0
76	N=18	1
77	N=21	0
Total	N=97	N=1

Self-Rating Problems, Work Habits		Incomplete
TCHR 53	N=33	0
73	N=25	1
76	N=18	1
77	N=21	0
Total	N=97	N=2

Self-Rating Problems, Skills Development		Incomplete
TCHR 53	N=33	0
73	N=22	3
76	N=16	2
77	N=21	0
Total	N=97	N=5

Total Number of Students		Complete Data of Students
TCHR 53	B=33	N=33
73	N=25	N=20
76	N=18	N=15
77	N=21	N=21
Total	N=97	N=89

Activity Evaluation Form		Number of Activities/Students
TCHR 53	N=33	33/1 ACT
73	N=11	5/1 ACT, 3/2 ACT, 1/4 ACT, 2/5 A
76	N=19	5/1 ACT, 10/2 ACT, 4/3 ACT
77	N=17	13/1 ACT, 3/2 ACT, 1/3 ACT
81	N=14	6/1 ACT, 6/2 ACT, 2/3 ACT
Total	N=94	62/1 ACT, 22/2 ACT, 7/3 ACT, 1

Problems to Solve		
TCHR 53	N=0	
73	N=13	6/1 ACT, 4/2 ACT, 2/4 ACT, 1/5
76	N=17	6/1 ACT, 7/2 ACT, 4/3 ACT
77	N=2	1/1 ACT, 1/2 ACT
81	N=14	6/1 ACT, 6/2 ACT, 2/3 ACT
Total	N=46	

What Did Students Achieve?

"Problems to Solve" quiz sheets were obtained for fifteen of the twenty activities in SURROUNDINGS. The number of student responses for these activities ranged from one to thirteen. The distribution of "Problems to Solve" sheets received is shown in Table 50. Thirteen sheets were received for each of two activities, but, as mentioned above, five activities were without achievement data. As can be seen in Table 50, any interpretations of student achievement must be qualified by the small number of cases upon which achievement was based. Only three activities had evaluation data for more than ten students, making mean achievement calculations meaningless.

TABLE 50
Frequency of "Problems to Solve" Sheets Received by Activities

Number of Student "Problems to Solve" Sheets	Number of Activities
13	2
11	1
8	1
5	3
4	3
3	2
2	2
1	1
0	5
47	20

Table 51, page 103, shows the distribution of the number of items correct (thirty-nine items total) by percentage of correct responses. Thirteen, or 33 percent, of the items were answered correctly by all (100 percent) of the students responding to them. Ten items, 29.6 percent, were answered correctly by 59.9 percent or less of the students responding to them. Half of the items were responded to by about 80 percent or more of those who answered the items. In general, the trend in SURROUNDINGS was toward a higher level of achievement than was recorded for the other Level III modules.

Work Habits

A "Work Habits" self-rating Likert-type scale was used in SURROUNDINGS. The scale was factor-analyzed, using the SPSS factor analysis program with principle components and RAO's cannonical subroutines, both with varimax rotations. Both subroutines produced similar solutions. The eleven items were analyzed into two factors (Table 52, page 103). The factor structure is sound, accounting for 60.1 percent of the variance. Conceptually, Factor 1 is designated "independence", and Factor 2, "participation." Students rated themselves on a scale from 1 to 5, described as (1) most of the time, (3) about half the time, (5) not often, and (2) and (4) at intermediate rating positions.

TABLE 51

The Number of Items Responded to in Relation to the Percentage of Students who Responded to the Item Correctly

Percent of Students of Those Responding Who Responded Correctly	Number of Items
100.0	13
90.0-99.9	2
80.0-89.9	4
70.0-79.9	4
60.0-69.9	3
50.0-59.9	1
40.0-49.9	0
30.0-39.9	3
20.0-29.9	3
10.0-19.9	0
0.0- 9.9	6
Total	39

Responses were generally in the "1" or "2" category. That is, students responded to such statements as Item 1, "I work independently," and Item 2, "I don't need to be supervised," by marking either (1) 51.2 percent, or (2) 25.5 percent. The factor structures were not simple since Items 3, 4, and 5 were split between the two related factors (r=.63). However, the single factor loadings of Items 1 and 2, and 6 through 12, support the conceptualization of the factors.

TABLE 52

Factor Structure of the Self-rating Scale, "Work Habits"

Item Number	Factor Loadings	
	Factor 1	Factor 2
1	.18553	.81399
2	.19905	.76065
3	.48376	.44139
4	.43311	.56130
5	.45644	.58320
6	.59832	.25755
7	.74098	.24263
8	.77834	.25300
9	.64763	.17848
10	.54714	.32816
11	.56585	.19421

Students were very positive about their work habits in SURROUNDINGS. The same scale was also marked by students with the statement, "In comparison to INVENTION, I have ________." This scale yielded a single factor that accounted for 61.4 percent of the variance. Students were also very positive in responding to this scale. Ten of the eleven items had ratings by more than 50 percent of the students for response choice (1) or (2). Only on Item 6, "I start a new activity when I finish one," did students mark choice (3) more heavily (42.6 percent), indicating "no improvement" compared to their actions during the INVENTION module.

Skills Development

Skills development was evaluated by a nine-item self-rating scale (Table 53). Mean responses show that students rated six of the nine responses as "Don't Know," with only two items with mean scores between "agree" and "don't know." Actual percentages of responses to each of the five response choices indicated that the modal group of students did not take a stand on six items, but chose the "didn't know" response. There seemed to be a group 30 to 40 percent of the students who indicated positive skill development, and a smaller group, 15 to 20 percent, who were negative about their development of skills.

TABLE 53
Mean Scores and Response-choice Scores for Eleven Self-rating "Skills Development" Problems for SURROUNDINGS, N=90

Self-rating Problems	Strongly Agree 1	Agree 2	Don't Know 3	Disagree 4	Strongly Disagree 5
I kept animals alive and have learned how to care for them.	23.1	24.2	36.2	5.5	11.0
I used books such as "How to Know the Insects" or "How to Know the Wild Flowers" and identified organisms successfully.	12.1	18.7	40.7	11.0	17.6
I improved my skill in taking photographs by using photography in SURROUNDINGS activities.	13.2	16.5	37.4	15.4	16.5
I made collections of plants or animals I had never collected before in SURROUNDINGS.	15.3	23.1	34.1	11.0	16.5
I can use a laboratory balance to accurately weigh materials.	12.1	25.3	38.5	11.0	13.2
I know how to measure the pH of liquids.	17.6	22.0	31.9	12.1	16.5
I can now introduce myself to strangers to conduct interviews.	22.0	33.0	28.6	7.7	8.8
I have successfully arranged to leave the classroom to do Human Sciences activities.	26.4	34.1	28.6	3.3	7.7
I have successfully arranged for myself (and a friend or small group) to leave the school grounds to do a Human Sciences activity.	13.3	24.4	37.8	11.1	13.3

Summary

During the school year 1975-1976 four Level III modules of the Human Sciences Program were produced for use in seven test schools.

Only three schools used any part of the SURROUNDINGS module, the last module prepared for testing during 1975-1976. Teachers were not confined to specific ending dates for the modules, but were to conclude module use when student interest lagged in selecting new activities.

Table 54 shows the duration of use of each module by each teacher in the test schools. End-of-year activities forced the termination of INVENTION in most test schools, so the time spent does not reflect the potential duration of this module. SURROUNDINGS was tested at only three schools. Time spent at School 8 is not known and data were incomplete. Since CHANGE was not available for testing until the first week in October, more school time for these modules would be available in a full school year.

TABLE 54
Number of Days Modules Were Used by Schools and Teachers

School	Teacher	Number of Days per Module				
		CHANGE	FEELING FIT	INVENTION	SURROUNDINGS	TOTAL
1	2	61	48	38		147
4	3	85	34	42		161
		87	42	42[1]		160
3	3	53	25	58	15	151
2	4	70	65	7		142
7		69	48	27	12	156
	6[2]	71	46[1]	27[1]	12[1]	156
		71	46	25	14	156
8	1	69	48	19	?	136
9	1	82	43	37		162
All Test Schools	Mean	70.7	44.5	32.2	13.3	
	Standard Deviation	9.3	10.3	14.3	1.5	

[1]Approximations based on the starting and finishing dates for other teachers in the same school using HSP.

[2]School 6 was closed during a teacher's strike for several weeks.

Modules varied in number of activities and in the length and duration of activities. CHANGE had forty-six activities and three skills booklets; FEELING FIT, fifty-three activities; INVENTION, twenty-seven; and SURROUNDINGS, twenty. Although CHANGE and FEELING FIT were somewhat comparable, CHANGE had several long-term activities and FEELING FIT had one problem area with many community-based activities. In some schools, the community-based activities in FEELING FIT were not encouraged and hence were not effectively used. Only twenty of the proposed thirty-eight to forty activities in SURROUNDINGS were produced for field testing due to limited time for testing.

INVENTION had a different design from the other Level III modules. It contained only twenty-seven activities. These activities were generally long, most having from four to six parts. In many of these activities two or three parts were comparable to a single activity in CHANGE or FEELING FIT. INVENTION was designed to make it possible for students to study a topic in depth. Students were encouraged, but not required, to do more than one part of the activity they chose.

The mean number of activities "done" for each module is shown in Figure 17, page 107. The data for CHANGE probably show the optimum amount of activities completed or "done," based on the criteria that "doing" an activity meant doing at least one part of the activity. The drop in number of activities done in FEELING FIT compared to CHANGE probably reflects the lack of use of community-based activities. The large decrease in doing activities in INVENTION is at least partially due to the lack of time in many classes. A second factor contributing to the decline is the way "doing" an activity was calculated. For activities with two or three parts, a student was considered to have done an activity if 50 percent of the parts were completed. For activities with more than three parts, a student was considered to have "done" the activity if four or more parts were completed. These two factors most likely account for the decrease in percentage of activities completed in INVENTION. The fact that students persisted with many parts of the activities they chose precluded their choosing more activities.

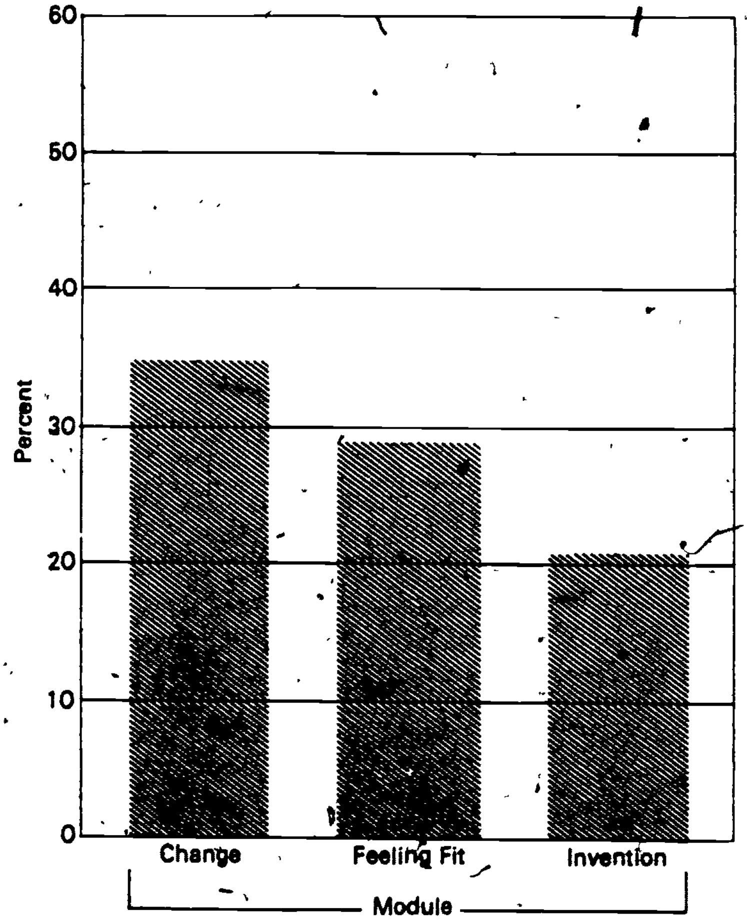
A comparison of activity choice by content of the activity for three of the four Level III modules again shows parallels between CHANGE and FEELING FIT with INVENTION showing differences. SURROUNDINGS' use was too limited for an analysis to be comparable.

For CHANGE and FEELING FIT, choice of activity by content paralleled the content proportion in the module.[1] However, potential for increased choice by students for each category remained high. The major difference between CHANGE and FEELING FIT appeared in the selection of social science oriented activities with proportionately less of these activities chosen than were available in CHANGE and proportionately more chosen than were available in FEELING FIT.

INVENTION shows a different pattern. The proportion of physical science and interdisciplinary activities chosen is much lower than the proportion available whereas the proportion of biological, earth and social science activities chosen is more than double the proportion available. This difference can be most simply explained as reflecting initial choice of activities since INVENTION was not used in test classes as long as the other two modules (Figure 18, page 108).

---

[1]It is important to recognize the arbitrary aspects of categorizing many of the activities by content. Operational criteria were not developed for this task and different observers might categorize activities differently. Teachers and students were not given information regarding content emphasis.



Modules

Figure 17. The percent of activities done in three Level III modules.

The final comparison of three of the four modules is presented in Figure 19. The difference between CHANGE and FEELING FIT is probably not statistically significant. That is, the percent correct and percent chosen are about the same. The same relationship probably holds between FEELING FIT and INVENTION. However, there is a decline in achievement from CHANGE to INVENTION that appears to be real. There is also a reversal in INVENTION. A smaller proportion of problems were answered correctly than were chosen. This shift may reflect end-of-year phenomena in most test schools, but there were not sufficient data to explore the problem in more detail.

CHANGE, FEELING FIT, and INVENTION have been compared on five variables: time in classrooms, proportion of activities done, proportion of objective problems chosen, and proportion of objective problems correct. Time spent in each module and achievement declined, as measured by the proportion of objective problems answered correctly. Given enough time in a module, students seem to choose activities from the full range available. It does not seem likely that the content of the activity, per se, was an important criterion in determining student choice.

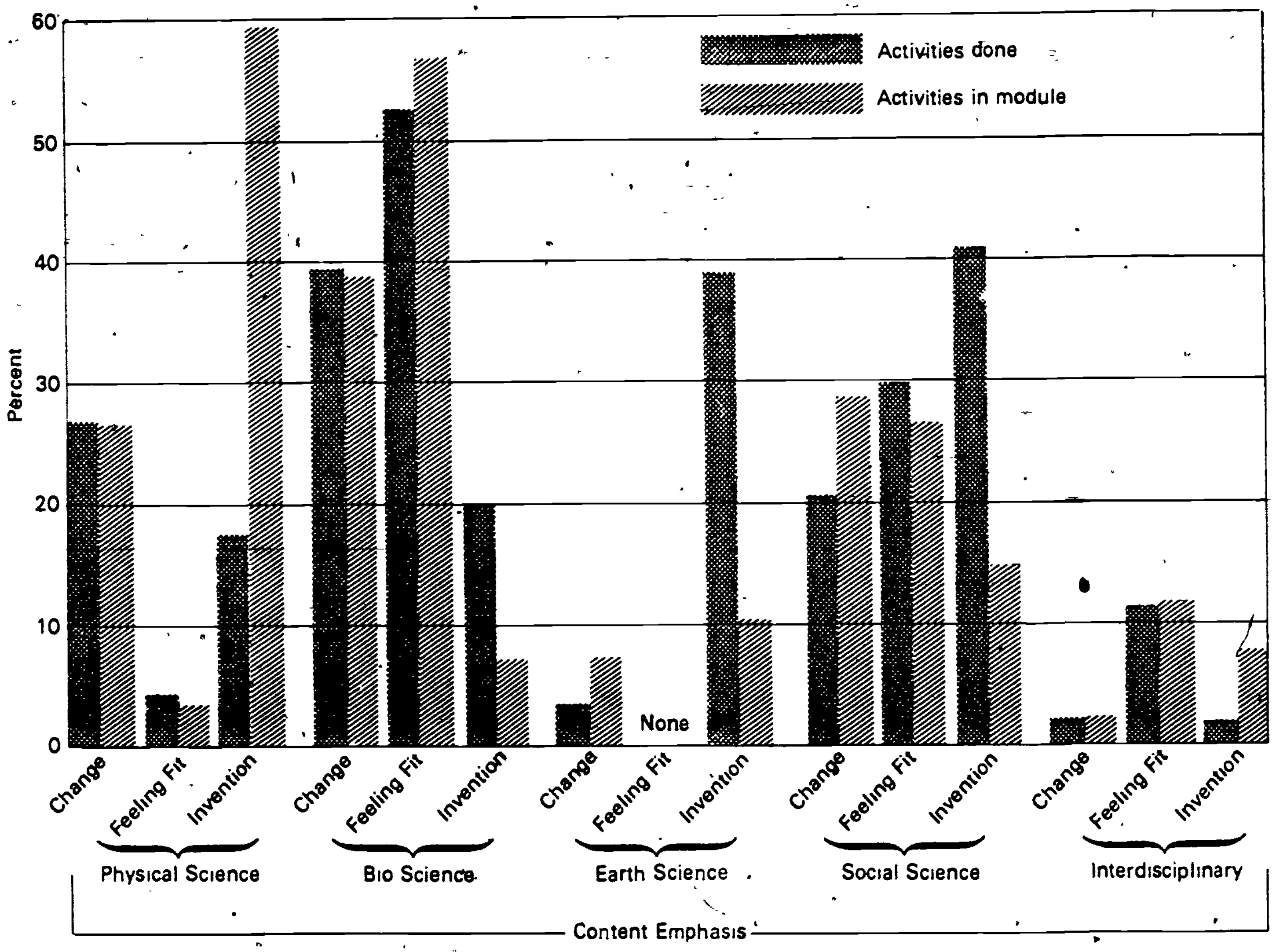


Figure 18. The proportion of activities with each of five content emphases compared to student choice of these activities for three Level III modules.

SURROUNDINGS was not adequately tested in 1975-1976. First, the proposed final number of activities for the module was not produced for testing as few schools had time for testing, and those that did have time did not have adequate time to test the twenty activities provided. Achievement, though based on limited data, seemed higher than for the other Level III modules. Discussions with test teachers indicated that the module was too simplistic for eighth-graders. As a result of the limited evaluation data, but based primarily on teacher feedback, SURROUNDINGS was revised to be used with sixth-graders.

Evaluation data from CHANGE, FEELING FIT, and INVENTION, indicated that these modules were much closer in activity structure to the final model adopted for commercial revision of the materials and would not need as much revision as the Level I and Level II materials. This result was one of the advantages of the curriculum development model that provided for continuous feedback from the formative evaluation into the design of the next set of materials to be produced.

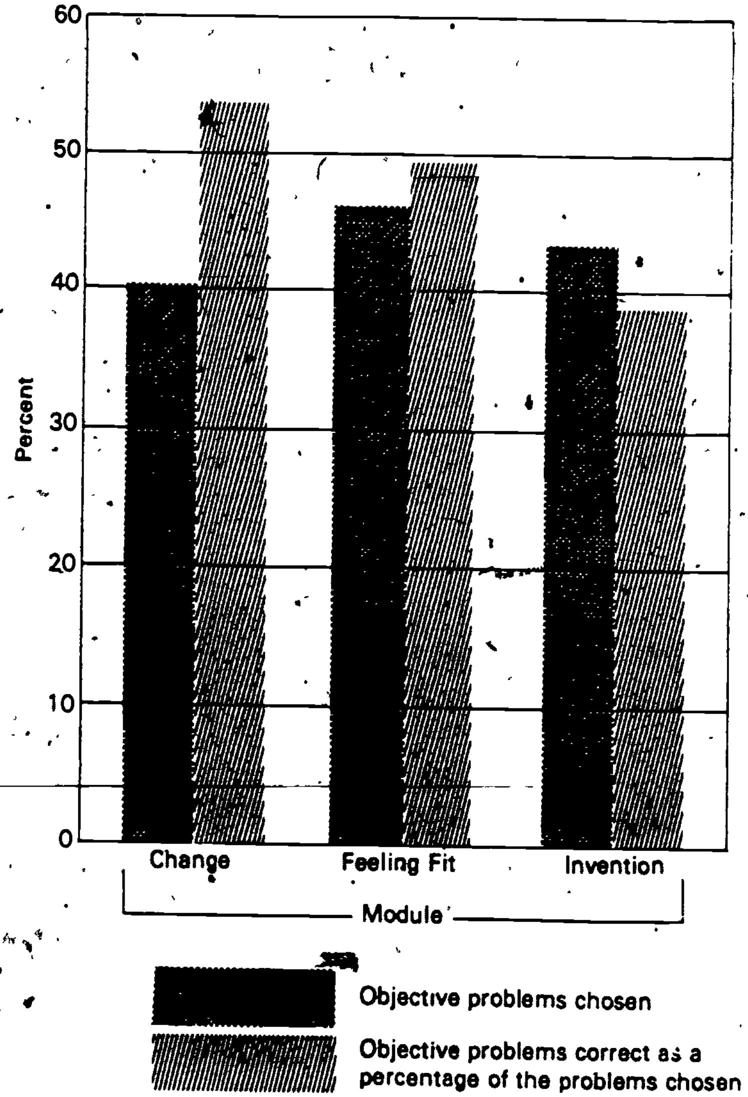


Figure 19. Percent of objective problems chosen and percent of objective problems correct for three Level III Human Sciences modules. There were forty-five objective problems for each module.

CHAPTER 7

EVALUATION OF KNOWING, A LEVEL III MODULE

The results of the Level III field test indicated that SURROUNDINGS would be a more appropriate module for younger students, either sixth or seventh graders. This finding made necessary the development of an additional Level III module in order to provide a full-year, eighth-grade curriculum.

KNOWING was conceptualized in the spring of 1976, and written, developed, and produced for field testing in 1976-77. Delays in funding delayed field testing until April 1977. When funding for field testing was received, selecting test sites and implementing a formative evaluation plan had to be completed quickly.

Evaluation Plan

Time limitations, both for the field test and for testing students, resulted in a pre-posttest evaluation design, using a 50% random sample of students for each of two pretest measures and a 50% random sample for two posttest measures.

Achievement was measured by "Problems to Solve" quizzes made specifically for each activity. When students completed an activity they were to get a "Problems to Solve" quiz for that activity and to complete an "Activity Evaluation Form" specifically for that activity. Table 55

TABLE 55
Evaluation Instruments for the One-group, Pretest-posttest Design

Pretest	"KNOWING Pretest" (35 items) (50% of students)	"How Is Your Logic?" (30 items) (50% of students)
Throughout module	"Problems to Solve" quiz for each activity studied (all students)	"Activity Evaluation Form" for each activity studied (all students)
Posttest	"KNOWING Posttest" (35 items) (50% of students)	"Science Questionnaire" (36 items) (50% of students)

shows the instruments used during the pretest and posttest, and during the use of the module. The design was quasi-experimental, following Campbell and Stanley (1963), designated as a one-group, pretest-posttest design. The treatment period lasted from April 7, 1981 to the end of the school year, a period of from 9 to 11 weeks in the selected test schools.

Selection of Field Test Sites

The selection criteria described for the selection of schools for the three-year field test (Chapter 3) were modified because of the necessity to make arrangements quickly after notification that funding was received and field testing could be done. Three-year test school teachers and administrators were contacted by questionnaire to determine if they would be interested in testing KNOWING during April to June, 1977. In addition, college and university science and social science educators who had given workshops on Human Sciences were contacted to recommend teachers and schools who might be interested in testing the module. Questionnaires were also sent to those schools.

Final site selection was made to include seven experienced Human Sciences teachers and seven teachers who had not had training from the BSCS staff. This criterion plus geographical distribution were the major site selection criteria. Table 56, page 113, presents the KNOWING test sites. The twelve schools were widely distributed geographically and represented variation in school type, community served, and ethnic backgrounds of students.

Instruments Used in the Field Test

The forty-four-item KNOWING Pretest and Posttest were designed to control for prior knowledge of key ideas developed in the KNOWING module. The tests contained the same items, with items or item groups reordered. Both instruments consisted of two sections: a thirty-five-item set of groups of statements marked either "agree" or "disagree," and a nine-item multiple-choice section (four choices per item). The "agree" or "disagree" section included three groups of items (twenty-one items) related to an expository section and 14 unrelated items. Each of the multiple-choice items was independent of the others.

"How Is Your Logic?" 1976 edition was a twenty-item, two-form (A and B) test of logical competence. The test had been used and validated during the three-year Human Sciences field test.

A "Problems to Solve" quiz of from three to five items was developed for each activity. The items were a mixture of objectively-scored and essay problems. Scoring keys and suggestions for scoring were

TABLE 56
Field Test Sites and Descriptions of Schools, Students, and Classrooms Testing KNOWING

Identification Numbers			Ethnic Background of Students						
School Number	Teacher Number	Number Students	Anglo	Hispanic	Black	School Type	Grades Served	Community Served	Geographic Region
16	04	35	100%			Middle	6,7,8	Suburban	Midwest
21	03	30	51%	23%	23%	Junior high	7,8,9	Inner city	Northeast
22	05	64			100%	Junior high	7,8,9	Inner city	Pacific Northwest
17	07	30	65%		35%	Middle	6,7,8	Urban	North Central
14	01	60	100%			Middle	6,7,8	Suburban	North Central
	06	30	100%						
23	08	30	100%			Junior high	7,8,9	Suburban	North Central
	09	30	100%				7,8,9		
24	10	30	60%	38%	2%	Middle	6,7,8	Inner city	Southwest
19	02	30	75%	25%		Junior high	7,8,9	Suburban	Pacific Southwest
25	11	30				Middle	6,7,8	Inner city	Southeast
26	12	30	100%			Middle	6,7,8	Suburban	Pacific Northwest
27	13	30	100%			Middle	6,7,8	Rural	Midwest
28	14	60	100%			Junior high	7,8,9	Suburban	Rocky Mountains

provided to the field test teachers for their use in evaluating student achievement.

A general "Activity Evaluation Form" was prepared on optical scanning sheets for use with all activities in KNOWING. The form asked for activity title and time spent in and out of class on the activity, answers to eight Likert-type statements, and two open-ended essay problems.

The pre- and posttests were scored using the LERTAP program (University of Colorado, Laboratory of Educational Research). "How Is Your Logic?" was scored by Dr. William Gray and his graduate students, University of Toledo. The "Problems to Solve" form was coded and processed using the LERTAP program and various SPSS subroutines. The "Activity Evaluation Form" was processed by SPSS subroutines.

The "Science Questionnaire" from previous studies (Robinson, 1980) was used to determine student attitudes. This instrument contained two semantic differential scales, each with 18 bipolar adjective pairs. Conceptually, these adjective pairs were designed to measure four aspects of attitude: evaluation, value, activity, and judgment. The adjective pairs for each subscale are shown in Table 57. One semantic differential asked students to "Circle the number that expresses how strongly you feel about Human Sciences." The second was addressed to "your regular science course (before you started KNOWING)." The "Science Questionnaire" also had a section asking students to list and then to rank-order all the classes they were enrolled in during the year.

TABLE 57
Conceptual Design of the Semantic Differential Scales of the "Science Questionnaire"

Evaluation	Value	Activity	Interest
good-bad	close-distant	not active-active	interesting-boring
pleasant-unpleasant	full-empty	slow-fast	dull-exciting
sad-happy	worthless-valuable	still-moving	never fun-always fun
nice-awful	useful-useless	tired-lively	
fair-unfair	not important-important	listening-doing	

Results of the KNOWING Field Test

The KNOWING module contained forty-four activities grouped into eight topical clusters. For purposes of field testing, the clusters were grouped into two groups of four clusters each. Half of the field test classes began the study of KNOWING by making only Group I activities and clusters available. The other half of the classes began with the activities and clusters in Group II (Table 58). With these arrangements, both groups of activities were used early and later in the field test. The integrative activity, "Ways of Knowing," was used to initiate study of the module and as an optimal integrative activity toward or at the termination of module study.

TABLE 58
Cluster Titles and Activity Numbers in Group I and Group II Clusters, KNOWING Module

Group I Cluster Title	Number Activities	Group II Cluster Title	Number Activities
Knowing the Past	7	Knowing About People	8
Heavens Above	4	Knowing About Mars	3
The Human Body	6	Notions of Motion	6
Materials and Shapes	4	Whatever the Weather	5
	21		22

Ways of Knowing -- Integrative

Activity Choice Patterns

The eight to ten weeks during which KNOWING was used seemed to be adequate time for testing, although no data were gathered about whether time was a factor in non-choice of activities. Data for activity choice were obtained from two sources, completed "Activity Evaluation Forms" and completed "Problems to Solve." If all students followed directions completely, the numbers obtained by counting each of these forms would have been identical.

Table 59 shows the case of activities in Group I clusters. The discrepancies between the number of students completing the two forms is an error in failing to mark both forms. As can be seen from Table 59, the marking error ranged from 2.6% to 20.3%, a large discrepancy. (Discrepancy was calculated by subtracting the smaller number of forms

TABLE 59

Use of KNOWING Activities in Group I Clusters as Reflected in Data Collected from the Number of "Activity Evaluation Forms" Received (N=538) and the Number of "Problems to Solve Forms" Received (N=464)

Activity	Card No.	Completed Activity Evaluation Forms N	Completed Activity Evaluation Forms %	Completed Problems to Solve Forms N	Completed Problems to Solve Forms %	Discrepancy %
1 Strange Fossil	01	66	12.3	62	13.4	6.1
2 Time Travel into the Paleozoic	02	52	9.8	56	12.1	7.1
3 Counting with Carbon	03	48	8.9	44	9.5	8.3
4 Rosetta II	04	136	25.3	110	23.7	19.1
5 Where Did We Come From?	05	26	4.8	24	5.2	7.8
6 Patterns in Your Past	06	54	10.0	51	11.0	5.6
7 Pueblo People of the Past	07	39	7.2	37	8.0	2.6
8 The Solar Merry-Go-Round	11	40	7.4	33	7.1	17.5
9 Sun watch	12	103	19.1	92	19.8	10.7
10 The Star Gazers	13	75	13.9	62	13.4	17.3
11 What Do the Stars Know?	14	128	23.8	117	25.2	8.6
12 Human Ideas About Disease	15	57	10.4	52	11.2	8.8
13 Knowing About the Brain	22	81	15.1	74	15.9	8.6
14 Hot Spit	17	32	6.0	33	7.1	3.0
15 Levers of the Body	21	31	5.8	28	6.0	9.7
16 Farther and Faster	16	111	20.6	106	22.8	4.5
17 A Way of Seeing Inside the Body	23	118	21.9	94	20.2	20.3
18 Building with Bricks	24	71	13.2	67	14.4	5.6
19 Building Materials: How Good?	27	34	6.3	32	6.9	5.9
20 Foiled Again	25	83	15.4	75	16.2	9.6
21 Materials in Space	26	110	20.4	105	22.6	4.5

turned in from the larger and dividing by the larger.) This does not, of course, provide any data on students who failed to mark both forms or who misrepresented themselves in marking the forms tallied.

The use of all Group I activities with ranges of 24 to 136 students again confirms the diversity of interests of early adolescents. Similar diversity of use was found in the Group II activity clusters (see Table 60 with a range of 29 students choosing the least used activity and 160 choosing the most used. Five to six activities in each group were chosen by more than 19% of the students and four in each group were chosen by 7% or less. Splitting the module into two groups and reversing their classroom use gave more even usage of activities across the module than was found in the 1973 to 1976 field test. This difference indicates that perhaps the lesser use of activities found in the last problem area of most modules was affected more by lack of time than by lack of interest.

TABLE 60
Use of KNOWING Activities in Group II Clusters as Reflected in Data Collected from the Number of Activity Evaluation Forms Received (N=538) and the Number of Problems to Solve Forms Received (N=464)

Activity	Card No.	Completed Activity Evaluation Forms N	Completed Activity Evaluation Forms %	Completed Problems to Solve Forms N	Completed Problems to Solve Forms %	Discrepancy %
22 Images of Brush and Pen	31	46	8.6	45	9.7	2.2
23 How Old Are They?	32	160	29.7	144	31.0	0.1
24 The Unknown Millions	33	37	6.9	31	6.7	16.2
25 Vital Statistics	34	34	6.3	29	6.2	14.7
26 The Very Different Ones	35	67	12.4	67	14.4	0.0
27 Surveys, Samples, and Schools	36	78	14.5	71	15.3	9.0
28 Size Wise	37	145	27.0	133	28.7	8.3
29 Knowing Yourself	41	93	17.3	79	17.0	15.1
30 A Martian Test	42	126	23.4	118	25.4	6.5
31 Martian Tales	43	88	16.4	76	16.4	13.6
32 Four Views of Mars	44	78	14.5	73	15.7	6.4
33 Moving Words	45	106	19.7	103	22.2	2.8
34 Dancing Motion	46	39	7.2	37	8.0	5.1
35 Vibes	47	30	5.6	29	6.2	3.3
36 Rolling Along	51	80	14.9	78	16.8	2.5
37 Heavenly Motion	52	47	8.7	43	9.3	8.5
38 Magic Motion	53	115	21.4	107	23.1	7.0
39 The Rainmakers	54	66	12.3	55	11.8	16.7
40 Weather Music	55	47	8.8	53	11.4	11.3
41 Weather According to Granny Oakes	56	82	15.2	75	16.2	8.5
42 Do Dew Drops Drop?	57	62	11.5	60	12.9	3.2
43 The Storm	61	74	13.8	67	14.4	9.5
44 Ways of Knowing	62	29	5.4	-	-	-

What Did Students Accomplish?

Students completed a quiz specifically designed to assess the major objective of activities they chose to study. General achievement for all multiple-choice problems by each activity cluster is shown in Table 61. Mean achievement varied from 70% correct for the cluster "Knowing the Past" to 53% for the cluster "Heavens Above."

TABLE 61
Mean Scores and Percent Correct for Multiple-choice Items for the KNOWING Module

Activity Cluster	No. Items	Mean Score	Standard Deviation	Percent Correct
Knowing the Past	14	9.80	.90	70.0
Materials and Shapes	6	3.38	.72	56.4
Heavens Above	8	4.23	1.00	52.9
The Human Body	12	7.58	.94	63.2
Knowing People	15	9.51	.87	63.4
Knowing About Mars	5	3.12	1.00	62.4
Notions of Motion	10	5.92	.85	59.2
Whatever the Weather	8	4.51	.91	56.4
Averages	9.8	6.87	.90	61.6

Logical Competence

"How Is Your Logic?" was administered in April, 1977 to most students studying KNOWING. The item means are shown in Table 62, page 118. Concrete responses received scores of four, and most students responded appropriately on most concrete items, as shown by the means and standard deviations.

Concrete reponses were the mean responses for nearly all of the formal items as well, but the standard deviations show that many students were scoring at the "Concrete II" level (see Chapter 9) and some at formal levels of attainment (maximum scores of 7 or 8). These data are consistent with the scores of the eighth-graders in the three-year Human Sciences test classes.

Can Students Make Useful Judgments About Curriculum Materials?

The purpose of this part of the evaluation of KNOWING was to analyze data collected during the field test to determine the usefulness of student ratings of curriculum materials. Utility or usefulness was defined as making discriminating or differential judgements on various questions about activities they chose to do. Additionally, "useful" entailed that student ratings generate hypotheses and/or less formal conjectures about activity characteristics that could be verified in some way and that could provide new insights into designing student activities for emerging adolescents.

TABLE 62
Means and Standard Deviations for Twenty-six Items in "How Is Your Logic?" 1976 Edition (N=442)

Item	Concrete or Formal	Logic Description	Mean	S.D.
A1	C	Increasing series	4.0	.31
A2	F	Making correct implication	4.7	2.15
A3	F	Making correct implication (x)	4.7	2.16
A4	C	Decreasing series	3.8	.44
A5	F	Complete combination	5.1	1.70
A6	F	Permutation	4.6	2.31
A7	C	Correspondence of classes	3.7	.70
A8	C	Increasing/decreasing series	2.5	1.02
A9	F	Denying correct implication	4.4	2.15
A10	F	Denying correct implication (x)	3.8	2.27
A11	C	Many-to-one correspondence	3.6	.81
A12	F	Proportional reasoning	3.4	2.11
A13	F	Proportional reasoning (x)	2.5	1.30
B1	C	Increasing series	3.9	.34
B2	C	Correspondence of classes	3.6	.80
B3	F	Denying correct implication	4.0	2.18
B4	F	Denying correct implication (x)	3.2	2.22
B5	C	Decreasing series	3.8	.46
B6	F	Complete combinations	5.0	1.63
B7	F	Permutations	3.6	2.40
B8	C	Decreasing/increasing series	2.7	1.06
B9	F	Denying correct implication	4.8	2.25
B10	F	Making correct implication (x)	3.8	2.48
B11	F	Proportional reasoning	2.7	1.94
B12	F	Proportional reasoning (x)	2.4	1.32
B13	C	Many-to-one correspondence	3.7	.73

The data base for exploring the usefulness of student ratings of curriculum materials included data from "Activity Evaluation Forms" and "Problems to Solve." With this data base the following questions were asked:

1. What activities in KNOWING were the most highly chosen by students?
2. What activities were least chosen by students?
3. What relationships were found between activity choice data (questions 1 and 2) and student ratings of the activities on Likert scale statements, such as "This activity made me think" or "This activity was enjoyable."?

4. What response patterns were found in multiple-choice and essay problems that would support or falsify inferences obtained from the analysis of "Activity Evaluation Form" data?
5. What were the characteristics of activities rated by students at the extremes (high-low) on different statements?
6. What activity design characteristics could be formulated that are consistent with analysis of the data?
7. Can students make useful judgements about curriculum materials?

Each of these questions will be explored in turn.

The "most highly chosen" activities were arbitrarily limited to the activities chosen by 19%, or more, students. These activities are listed in Table 63 by the group within which they were chosen.

TABLE 63

The Most Highly Chosen Activities in KNOWING

Group I Activities		Group II Activities	
4.	Rosetta II	23.	How Old Are They?
11.	What Do the Stars Know?	28.	Size Wise
17.	A Way of Seeing Inside the Body	30.	A Martian Test
16.	Farther and Faster	38.	Magic Motion
21.	Materials in Space	33.	Moving Words
9.	Sun Watch		

The "least chosen" activities, arbitrarily designed as activities chosen by 7% or less of the student group, are listed in Table 64.

TABLE 64

The Least Chosen Activities in KNOWING

Group I Activities		Group II Activities	
5.	Where Did We Come From?	35.	Vibes
15.	Levers of the Body	25.	Vital Statistics
14.	Hot Spit	24.	The Unknown Millions
19.	Building Materials: How Good Are They?	34.	Dancing Motions

The most highly chosen activities varied considerably in their subject matter, from linguistics analysis in decoding several different languages in "Rosetta II," to plotting the movement of sunspots on a transparency using a time-series set of photographs of the sun and trying to deduce the sun's behavior from the plot in "Sun Watch." Careful reading of these activities indicated that the value of the activity was implicit in and simultaneous with the action itself. It was not separated from the action in the sense that the activity would derive its value solely from reference to a future goal or end. Nor did the essential learning come from questions that were answered at the end of the activity, after it was completed. Rather, the values were integral to the accomplishment of the activity, and knowledge was developed as various phases of the activity were developed. For example, in "Sun Watch" where the sunspots were plotted on a transparent acetate sheet, one saw a sunspot in photographs taken at "Time 1." A trend emerged after four or five of the eight photographs were plotted. Alternate hypotheses could be generated at this time and further checking could help the student deduce the sun's behavior.

Activities that entailed only reading were not chosen by as many students as those in which reading may have been a large part, but in which other kinds of action were essential. Another tentative finding from activity choice patterns was that sometimes too much choice was provided within an activity. An example was "Building Materials: How Good Are They?" in which a sequence of eight or nine different tests were provided and the student was given the option of choosing to do two or three of those tests to determine how good two different kinds of building materials were. In this activity it took a great deal of reading before students could decide what they wanted to do. Students found this kind of activity tedious.

Analysis did not indicate that there was a bias toward a particular content source for activities, but many activities in the KNOWING module were interdisciplinary. There did not seem to be a pattern of choosing natural science over behavioral science, or social science over fine arts-oriented activities. The pattern with regard to the content of the activities seems to be quite diffuse.

A detailed review of one activity illustrates the relationships between activity choice data and student ratings of activities. The most highly chosen activity in Group I was "Rosetta II."

Eight Likert-type items, two sentence-completion problems, and a section for "other comments" were used to enable students to evaluate each activity they chose in KNOWING. Each of these sources of data will be presented and then interpreted in an interrelated way to reflect student evaluation of the activity.

The Likert statement responses were factor analyzed by three different methods. This procedure was used to seek the most stable relations between statement responses. The similar patterns of factors permit the following interpretation.

Factor 1. <u>Satisfaction</u>. This factor accounted for 34.6% of the variance among factors. Satisfaction seemed an appropriate conceptualization of this factor. Items 1 and 8 had the highest loadings (see Table 66, page 121) with means between "agree" and "strongly agree" (see Table 65, page 121) Item 7, "I already knew most things in the activity," also loaded highly on this factor, as did Item 4, "The activity was too long." Mean scores for these items (see Table 65) show that students disagreed with both statements, but more strongly with Item 7. These responses contribute to activity "satisfaction." Item 3, "The activity made me think," had the lowest item loading on this factor. The inverse relation between Items 3 and 7 further supports the inclusion of Item 3 on the satisfaction factor.

Factor 2. Personal relevance. Conceptually, this factor combines both utility and importance, or value. It consists of two items, 5 and 6 (see Table 66), and accounts for 20.8% of the variance among factors. The mean response values for the item fell near the midpoint of the scale (Table 65) indicating that, for Rosetta II, students were undecided as to whether the activity had personal relevance.

TABLE 65
Means and Standard Deviations for Each of Eight Likert-type Statements Students Marked for Rosetta II

Statement	Item	Mean[1]	Standard Deviation
Important to me	5	2.8923	1.0583
Useful to me	6	2.8077	1.0199
Made me think	3	2.0154	.9063
Enjoyable	1	1.8615	.9544
I recommend it	8	1.8769	.0151
Too long	4	3.3692	1.3067
Difficult for me	2	3.5923	1.0833
Already knew it	7	4.1077	.8468

[1]Strongly agree = 1; undecided = 3; strongly disagree = 5

Factor 3. Difficulty. This factor included only one item (2) about which students responded toward the disagree side of undecided. This "factor" accounted for 14.2% of the variance among factors.

Student responses to the statement, "I chose this activity because ______" were categorized into three types: cognitive, attitudinal and logistic. Further differentiation was made within each area. Most of

TABLE 66
Factor Analyses of the Eight Likert-type Statement for Rosetta II

Statement	Item	RAO's Canonical Factor 1	Factor 2	Factor 3
Important to me	5		.79847	
Useful to me	6		.64310	
Made me think	3	(.39509)		
Enjoyable	1	.65459		
I recommend it	8	.78068		
Too long	4	.61144		
Difficult for me	2			.56585
Already knew content	7	-.70798		
Eigenvalue		2.76602	1.66795	1.13810
Percent of variance		34.6	20.8	14.2

the information provided by students was codeable as attitudinal (89%), with 14% in the cognitive area, and 26.5% logistic. Reasons such as "I liked it," "It was enjoyable," and "interesting," were the predominant attitudinal reasons. In the cognitive area, two kinds of statements were common: a general comment of wanting to learn (about two-thirds of those giving a cognitive comment), and a comment of wanting to learn with a specific referent. All logistic comments were categorized as "doing."

Students were asked to complete the statement "The most important thing I learned was ______" for the activity. Student reponses were coded into seven categories, as shown in Table 67. Most student responses were specific; for example, learning how to code, to decode, or to write. Two kinds of more general responses were provided, however. Twenty-one percent of the students doing the activity noted that different people communicate in different ways, or with similar statements; and 4.4% of the students stated that they learned to concentrate, or to use logic.

TABLE 67
Frequencies of Student Responses to the Statement, "The Most Important Thing I Learned Was ______" (N=136)

Categories of Student Response	Student Responses	
	Frequency	Percent
How to decode/what different shapes mean	52	38.2
About early people	17	12.5
Different people communicate in different ways	29	21.3
That there really was a Rosetta stone	5	3.7
To concentrate/use logic	6	4.4
How to write in Nomo, Skribly, or Wosak	6	4.4
No response/not codeable	21	15.4
Totals	136	99.9

The "comment" section of the "Activity Evaluation Form" was used by 30.9% of the students choosing the activity. The most common comment (16%) was that the activity was interesting, fun, or enjoyable. Eight percent of the students commented that the activity was too long, too hard, or boring.

Data have been analyzed from the "Activity Evaluation Forms" completed by 136 students who did the activity Rosetta II. These data indicate that students did make a discriminating evaluation of the activity. Rosetta II was a highly chosen activity (25% of test class students chose the activity). Student satisfaction with the activity, as defined by the description of Factor 1, was high. Personal relevance of the activity was more modest, but still positive, and the activity was not perceived as being too difficult. Responses to sentence completion problems were consistent with the results of the factor analysis of eight Likert-type questions.

The example provided by Rosetta II in response to the statement "I chose this activity because ______ " was typical of activities in the module. For the seven activities in the cluster "Knowing about the Past," 80-85% of the responses to this item were attitudinal. In the activity "Patterns in Your Past," cognitive responses amounted to 48% as compared to 57% for attitudinal. The next highest response for cognitive was 22% as compared to 85% attitudinal for "Counting with Carbon."

To secure additional data regarding the capability of students to make useful judgments about curriculum materials, two items from the Likert scale of the "Activity Evaluation Form" were analyzed: "The activity made me think" and "The activity was enjoyable." The percentage of students who agreed or strongly agreed with the statement "The activity made me think" was added together and divided by the sum of the percent marking "disagree" and "strongly disagree" to arrive at a Think Index. An Enjoyable Index was computed in the same manner. All Enjoyable Indexes for the twenty-one activities in Group I were positive. That is, each Enjoyable Index was greater than 1.0. All but one of the Think Indexes was positive. The only one being lower than 1.0 was that for the activity "Hot Spit," which was, incidentally, one of the least chosen activities. To further reduce the data, a Composite Index was calculated, dividing the Enjoyable Index by the think Index to achieve the indices shown in Figure 20, page 124. These composite indexes are for each of the twenty-one activities in the four clusters of the Group I activities for the KNOWING module. Each activity was rated only by those students who chose the activity. Few activities that were rated by students as making them think were equally enjoyable. This relationship would be reflected by a composite index of 1.0. Note on the histogram that such is the case for relatively few activities. What results is a ratio of "enjoyable" to "think" from about 32 to 1 at the greatest magnitude down to 1.2 to 1, and then with "think" larger, an index ranging from .88 to .32.

Asterisks have been placed beside the number of the activities that were the most highly chosen, that is, were chosen by 19% or more of the 538 students in the study. Note that these highly chosen activites spread across this display of the Composite Indexes. Several interpretations of this information are possible. First, of course, the Enjoyable Index displays a much greater range than the Think Index. Students marked "strongly agree" more frequently for "The activity was enjoyable" than they did for the statement "The activity made me think." This distribution of the Composite Index relates to the previous analysis that dealt with comments by students when they were asked why they chose the activity. There, as is found here, affect predominates over cognition.

Could students make the distinction between think and enjoy? In looking over the complete data for KNOWING, a Composite Index of .19, indicating a high "think" to "enjoy" ratio, was found for "Heavenly Motion," an activity done by about 9% of the students. This activity was complex, with photographs of the comet Kohouteck and data on the comet's position at different times. Students are led to discover the "equal areas in equal time" pattern of motion first recognized by

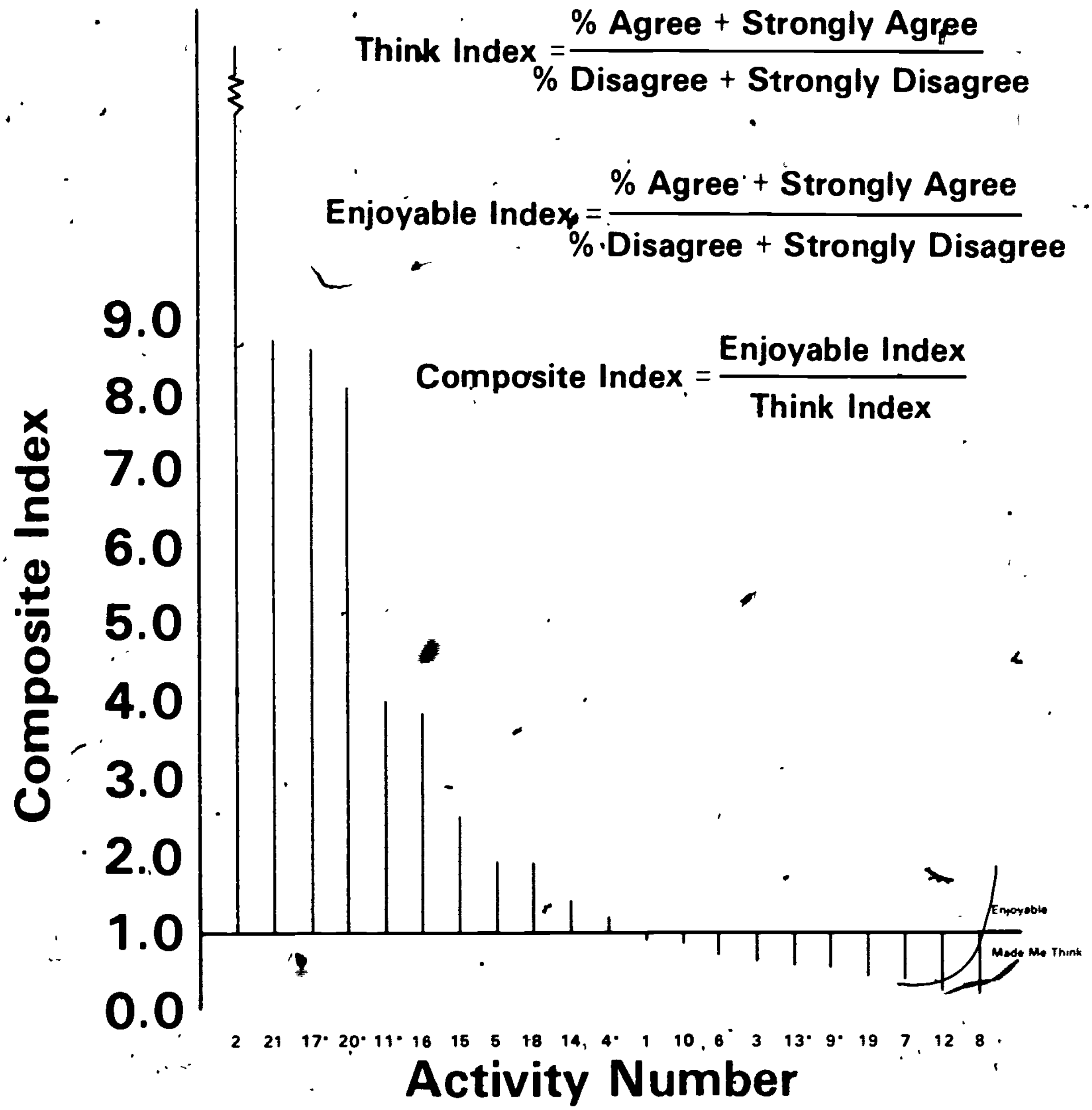


Figure 20. Rank order distribution of twenty-one activities from KNOWING on the composite index.

Johannas Keppler. The activity is indeed complex and does require students to think.

The ability of some students to make the "think-enjoy" distinction is affirmative. The finding that the most highly and least highly chosen activities are found both in the "enjoyable" composite group and the "made me think" composite group also indicate that students were not polarized to choose the activities where "enjoyable" predominates highly over "think."

What Were Students' Attitudes Toward Science Courses?

The "Science Questionnaire," as previously discussed, was administered to a 50% random sample of KNOWING students at the end of the field test period.

The conceptual structure of four subscales on both administrations of the semantic differential instruments was confirmed. The scale scores and scale analysis (Scott, 1968) results are shown in Table 68. Reliabilities (r=Cronbach's alpha) are satisfactory and the homogeneity ratios (HR) confirm the homogeneity of the subscales. Mean scores show that students rated both their regular science course and Human Sciences positively. However, the standard deviations indicate that students varied more on their ratings of regular science than they did on their ratings of Human Sciences. The subscales on both administrations of the semantic differential instrument had high intercorrelations, but correlations were low between the Human Sciences subscales and the 8th Grade Science subscales, as shown in Table 69, page 126.

TABLE 68
Scale Analysis of the Four Subscales of the Semantic Differential Instruments by Course Type

Course	Subscale	n	r	HR	M	S.D.
HUMAN SCIENCES	Evaluation	268	.84	.51	27.88	5.49
	Value	268	.80	.45	27.20	5.46
	Activity	268	.72	.34	26.55	5.46
	Interest	268	.84	.64	16.25	4.15
8th GRADE SCIENCE	Evaluation	268	.91	.67	22.09	7.91
	Value	268	.87	.56	24.52	7.37
	Activity	268	.82	.49	20.35	7.74
	Interest	268	.88	.71	12.33	5.40

Note: Means greater than 20.00 reflect positive values on all scales except Interest (greater than 12.00).

Students rated both courses positively on each of the four subscales of the instruments. They rated the Human Sciences course significantly higher than their eighth-grade science course on each of the four subscales. Results of the t-test for paired samples is shown in Table 70, page 127.

Sex differences were examined by using the one-way ANOVA statistic in the Statistical Package for the Social Sciences. Girls' attitudes toward the Human Sciences course, as expressed on each of the four subscales, were significantly higher than those of boys. There were no significant differences (p=.05) between boys and girls on three of the four scales for the 8th-Grade Science Course, but girls had lower mean scores on all four subscales. Results of the one-way ANOVA are displayed in Figure 21, page 127.

TABLE 69

Correlation Matrices for the Two Semantic Differential Subscale Sets

Course	Subscale	Human Sciences 1	Human Sciences 2	Human Sciences 3	Human Sciences 4	8th Grade Science 1	8th Grade Science 2	8th Grade Science 3	8th Grade Science 4
	1		.734	.705	.827	.229	.294	.087	.167
HUMAN	2			.661	.703	.210	.234	.092	.160
SCIENCES	3				.768	.198	.199	.117	.172
	4					.199	.187	.016	.162
8th	1						.767	.772	.867
GRADE	2							.684	.761
SCIENCE	3								.822
	4								

Subscales: 1 = evaluation; 2 = value; 3 = activity; 4 = interest

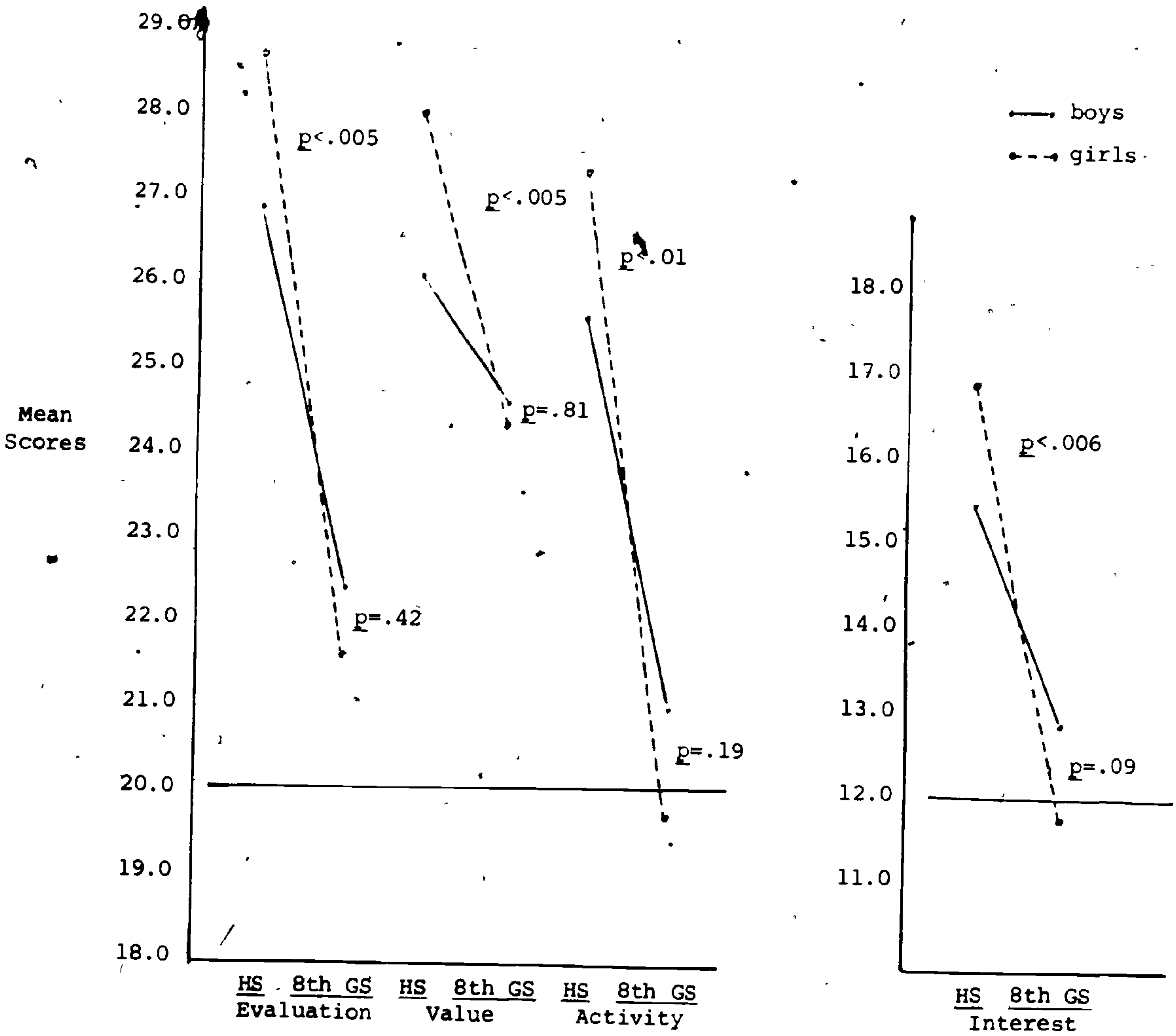
Results from the analysis of the two semantic differential sets from the "Science Questionnaire" indicate that students in KNOWING test classes rated Human Sciences more positive than they rated their regular science program prior to changing to the KNOWING module. This higher rating was found on each of four subscales of the semantic differential instruments. Not only were attitudes toward Human Sciences more positive than for the regular science program, but girls were significantly more positive toward Human Sciences than boys on all four subscales, and significantly more negative toward regular science than boys on one subscale, interest. The trends in the other three subscales were similar.

This finding gives a positive assessment of one of the goals of Human Sciences: to develop and maintain positive attitudes toward science courses. The finding, after eight to ten weeks of using the KNOWING activities, is similar to the findings at the end of the three-year field test of Human Sciences (Robinson, 1980), giving additional support to the effect of the Human Sciences Program on student attitudes.

TABLE 70

Comparison of Mean Scores of Student Attitudes Toward Human Sciences and Eighth-Grade Science on Four Subscales of the "Science Questionnaire"

Subscale	M	S.D.	M	S.D.	N	t	p
Evaluation	27.88	5.50	22.10	7.93	267	5.54	.001
Value	27.21	5.47	24.52	7.39	267	2.72	.005
Activity	26.56	5.47	20.36	7.75	267	5.66	.001
Interest	16.25	4.16	12.33	5.41	267	5.12	.001



Note: The horizontal line is the neutral mean score on the subscales.

Figure 21. A comparison of mean scores of boys and girls on four attitude subscales toward two science courses, N=268.

Summary

The field test of the KNOWING module showed that students new to Human Sciences could be successful with a single module. Half the teachers were not given any orientation to the module but had experienced a workshop in a nearby college or university. Their success with the module showed that with minimum orientation, teachers who volunteered to teach this innovative program could master it.

The students in field test classes had similar logical competence to eighth graders in the three-year field test classes. They responded to the program with attitudes similar to those produced with students who studied Human Sciences in the three-year field test.

CHAPTER 8

ADULT EVALUATION OF HUMAN SCIENCES

The first test plan for Human Sciences included activities to involve a variety of adults in the evaluation process. Community seminars were held in the spring of 1973 at the seven proposed test sites. The seminars were designed to inform school and community personnel about Human Sciences and to engage them in participating in and evaluating the program. The community seminars were held on a Saturday with parents of students who would be in test classes, potential test teachers, building and district administrators, school board members, and science and social studies educators from a nearby college or university.

The seminars were held as planned, but funding restrictions reduced staff work on maintaining adult involvement. The seminars were evaluated and raw data were circulated to staff members. No formal summary of these data has been prepared. In addition to informal evaluation obtained during two site visits in 1973-74, two formal evaluation activities were conducted in 1973-74. Parental evaluation of Human Sciences was solicited through the use of a Parent Report Form, and evaluation data from parents, teachers, administrators, and students was obtained through a series of instruments used in May, 1974.

On the first site visit, an "open house" for Human Sciences was held at each test site. At least one staff member was present and submitted written reports of the visitation.

An external reviewer was retained in 1976 to review seven Human Sciences modules from a developmental perspective to provide information to the staff for revision of the module material. The full report is in EP 7610-05. Following the field testing of all Human Sciences modules, and to prepare for revisions of the modules and activities a content and public review was conducted for selected activities from all modules. The complete results of the content review conferences is presented in EP 7704-19; that of the public review conference in EP 7704-18.

Parental Evaluation

In late October, 1973, Parent Report Forms (see Appendix XX) were sent in classroom quantities to each field test teacher. The forms were distributed in Human Sciences classes to be carried home by students and returned by mail to the Human Sciences project. Envelopes with postage were provided for responses. Responses were received from only four of the seven test sites. The seventy-six forms received represented an 18-percent response from the four sites; an 11 percent response from the total test group.

The responses represent a probable bias toward positive responses since negative responders tend to respond in fewer numbers. Only one of the seventy-six parents responding indicated that her or his child had failed to mention Human Sciences (Item 1, Parent Report Form).

A content analysis was conducted on Item 2 of the form to determine to what extent children reported positive, negative, or neutral information as perceived by parents. Written statements were categorized as being descriptive or comparative. Descriptive statements mentioned only Human Sciences. Comparative statements had to mention Human Sciences in relation to some other course or activity. Descriptive comments were categorized into four subgroups: positive, negative, neutral, and discriminative (or mixed). Comparative comments were coded into one or more of the three categories of descripters shown in Table 71, page 131.

Responses to Item 2 were coded, either in a single category or into several categories. The majority of responses to Item 2 were coded into one category. The maximum number of codings for any one response was five.

Positive codings outnumbered negative codings by over 30 to 1. Positive codings were about three times more numerous than neutral codings. The response to Item 2 is interpreted as reflecting a positive student-to-parent-to-developer report regarding Human Sciences.

There were seven questions seeking parental responses on the Parent Report Form. Only Item 2 has been formally evaluated. A reading of all of the responses to the eight items on the Parent Report Form supports the positive response interpretation of students toward Human Sciences, as seen through the reports of parents. When parents expressed their personal opinions, they were also positive in similar proportions. Complete responses to the Parent Report Form for two schools are presented in Appendix YY.

Informal comments from parents were documented by observers at one test site. In reporting about an "open house" for Human Sciences parents, one staff observer found parents concerned about how their children were achieving academically in comparison to their peers. Their other concern seemed to be with behavior: were their children causing problems or were they well behaved? The following is a selection from the observer report.

"We're hearing more about this program than we've heard about school since he started back in kindergarten." This comment was similarly echoed by many parents throughout Open House Night at Test Site 2.

TABLE 71

Response of Parents to Item 2 of the Parent Report Fo[illegible] your child has told you about Human Sciences, what were some [illegible] or her comments?"

	Schools				Totals		
	2	3	4	5	Posi-tive	Nega-tive	Neu-tral
	N=16	N=34	N=8	N=18			
Descriptive							
Positive							
Liked/enjoyed/fun/interesting							
Unspecified	4	10	1	3	18		
Specified							
Choice	1	2			3		
Close to life experience		1			1		
Personal responsibility		2			2		
Particular activities/tasks	5	7	6	3	21		
The vanity		2			2		
Working with others		1	1		2		
Special test group	2	1			3		
Way to learn science		3	1		4		
Understands better/learns a lot	1	2			3		
Easier	1				1		
Total Positive, Descriptive	14	31	9	6	60		
Negative							
Disliked/boring							
Unspecified		2					
Total Negative, Descriptive	0	2	0	0		2	
Neutral	1	3		2			6
Specified				6			6
Total Neutral, Descriptive	1	3	0	8			12
Discriminative (some interesting/some boring)	3	4		4			
Total Discriminative	3	4	0	4			11
Comparative							
Likes better than other courses/science	1	3		1			
Likes less than other courses/science							
Likes about the same as other courses/science		1					1
Total Positive, Comparative	1	3	0	1	5		
Total Negative, Comparative	0	0	0	0		0	
Total Neutral, Comparative	0	1	0	0			1
Column Totals	19	44	9	19	65	2	24

"Parents are finding themselves actively involved in the Human Sciences program as sources of information, as listeners, and as participants in discussions. One parent was intrigued by her daughter's thoughts on 'Boy Or Girl?' Apparently they both had gotten into quite a discussion over a common adult response of, 'A boy, to carry on the family name.' Another parent participating in the same activity found himself giving considerable thought to the question, and was interested to see the responses of other persons interviewed by his son. In accordance with the activity, 'Jabberock,' one student had his family conversing freely and frequently with an adopted family of rocks. It is apparent from these comments and others, that students are sharing many of their Human Sciences experiences with their parents. Reactions to the materials are being received at home."

"There seem to be several parental concerns that need to be attended to at this point. First, parents are interested in finding out more about the Human Sciences program. What is the Human Sciences? Who's involved in its development? How is it different?, etc. Certainly, a test teacher's notion of a questions and answer parent pamphlet would provide a basic overview of the program, answering many and most of these questions. These could be sent home to all parents via their children prior to the Community Seminar meeting.

Second, those parents at Test Site 2 who attended the Community Seminar last spring are now eager to be involved in the program. It is important that interested parents not be left floundering for what they can do to be of help. One area that these parents could be of invaluable service would be in recording student comments made at home. These records could become part of the student's longitudinal files. It would seem appropriate that these parents encourage other parents of students in the Human Sciences program to note student reactions brought home. Indeed, they could be given the opportunity to set up some system of collecting these comments on a regular basis (i.e., possibly after each module). Another area of parental involvement might be at the coffees at the school. Over the next three years it would be exciting to see parents as well as teachers answering parent questions. The needs of each school system are varied, but in each where there are interested and concerned parents, opportunities for meaningful involvement should be made available.

Supporting the interest and participation of parents in the Human Sciences Program was a source of frustration to the staff. Parents could have become a valuable source of ideas for program improvement, but the time and energies of the staff were drawn back to development. One information sheet about Human Sciences was prepared and sent to parents. One test school held a curriculum fair in the spring of 1974 in which Human Sciences was a part. Other parent events were also accomplished at the school level. The final contact with parents was made in another evaluation activity in May, 1974.

Evaluating Innovative Science Curricula

In the spring of 1974 the BSCS Human Sciences Program was funded to establish four centers in different regions of the United States in order to test a model for the dissemination and implementation of innovative science curriculum materials. In preparation for this addition to the project, instruments were developed for gathering data on adult and student perceptions and opinions about Human Sciences were developed. The seven field test sites were visited by staff members in May 1974 to test the instruments by gathering data from teachers, administrators, parents, and students.

One instrument, "Developing Human Sciences," was mailed to teachers to enable students to use it as an activity. Results from this instrument could not be quantified reliably for reporting. A second instrument, "Evaluation of Human Sciences" (see Appendix ZZ), was mailed to teachers and administrators. The other instruments (see Table 72) were carried by staff members on a site visit and administered personally by them.

TABLE 72
Instruments Administered in May, 1974 to Selected Individuals at the Seven Human Sciences Test Sites and the Groups to Whom the Instruments Were Administered

Instrument	Teachers	Principal	Parents	Students	Central Office
What is your opinion of Human Sciences?	X	X	X	X	
Goals of education	X	X	X		X
Evaluation of Human Sciences					
1. Amount of change	X	X			X
2. Kinds of change	X	X			X
3. Communications	X	X			
Activity "Developing Human Sciences"				X	

Each instrument was designed to gather particular data from one or more target groups. Data gathered on this first administration were to test the instruments and, if they were found to be useful, to establish base-line data to be used as a comparison group when the instruments were used in subsequent years as part of the evaluation of the dissemination-implementation model. The findings from each of the instruments will be presented independently in subsequent sections.

What is Your Opinion of Human Sciences?

This twenty-five-item Likert scale (see Appendix AAA) was developed from comments or answers to questions asked about Human Sciences during interviews or conversations with teachers, principals, students, and parents. The instrument was administered to all test teachers, all principals, a random sample of parents, and a random sample of students, at each of the seven test sites. There were 87 boys and 82 girls in the student group. Table 73 shows the categories and numbers of individuals who provided completed _What is Your Opinion of Human Sciences?_ questionnaires.

TABLE 73
Types and Numbers of Individuals from Whom "What is Your Opinion of Human Sciences?" Data Were Obtained

School	Students	Teachers	Number of Administrators	Parents	Total
1	18	3	1	3	25
2	28	3	1	10	42
3	30	3	1	4	38
4	27	2	2*	9	40
5	24	2	1**	13	40
6	20	2	2*	20	44
7	22	3	1**	14	40
	169	18	9	73	269

*Includes one principal and one central office administrator
**Assistant principal

Data from all respondents were pooled for preliminary analysis. Further examination by students, parents, and educators indicated that the response patterns of the adults were so similar that they could be combined as one group.

The responses to each statement were weighted as follows for scoring: strongly agree, 5; agree, 4; neutral or uncertain, 3; disagree, 2; and strongly disagree, 1. Examination of mean item scores of adults as compared to students suggested that these two groups were responding to the items diffferently. To test this hypothesis, a multiple discriminant analysis was made to determine which items discriminated between the two groups most effectively. One discriminant function used all items as significant discriminators (p .0000). This analysis was not usable, however, because the test for equality of the groups' covariance matrices resulted in a significant difference (F=1.96, p .0000), indicating that the relationships between the groups were not linear. A t-test of the difference between means indicated that the mean scores of students and adults on the twenty-five items were not different (t=1.50, t=1.96). However, a one-way analysis of variance test for equality of group means for each discriminating variable indicated that the groups differed significantly on sixteen of the twenty-five items (p .02). Results of the analysis are shown on Table 74.

TABLE 74

Test for Differences Between Students and Adults on Each Item of "What Is Your Opinion of Human Sciences?"

Wilk's lambda (U-statistic) and univariate F-ratio with 1 and 267 degrees of freedom			
Variable	Wilk's Lambda	F	Significance
V1	.99781	.5858	.4447
V2	.96696	9.122	.0028
V3	.97104	7.962	.0051
V4	.86394	42.05	0
V5	.99714	.7663	.3822
V6	.99768	.6206	.4315
V7	.99962	.1013	.7505
V8	.97280	7.464	.0067
V9	.90134	29.23	0
V10	.98003	5.441	.0204
V11	.97794	6.023	.0148
V12	.96213	10.51	.0013
V13	.99903	.2591	.6221
V14	.97928	5.650	.0182
V15	.77821	76.09	0
V16	.95583	12.34	.0005
V17	.99973	.7180E-01	.7889
V18	.87562	37.93	0
V19	.97707	6.267	.0129
V20	.95109	13.73	.0003
V21	.99702	.7968	.3729
V22	.99190	2.182	.1408
V23	.82265	57.56	0
V24	.99889	.2970	.5863
V25	.95986	11.17	.0010

Of greatest interest for the formative evaluation of Level I was the way in which students and adults differed in their responses to the items on the questionnaire. To investigate these differences, Alpha factor analysis of student responses and adult responses were separately computed. Alpha factor analysis was chosen because this kind of analysis emphasizes the maximum generalizability in the coefficient alpha sense (Harman, 1976, p. 231). This is due to the determination of common factors which have maximum correlation with the corresponding universe common factors, a psychometric inference.

The adult group included seventy-three parents, seventeen teachers, and seven administrators. The students (N=169) were a random sample from each of the seven test site classes. These students were completing their first year in Human Sciences test classes. The parents group was a sample of parents of students in the Human Sciences classrooms. The teachers were seventeen of the nineteen Human Sciences teachers and the administrators were building level administrators who were most knowledgeable about Human Sciences. The interpretations will use the factors identified from the student analysis. In one case, both students and adult factor structures will be presented and discussed because these factors overlapped. Both factor matrices are reproduced in Appendix BBB.

Ten factors with eigenvalues of 1.0000 or greater were identified from the student data; eight factors were computed from the adult data. Factor 1 included three items in the student analysis and was conceptually similar to Factor 2 with six items in the adult analysis. Both factors are conceptualized as attitudes about student learning in Human Sciences. Table 75 presents the items and relevant statistical data for students' and adults' responses on the three items common to both factors (items 1, 8, and 11). The mean of the item mean scores is similar for both groups, with both disagreeing with the negatively worded items. The adults were significantly more negative than students in two of the three items. Since all statements in this factor were negatively worded, both students and adults were positive in their attitudes toward what students were learning in Human Sciences.

TABLE 75

Factor 1, Student Learning, of "What is Your Opinion of Human Sciences?" Showing the Items, Factor Loadings, and Means for Student Responses; Means for Adult Responses; and the Wilke's Lambda Significance Test for Differences Between Means

Items	Student Mean Scores	Factor Loading	Adult Mean Scores	Desired Response	Significance of any Differences
1. Students in regular science classes learn more than students in Human Sciences classes (-)	2.65	.48	2.54	3.00	n.s.
8. Students don't learn as many important science facts in Human Sciences as they would in regular science courses (-)	2.90	.64	2.47	3.00	.007
11. Students in Human Sciences classes aren't learning things that will help them in their high school science courses (-)	2.55	.76	2.20	3.00	.015
Mean of the item mean scores	2.70	-	2.40	3.00	-

Table 76 shows Factor 2, Student learning, from the adult response factor analysis. Although there were three more items than in the comparable student factor, responses of both adults and students were similar and positive about student learning. Only on item 16 did students and adults disagree on their responses, with students disagreeing and adults agreeing with the item. This difference was statistically significant with responses in different directions from a neutral score. By converting item mean scores to positive equivalents, the responses of

TABLE 76

Factor 2, Student Learning, of "What Is Your Opinion of Human Sciences?" Showing the Items and Means for Adult Responses, Means for Student Responses, and the Wilkes Lambda Significance Test for Differences Between Means

Items	Student Mean Scores	Factor Loading	Adult Mean Scores	Desired Response	Significance of any Differences
1. Students in regular science classes learn more than students in Human Sciences classes (-)	(3.46) 2.54	.49	(3.35) 2.65	(3.00) 3.00	n.s.
8. Students don't learn as many important science facts in Human Sciences as they would in regular science courses (-)	(3.53) 2.47	.45	(3.10) 2.90	(3.00) 3.00	.007
11. Students in Human Sciences classes aren't learning things that will help them in their high school science courses (-)	(3.80) 2.20	.49	(3.45) 2.55	(3.00) 3.00	.015
16. Human Sciences does not give enough attention to important science activities such as dissecting animals (-)	(3.41) 2.59	.64	(2.83) 3.17	(3.00) 3.00	.000
19. More attention to science topics such as weather or chemistry should be included in Human Sciences (-)	(2.82) 3.18	.52	(2.45) 3.55	(3.00) 3.00	.013
22. Students in Human Sciences have more opportunity to learn from each other than in other science classes (+)	3.86	-.54	4.05	3.00	n.s.
Mean of the item mean scores	3.48	-	3.20	3.00	-

Note: Mean scores in parentheses are transformed to their positive reciprocal in order to make all items directionally positive for purposes of averaging.

both students and adults were positive regarding student learning as measured by the items in this factor. Adults were significantly more positive in four of the six items. Students were negative on one item.

Student Factor 2 has no comparable factor in the adult response pattern. It does include one item in adult Factor 2 and one item not included in any adult factor. This factor, with two items (see

Table 77), is conceptualized as "Physical science content for sixth-graders. Students were in significantly stronger agreement with Item 19 than adults and adults were in significantly stronger disagreement with Item 20 than were students. This finding was counter to the guidelines, Life Sciences for the Middle School and counter to the curriculum framework for Human Sciences, both of which proposed a life science program for the middle school/junior high school student.

TABLE 77
Factor 2, Physical Science Content for Sixth-Graders, of "What is Your Opinion of Human Sciences?" Showing Items, Means, and Factor Loadings for Student Responses; Means for Adult Responses; and the Wilke's Lambda Significance Test for Differences Between Means

Items	Student Mean Scores	Factor Loading	Adult Mean Scores	Desired Response	Significance of any Differences
19. More attention to science topics such as weather or chemistry should be included in Human Sciences (-)	3.55	.59	3.18	3.00	.013
20. Most students in the sixth grade are not really interested in the study of topics such as weather or chemistry (+)	(3.13) 2.87	.82	(3.68) 2.32	(3.00) 3.00	.000
Mean of the item mean scores	3.34	-	3.49	3.00	-

Note: Mean scores in parentheses are transformed to their positive reciprocal in order to make all items directionally positive for purposes of averaging.

Factor 3 from the student data is conceptualized as measuring attitudes about student responsibility. The two items in this factor were responded to with strong agreement among students and adults for the factor (see Table 78, page 139). As with the previous factor, adult responses did not place items 2 and 6 in the same factor. These two important characteristics of Human Sciences were viewed positively by both students and adults.

Student Factor 4 is conceptualized as a "Teaching Human Sciences" factor (see Table 79, page 139). Both students and adults were in agreement on items 7 and 23. However, students felt significantly stronger about their disagreement with item 23 than did adults. There was a signifcant difference in responses to the problem of getting materials needed to do activities. Since students were much closer to the classroom situation, their view--agreement with item 4--indicates that the management of scarce resources in overpopulated classroom environments was not resolved at the end of Level I testing. The adult response can be interpreted as a positive attitude toward the program, but a naivete with regard to the details of classroom management. This explanation is consistent with the signifcant difference in the magnitude of the responses of adults and students.

TABLE 78

Factor 3, Student Responsibility, of "What Is Your Opinion of Human Sciences?" Showing the Items, Means, and Factor Loadings for Student Responses; Means for Adult Responses; and the Wilke's Lambda Significance Test for Differences Between Means

Items	Student Mean Scores	Factor Loading	Adult Mean Scores	Desired Response	Significance of any Differences
2. Teachers of Human Sciences rely on students to care for the plants and animals (+)	4.13	.53	3.80	3.00	.003
6. Human Sciences helps students become responsible for their own learning (+)	4.02	.74	4.11	3.00	n.s.
Mean of the item mean scores	4.08		3.96	3.00	-

Student factor 5, conceptualized as "Student Self-Direction," indicated agreement of students and adults with the self-directive charcteristic of Human Sciences (see Table 80, page 140) and with this context, both perceive teachers as being positive toward Human Sciences.

TABLE 79

Factor 4, Teaching Human Sciences, of "What Is Your Opinion of Human Sciences?" Showing the Items, Means, and Factor Loadings for Student Responses; Means for Adult Responses; and the Wilke's Lambda Significance Test for Differences Between Means

Items	Student Mean Scores	Factor Loading	Adult Mean Scores	Desired Response	Significance of any Differences
4. Students often have difficulty getting materials needed to do Human Sciences activities	(2.71) 3.29		(3.64) 2.36	(3.00) 3.00	.000
7. It is harder for teachers to grade students in Human Sciences than in other courses	3.14	.58	3.19	3.00	n.s.
23. Teachers should tell students what activities they should do in Human Sciences (-)	(4.59) 1.41		(3.66) 2.34	(3.00) 3.00	.000
Mean of the item mean scores	3.48		3.06	3.00	-

Note: Mean scores in parentheses are transformed to their positive reciprocal in order to make all items directionally positive for purposes of averaging.

TABLE 80

Factor 5, Student Self-direction, of "What is Your Opinion of Human Sciences?" Showing Items, Means, and Factor Loadings for Student Responses; Means for Adult Responses; and the Wilke's Lambda Significance Test for Differences Between Means

Items	Student Mean Scores	Factor Loading	Adult Mean Scores	Desired Response	Significance of any Differences
21. Human Sciences allows students to look for answers to questions that they decide are important to answer	3.70	.45	3.79	3.00	n.s.
12. Teachers of the Human Sciences classes would rather teach the program than regular science (+)	3.72		3.35	3.00	.001
15. Students in the sixth grade are old enough to know what they want to learn in Human Sciences (+)	4.34		3.34		.000
Mean of the item mean scores	3.88	-	3.49	3.00	-

Students were significantly more positive about two of the three items than were adults.

"Wasting time" is the conceptualization of student Factor 6, consisting of two negatively worded items (see Table 81). The desired response was disagreement with both of these items. Both students and

TABLE 81

Factor 6, Wasting Time, of "What Is Your Opinion of Human Sciences?" Showing Items, Means, and Factor Loadings for Student Responses; Means for Adult Responses; and the Wilke's Lambda Significance Test for Differences Between Means

Items	Student Mean Scores	Factor Loading	Adult Mean Scores	Desired Response	Significance of any Differences
3. During Human Sciences classes, students spend their time doing things that are not important	2.41	.59	2.00	3.00	.005
10. Many students in Human Sciences classes waste their time when they get to choose what they do	2.93	.57	2.58	3.00	.020
Mean of the item mean scores	2.67	-	2.29	3.00	-

adults who did agree with the "wasting time" concept expressed, indicating that they would not characterize Human Sciences classes in those terms. In this instance, however, students were significantly different from adults on both items, tending to be closer to "undecided" in their mean response.

The seventh student factor includes two statements and is conceptualized as "Parental Involvement." Both statements in this factor are positive in wording and the desired response, agreement, was obtained with both students and adults (see Table 82).

Student Factor 8 consists of two items and is characterized as a "teacher control" factor. Students disagreed with the two statements, each stating a teacher control practice that contrasted markedly with the desire practices of teachers in Human Sciences (see Table 83, page 142). Item 17 (not shown) had its highest loading (.38) on this factor but since it was split between factors eight and nine, and it has no conceptual relevance to the other two items in the factor, it was excluded from the factor. Responses to Factor 8 indicate that both students and adults disagreed with teacher control of activity selection in Human Sciences. For both items, students were significantly in greater disagreement with the statements than were adults and were, therefore, less favorable toward teacher selection of students' activities than were adults. Factor 9 is conceptually related to Factor 8 in being concerned with teacher control. Factor 9 is conceptualized as "teacher control of content." Both students and adults again disagreed with teachers choosing content through activity choice for the Human Sciences curriculum (see Table 84, page 142). Students again were stronger in their opinions on this issue than were the adults.

Factor 10 is conceptualized as a "self-direction" factor (see Table 85, page 143). Both students and adults were in agreement with

TABLE 82
Factor 7, Parental Involvement, of "What Is Your Opinion of Human Sciences?" Showing Items, Means and Factor Loadings for Student Responses; Means for Adult Responses; and the Wilke's Lambda Significance Test for Differences Between Means

Items	Student Mean Scores	Factor Loading	Adult Mean Scores	Desired Response	Significance of any Differences
13. Most parents of students in the Human Sciences classes don't know much about what their sons or daughters are learning	3.25	.74	3.32	3.00	n.s.
24. Most of the parents of students in Human Sciences think that their sons or daughters really enjoy Human Sciences	3.65	.40	3.72	3.00	n.s.
Mean of the item mean scores	3.45	-	3.52	3.00	-

TABLE 83

Factor 8, Teacher Control, of "What Is Your Opinion of Human Sciences?" Showing Items, Means, and Factor Loadings for Student Responses; Means for Adult Responses; and the Wilke's Lamda Significance Test for Differences Between Means

Items	Student Mean Scores	Factor Loading	Adult Mean Scores	Desired Response	Significance of any Differences
18. Teachers should decide what Human Sciences activities should be done and teach one activity at a time to the whole class (-)	1.42	.66	2.18	3.00	.000
23. Teachers should tell students what activities they should do in Human Sciences (-)	1.41	.41	2.34	3.00	.000
Mean of the mean item scores	1.42	-	2.26	3.00	-

the opportunities for student initiative that characterizes Human Sciences. As with the items in Factors 8 and 9, students had significantly stronger opinions than did adults.

Responses from "What is Your Opinion of Human Sciences?" provided data for comparing student attitudes with adult attitudes. Data were analyzed from 169 students and 100 adults. Multiple discriminant analysis showed that adults and students differed significantly in 16 of the 25 items on the instruments. Factor analysis produced 10 student factors and eight adult factors. Twenty-three of the 25 items on the

TABLE 84

Factor 9, Teacher Control of Content, of "What Is Your Opinion of Human Sciences?" Showing Items, Means, and Factor Loadings for Student Responses; Adult Means; and Desired Response Means

Items	Student Mean Scores	Factor Loading	Adult Mean Scores	Desired Response	Significance of any Differences
9. Teachers should pick out the Human Sciences activities that are best for each student (-)	1.56	.57	2.26	3.00	.000
16. Human Sciences does not give enough attention to important science activities such as dissecting animals (-)	3.17	.42	2.59	3.00	.000
Mean of the item mean scores	2.36	-	2.42	3.00	-

questionnaire were included in the factor structure. Four items were complex in that they were included in two different factors. The 10 student factors were used to compare student and adult responses. A central theme of Human Sciences--student responsibility for learning, and student self-direction--was prominent in eight of the 10 student factors. On these factors, both students and adults agreed with the Human Sciences theme--Human Sciences should promote the development of student self-direction and responsibility for learning. On many items students had stronger opinions in this direction than did adults. In comparing Human Sciences with regular science, both students and adults expressed opinions that they were equivalent, but that Human Sciences offered more opportunities in regards to the central theme, as expressed above. In the area of content, both students and adults felt that physical science content was important for sixth-graders. This content was not included in the Level I materials. Students, but not adults, felt that more attention should be given to regular science activities, such as dissection.

TABLE 85

Factor 10, Self-direction, of "What Is Your Opinion of Human Sciences?" Showing Items, Means, and Factor Loadings for Student Responses; Adult Means; and Desired Response Means

Items	Student Mean Scores	Factor Loading	Adult Mean Scores	Desired Response	Significance of any Differences
14. The teachers of Human Sciences let students decide which activities they want to do and how they want to do them	3.89	.62	3.57	3.00	.018
22. Students in Human Sciences have more opportunity to learn from each other than in other science classes	4.05	.46	3.86	3.00	n.s.
Mean of the item mean scores	3.97	-	3.72	3.00	-

At the end of the first year of testing, students and adults who responded to "What is Your Opinion of Human Sciences?" were positive about the program and in agreement with one of its central themes.

"Goals of Education" Priorities

The goals of the Human Sciences Program are more developmentally oriented than the goals of most science curricula in schools (see Chapter 2). In order to determine if the high-priority goals of the Human Sciences staff were similar to or in conflict with the goals of parents, teachers, and administrators, a card sort task ("Goals of

Education") was administered to the four Human Sciences staff members, test teachers, administrators, and parents. Table 86 shows the number of individuals from whom data were obtained. School 3 and School 4 were seriously underrepresented in number of parents.

TABLE 86
Individuals Providing Data from the Card Sort Task, "Goals of Education," N=106

School	Teachers	Administrators	Parents	Total
1	2	2	74	6
2	3	1	10	14
3	3	1	4	8
4	3	3	9	15
5	2	3	14	19
6	2	3	20	25
7	3	1	15	19
TOTAL	18	14	74	106

The theoretical basis of the card sort task was presented by Kohlberg and Meyer (1972). They proposed three distinctive goals of education: developmental, romantic, and cultural transmission. Twelve statements were prepared, five consistent with the developmental goals description, three consistent with the romantic goals description, and four consistent with the cultural transmission goals description (see Appendix CCC for the reading, forms, and goals statements).

Single goal statements were taped on cards and randomly assigned letter identifiers from A to L. Respondents were asked to sort the cards into their own order, from highest to lowest priority, and then to record the letter identifier of the statement on a form provided (see Appendix CCC). Only the column, "Your personal preference" was used. "Goals of Education" was administered in May, 1974 during site visitations.

Table 87, page 145, shows high (rankings 1-5), medium (rankings 6-8) and low (rankings 9-12) categories of the four groups who did the card sort ranking. As seen by the percentage figures, there was considerable consensus among staff, administrators, teachers, and, to a lesser degree, parents, which places a high priority ranking on "developmental" goals of education. Parents differed from the other groups by ranking two of the cultural transmission items higher than any of the other groups. There was also considerable agreement among the four groups as to low priority items with teachers and administrators both making judgments similar to those of the Human Sciences staff.

Several goal statements are of interest in regard to the spread that occurred. Item G--"Education should help a student develop progressively more complex ways of logical thinking and moral reasoning"--was seen as a high priority item by 100% of the staff, 76% of the teachers, 63% of the administrators, but only 48% of the parents. Since this is a major goal of Human Sciences, perhaps more effort needs to be placed on emphasizing this goal. Item K--"Education should help a student develop the ability to reflect upon and make up his or her mind

Table 87. Percentage of Human Sciences staff, administrators, teachers, and parents giving high, medium, and low priority to Twelve Goals of Education Statements.

B, C, D, G, K — DEVELOPMENTAL

Priority	Goal	HSP Staff %	Administrators %	Teachers %	Parents %
HIGH 1-5	B	100	100	95	86
	C	100	87	68	63
	D	100	87	74	60
	G	100	53	79	47
	K	50	67	42	23
MEDIUM 6-8	B			5	10
	C		13	32	29
	D		13	21	26
	G		40	16	29
	K	50	13	37	37
LOW 9-12	B				4
	C				8
	D			5	14
	G		7	5	23
	K		20	21	40

E, H, J — ROMANTIC

Priority	Goal	HSP Staff %	Administrators %	Teachers %	Parents %
HIGH 1-5	E	50	20	5	39
	H		33	58	48
	J			16	15
MEDIUM 6-8	E	50	20	37	26
	H	100	60	37	29
	J	50	33	37	31
LOW 9-12	E		60	58	35
	H		7	5	23
	J	50	67	47	54

A, F, L, I — CULTURAL TRANSMISSION

Priority	Goal	HSP Staff %	Administrators %	Teachers %	Parents %
HIGH 1-5	A		40	47	74
	F				5
	I		13	16	37
	L				3
MEDIUM 6-8	A	25	47	48	15
	F	25	47		21
	I		13	21	36
	L		0	16	16
LOW 9-12	A	75	13	5	11
	F	75	53	100	74
	I	100	74	63	27
	L	100	100	84	81

about controversial social issues"--was seen as a high priority item by 63% of the administrators, 50% of the staff, 41% of the teachers, but only 23% of the parents. It was seen as low priority by 40% of the parents, 24% of the teachers and 13% of the administrators. This finding is in contrast with the often expressed belief that administrators do not want their schools to deal with controversial issues. In Human Sciences trial schools, this item was given low ranking by only 13% of the administrator respondents. The goals of education, as viewed by parents in the seven trial schools, are not markedly different from those of teachers and administrators.

The results of this ranking exercise seem to indicate that teachers, administrators, and parents in Human Sciences trial schools have priorities in educational goals that are generally in keeping with those of the Human Sciences Program. This finding is also supported by findings in Section 4 of "Evaluation of Human Sciences" (next part of this chapter), which shows that on fourteen statements of educational values, conflict between HSP and respondent values was indicated on only seven items. In terms of significant numbers, there was conflict only on items dealing with teacher control and classroom discipline, and transmission of academic science knowledge.

Teacher and Administrator Evaluation

The third part of the group of evaluation instruments that was to be used in future evaluation studies in the Dissemination and Implementation Centers was "Evaluation of Human Sciences." This instrument was designed for responses by teachers and administrators to determine their perception of the kinds of changes required to adapt and to implement the Human Sciences program. The 14 section instrument was based upon a review of the literature on curriculum implementation and diffusion (Hurd, 1972) pointing out the major problems that needed to be resolved if new curricula were to be successfully adopted in schools (see Appendix DDD for a copy of "Evaluation of Human Sciences."

Copies of the instrument were mailed to 18 teachers and 15 administrators in May, 1974. The administrator mailings included only those administrators who were known by BSCS staff and/or test site teachers to be knowledgeable about Human Sciences. To meet this criterion, the administrator must have visited a Human Sciences classroom at least once during the year, to have met with BSCS staff members during site visits for orientation to the program, and to have reviewed Human Sciences materials. The distribution of administrators by school and position is shown in Table 88, page 147).

Complete "Evaluation of Human Sciences" instruments were obtained from all administrators and from 7 of the 18 test teachers. Limited analyses of the data were prepared in a previous report (see Appendix EEE). A brief analysis of the questionnaire data is presented here.

TABLE 88

Distribution of Administrators Who Completed Evaluation of Human Sciences by Administrative Position and Test Site

Administrative Position	Field Test School Number							
	1	2	3	4	5	6	7	Total
Principal	1	1	1	1		1	1	6
Vice Principal/Curriculum Coordinator	1				1	1		3
Science Chairperson				1				1
Science Supervisor				2	1	1		4
Central Office Supervision								
State Dept. of Education	1							1
TOTALS	3	1	1	4	2	3	1	15

The questionnaire was composed of 14 questions, each with a series of subquestions. This analysis will present the findings of the combined teachers and administrators for each of these questions. The following areas of concern were built into the structure of the questionnaire: the amount of change, and the kinds of change necessary since the Human Sciences Program was introduced into the school; and effectiveness and quality of communications about Human Sciences. The data on which these interpretations are based are presented in Appendix FFF.

The first question was concerned with the effect of Human Sciences on five factors in school operations: cost, personnel, space, consumable materials, and equipment needed in comparison to the regular science program in the schools.

Figure 22 presents the composite pattern of responses for each of the five factors in school operations. The total impact of Human

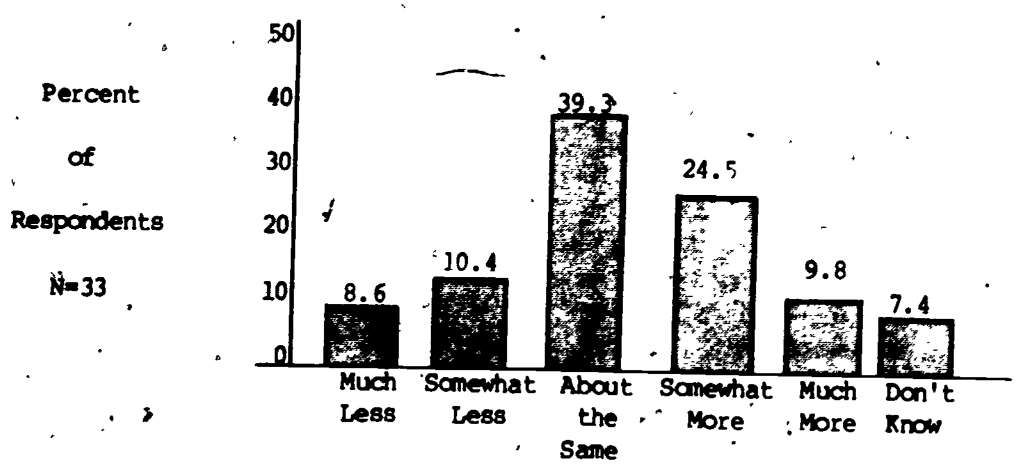


Figure 22. The requirements of Human Sciences as compared with regular science on school operations.

Sciences placed somewhat heavier demand on school operations than did the regular program. Both teachers and administrators agreed that HSP required somewhat more school equipment and consumable supplies than did their regular science program.

There were 11 items within the second question, all statements regarding teaching and teacher behavior (see Appendix FFF). Most of the statements were positive statements related to giving individual attention to students, to their differences, and to their needs. For these items, both teachers and administrators agreed that Human Sciences made these changes easier (see Figure 23). There was a wide divergence of opinion among both teachers and administrators about whether Human Sciences made classroom discipline easier or more difficult (data as shown in Appendix FFF). Another divergence from the general pattern was that related to evaluation, where evaluation was seen to be more difficult in Human Sciences than in regular science classes.

The third question focused on goals and objectives for education, asking if Human Sciences placed about the right emphasis on nine statements of goals and objectives. The general response to the nine statements was that Human Sciences gives about the right emphasis to the stated goals and objectives, but with more responses suggesting not enough emphasis rather than too much. "About right" was the modal response for every statement in this question. The second most frequent response choice was "not quite enough emphasis." The statement with the highest response in this regard by teachers was for "students learn the facts and principles of science disciplines." Administrators marked the "right" amount most frequently for this objective (see Figure 24, page 149).

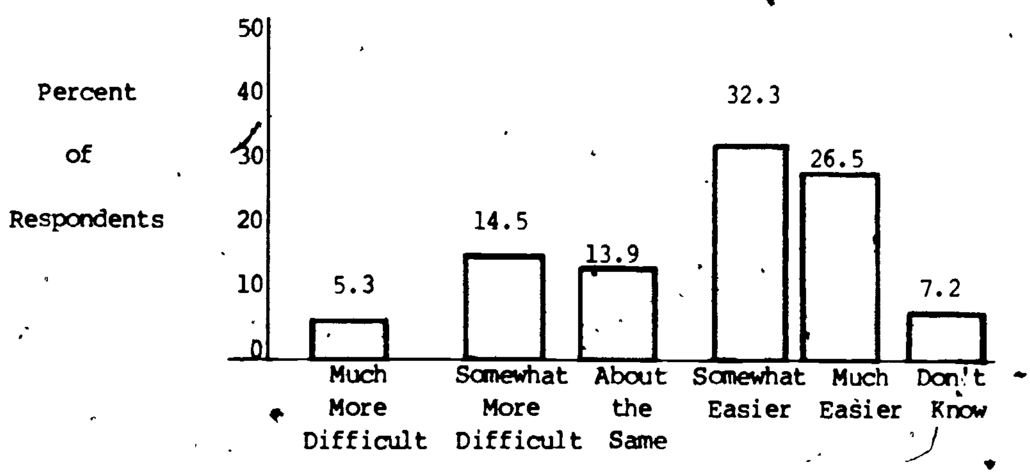


Figure 23. The requirements of Human Sciences as compared with regular science for changes in teaching.

The fourth question asked, "Does the Human Sciences curriculum conflict with or support your educational values?" There were enough comments on questionnaires to raise questions about ambiguity in the question and its relation to the fourteen statements to be evaluated. Analysis of the question in relation to values each respondent circled

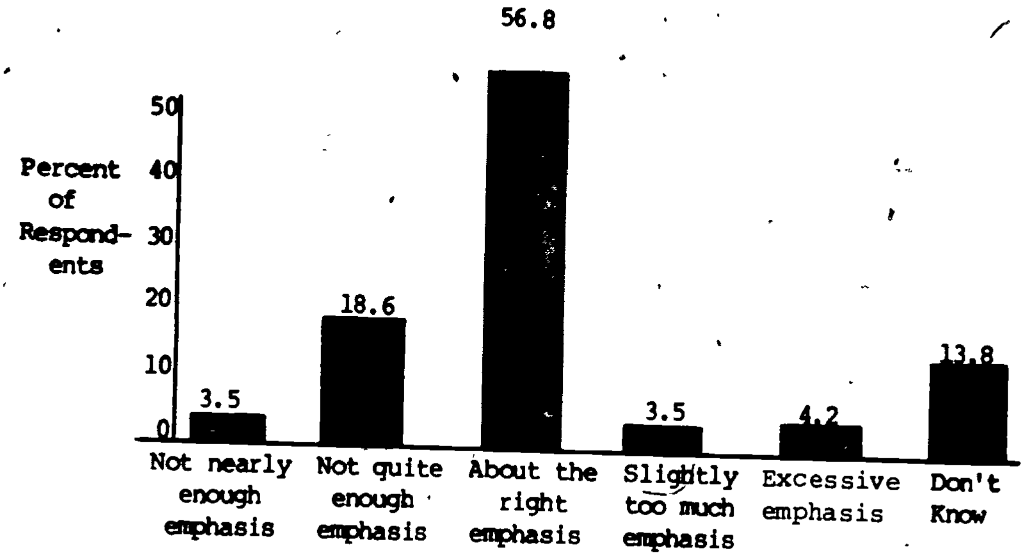


Figure 24. The emphasis Human Sciences gives to nine statements of goals and objectives.

has not been completed. A literal analysis was completed. This analysis ignores the idea of change and merely tabulates the response frequencies marked for each of the fourteen "educational values" statements (see Appendix FFF). Figure 25 presents the summed response frequencies for all statements in question 4. It may be interpreted as indicating that both teachers and administrators feel that Human Sciences gave strong support to such values as "individual standards of achievement," "individual personal development," and "cooperative social development."

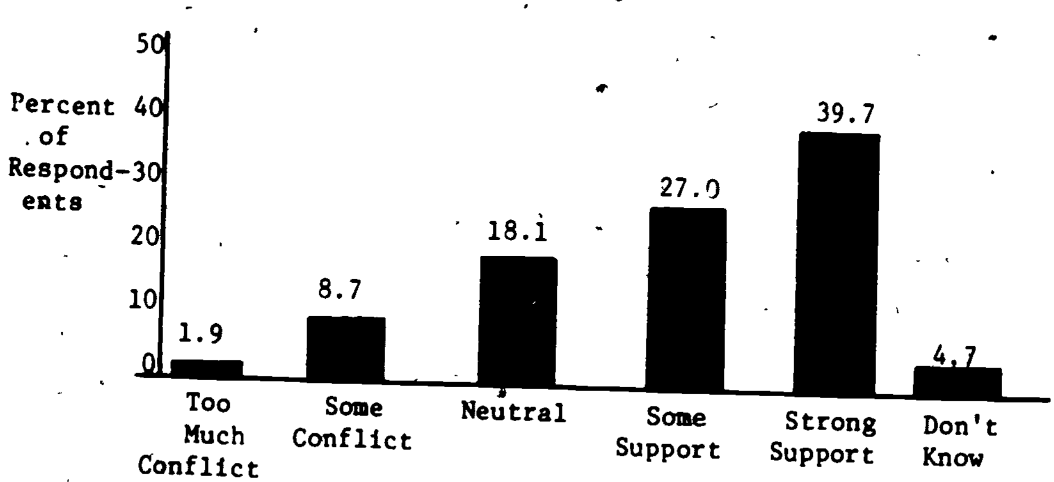


Figure 25. The degree to which Human Sciences conflicted with or supported fourteen statements of "educational values."

Two statements had the modal response of "some conflict." These were "good classroom discipline and order" and "teacher control of student activity." The latter is surprising since one would expect much conflict with this value in Human Sciences. A detailed analysis of responses in conjunction with circled values would be needed to clarify the interpretation of this question.

A second level analysis of the responses to the fourteen statements was made by calculating mean responses[1] for two types of statements. Five of the statements (A, C, E, G, and M: Item 4, Appendix FFF) were statements that contradicted the educational values of Human Sciences and had a desired conflict response (mean 3.0). Nine of the statements (B, H, D, F, H, I, J, K, L, and N: Item 4, Appendix FFF) were statements that the developers wanted to be evident in Human Sciences and had a desired support in Human Sciences (mean 3.0). The mean for the first group was 2.97 and for the latter group, 4.46.

These data indicate that Human Sciences was seen as neutral toward statements (teacher control and content transmission) with which the developers thought the program would conflict, and as being in support of statements (individual development and inquiry skills developement) that were planned to be effected in the Human Sciences Program. Thus, Human Sciences, from these statements, was not found to be in support or conflict with values characteristic of regular textbook science programs and in support of values of individual development that were intended by the developers.

The second section of "Evaluation of Human Sciences" was designed to gather opinions about the kinds of change required by the introduction of the program. Four questions (5-8) had tallies that conflicted with comments made. Examinations of the items indicated a confusion in marking that was caused by the layout and design of boxes to be checked and identification of the boxes. Those items will not be interpreted here. (See Appendix EEE for a preliminary analysis of the comments.)

Question 9 asked respondents to rate five statements on a five-degree rating scale in relation to whether Human Sciences had stopped five listed student behaviors, such as "depend on teacher for planning, direction, objectives," and "act as passive recipients of knowledge." Figure 26, page 151, shows the response frequencies for the sum of all questions across all respondents. These data show that Human sciences was perceived as having a strong effect, in total, on reducing or eliminating the student behaviors that can be characterized as "non-involved class membership."

The last question in this section, Item 10, was a list of ten statements of teacher behaviors designed to determine the degree to which these behaviors were supported or caused by Human Sciences. All of the statements were consistent with what the Human Sciences Program

---

[1]Means were calculated by giving the scale on Figure 20 a value of 1 (too much conflict) to 5 (strong support). "Don't know" responses were omitted.

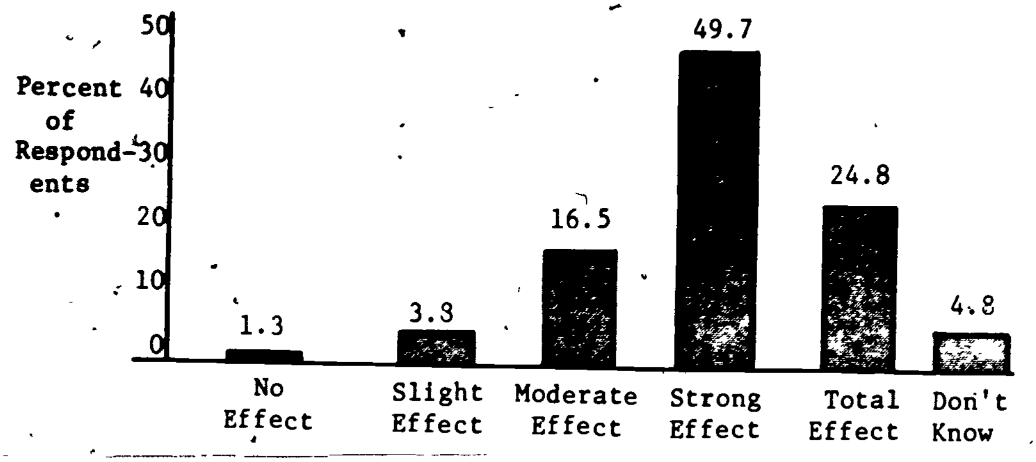


Figure 26. The degree to which Human Sciences stops student behaviors of "non-involvement."

was attempting to have teachers do in test classrooms. There was general agreement by both teachers and administrtors for most of the statements. The summary data for the total group of questions is shown in Figure 27.

The modal response for the statements was that Human Sciences gives strong support to such teacher behaviors as encouraging student self-evaluation, activity selection, action participation, pursuing multiple objectives, working with small groups, and class planning in terms of individual students. One-fourth of the respondents indicated that Human Sciences caused these teacher behaviors to occur. The behaviors described in part the meaning of "teacher as facilitator" in materials prepared for teachers. The only area of disagreement was on one statement where teachers differed from administrators as to whether the program supported student pursuit of objectives they (students) had established and planned.

The questions in the third and final section of the questionnaire were intended to give the Human Sciences staff opinions about the

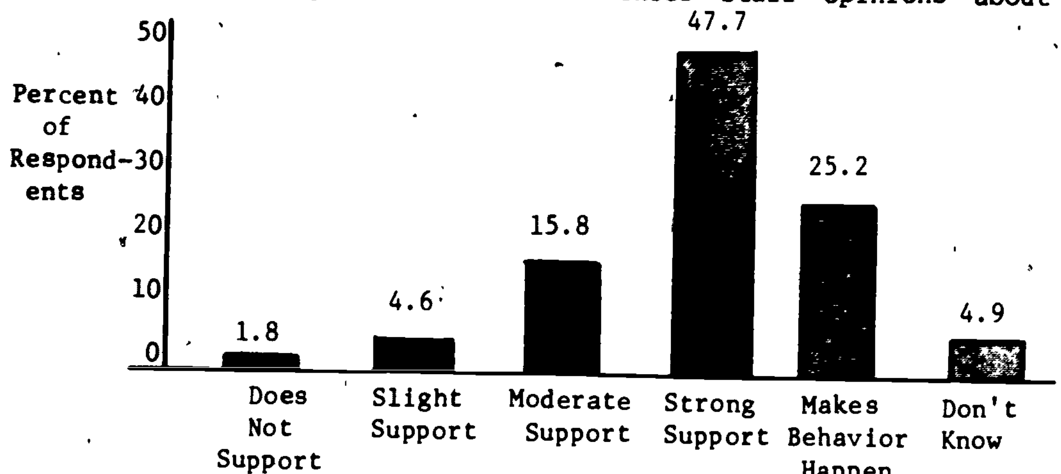


Figure 27. The degree to which Human Sciences supports teacher behaviors that are "facilitative."

communications between the staff and the test schools. Item 11 was concerned with communications about the community seminar held in the spring of 1973 prior to the initiation of testing in the fall. Figure 28 indicates that communications were from good to excellent in most cases.

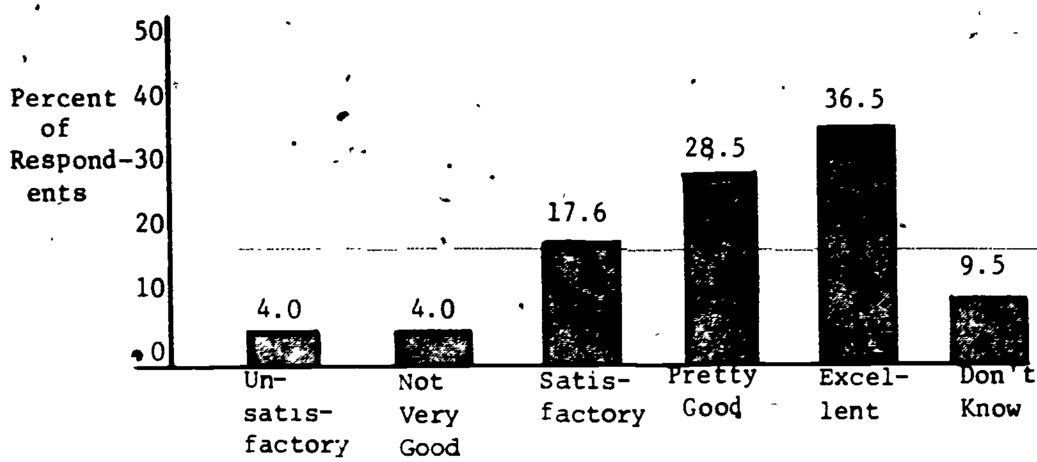


Figure 28. The effectiveness of communciations for arrangments for the Level I teacher orientation conference.

Only 8 percent of the respondents found communications poor. In several instances, poor communications were the result of an individual staff member at a single test site. The need for effective redundancy in communications was clearly apparent in these situations.

Item 12 was concerned with communications with teachers and administrators about the teacher orientation conference held in the fall of 1973. Data presented in Figure 29 support the interpretation that communications in this regard were good to excellent. Within this series of questions were questions about the content of the conference. The low ratings by about 8 to 9 percent of the respondents indicate that the content was not universally considered effective.

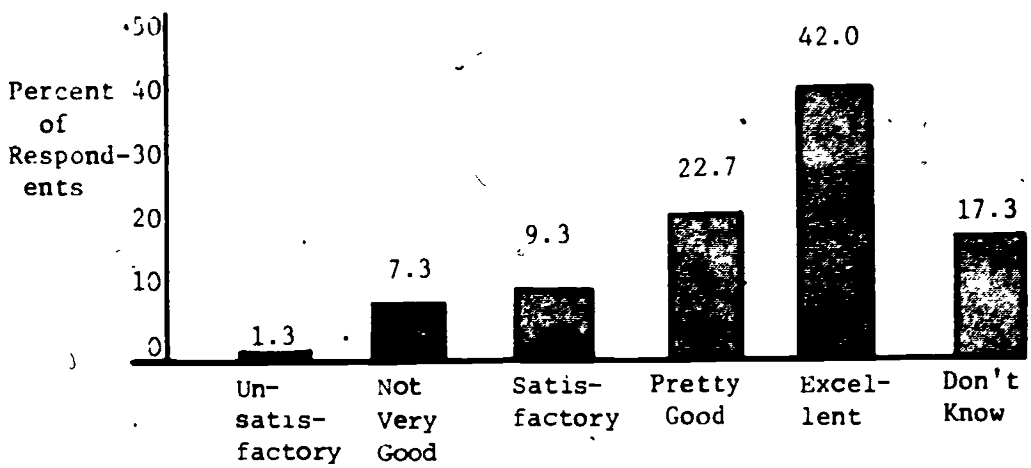


Figure 29. The effectiveness of communications for arrangements for the Level I teacher orientation conference.

On a separate question, respondents were asked to compare early (spring, 1973) with current (spring, 1974) communications. Over half were satisfied with both. Forty-four percent felt that present communications were most satisfactory (data not shown, See Appendix FFF for tabular data).

The next questions were concerned with communications by Human Sciences staff members during site visits. There were five statements to be rated on a five-point scale from unsatisfactory to excellent. Summary data for ratings of five statements about site visit communications are shown in Figure 30. The data show that the respondents to the questionnaire rated site visit communications good to excellent (but at one site communications were rated unsatisfactory).

On a separate question, respondents were asked to compare early and "present" communications. Only 9 percent rated earlier communications more satisfactory than present, indicating that steps taken to improve site visit communications were implemented (see Appendix FFF for data).

The final questions asked for ratings for each of the seventeen different communications materials and processes used to convey information about the Human Sciences Program to teachers and administrators. Complete data of these ratings are provided in Appendix FFF. Personal contacts with staff and personal correspondence with staff were the most highly rated forms of communcation with over 60 percent of the respondents rating them excellent. "Teacher Observation Records" and "Student Journals" received more "not very good" or "unsatisfactory" ratings than did other methods of communication.

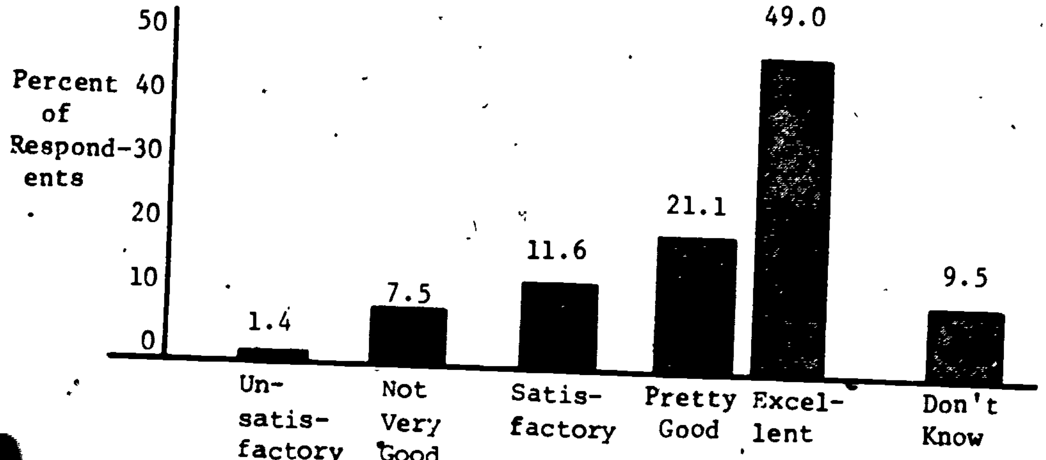


Figure 30. The effectiveness of site visit communications by Human Sciences staff.

The "Evaluation of Human Sciences" questionnaire was completed by thirty-three administrators and teachers at the seven Level I Human Sciences test sites. The questionnaire sought opinions about the amount of change and kinds of change necessitated by the introduction of Human Sciences, and the effectiveness of communications, both written and oral, about Human Sciences. Data from the questionnaire suggest that Human Sciences required more school support in terms of space and consumable materials; makes it easier for teachers to give attention to the development of students as individuals; gives about the right amount of emphasis to goals usually associated with student development, but not enough (from the teacher's viewpoint) to facts and principles of science; and provides support to the development of the student as an independent learner. Human Sciences also supported change in student behavior from passive learner to active participant in learning, supported teachers in moving away from teaching students as a class group and planning and presenting information to the class as a group, and supported the development of student self-evaluation and individualization of instruction. The variety of communications used in the Human Sciences Program was appraised as being from good to excellent. Particular instances of poor communication were exposed and steps were taken to improve them. In general, teachers, and administrators were positive about Human Sciences, as expressed in their responses to "Evaluation of Human Sciences."

Content Review Conferences

Two content review conferences were held at the BSCS building in Boulder, Colorado. Twelve reviewers participated in the first conference on April 18 and 19, 1977. On April 21 and 22, fourteen reviewers participated in the second conference. The number of reviewers in relation to the bulk of HSP materials and available meeting space were factors that suggested two conferences would be more effective than only one conference in which all twenty-six reviewers participated. The agenda and review procedures for the two conferences were the same.

Reviewers were given a brief introduction to the HSP program and received specific activity review assignments corresponding to their context expertise. Each participant reviewed thirty-five to forty-five activities. The review schedule was organized so that almost every activity in the program would be reviewed by at least two reviewers. Because of the interdisciplinary nature of the HSP activities, on an average, eighteen reviewers were involved in the review of each of the modules.

Each reviewer was asked to follow the procedures outlined below for each activity he/she reviewed:

1) Read the student activity card, worksheets and related readings which accompany the activity.
2) View and/or listen to audio-visuals that are included with the activity.
3) React to the materials by writing comments directly on the student activity card.
4) Complete an activity evaluation form (see Human Sciences Evaluation Materials).

The activity evaluation form was constructed to address two major concerns about an activity: a) is the content accurate?, and b) are the directions and procedures technically adequate and accurate? In addition to these concerns, the reviewers were encouraged to write specific comments about the activity in a space provided on the activity evaluation form. The evaluation forms and activity cards containing reviewer's comments were filed by activity to use in activity and module revision.

Several content reviewers were asked to study specific modules for purposes of reacting to the content organization and to suggest additional student experiences which could be included in the module or perhaps in some other part of the program. Those content areas in which input was specifically requested were medicine, ecology, anthropology, and the physical sciences. Reviewers with expertise in these disciplines interacted with staff personnel and provided some new ideas for strengthening the program.

Persons with expertise in developmental psychology were called upon to view the program in a different light. Their opinion was sought as to whether or not the cognitive level of activities was congruent with the developmental levels of 11- to 14-year-olds, and whether or not selected activities were structured to move from concrete to formal operations in a systematic manner.

Summary of Reviewers' Feedback

It is difficult to assess the reviewers' responses quantitatively. Each evaluation form and each comment were treated individually. However, an idea as to the extent of the revision task in regards to content may be inferred from the data in Table 89, page 156. Table 89 shows the number of responses and the percentage of total responses to recommending "some staff work," or "extensive work," the critical response choices about activities. For example, in the BEHAVIOR module, twenty-nine activities were evaluated and fifty-seven activity evaluation forms were completed by the reviewers. Two forms were completed for all but one of the activities. Six out of fifty-seven, or 10 percent of the reviewers' responses to the statement were, "needs some staff work to complete the revisions I have suggested." Generally, each activity was reviewed twice, thus two evaluation forms with each response selected could refer to the same activity. Therefore, as few as three activities (1/2 x 6) or as many as six activities may need some staff content revision work. Likewise, two out of the fifty-seven, or four percent of the responses to the statement were responses indicating "needs extensive work to be useful." Therefore, based on the content reviews, only one or two activities in the BEHAVIOR module needed extensive content revision. Similar analysis can be used for the responses to the statement about directions.

The column totals at the bottom of Table 89 give an indication of the content reviewers' responses to the total program. Looking only at

Table 89. Summary of Evaluation Form Responses

Module	Activity Content: Needs Some Staff Work	Activity Content: Needs Extensive Work	Activity Direction: Needs Some Staff Work	Activity Direction: Needs Extensive Work
BEHAVIOR (57 forms)	6 (10%)	2 (4%)	3 (5%)	1 (2%)
GROWING (53 forms)	5 (9%)	3 (6%)	4 (8%)	1 (2%)
LEARNING (74 forms)	8 (11%)	1 (1%)	8 (11%)	3 (4%)
SENSE...OR NONSENSE? (53 forms)	7 (13%)	0 (0%)	6 (11%)	0 (0%)
SURVIVAL (68 forms)	7 (10%)	2 (3%)	5 (7%)	3 (4%)
Totals for Level I	33 (11%)	8 (3%)	26 (9%)	8 (3%)
PERCEPTION (89 forms)	9 (10%)	1 (1%)	7 (8%)	2 (2%)
RULES (85 forms)	6 (7%)	3 (4%)	4 (5%)	1 (1%)
WHERE DO I FIT? (121 forms)	7 (6%)	3 (2%)	10 (8%)	3 (2%)
REPRODUCTION (64 forms)	3 (5%)	1 (2%)	2 (3%)	0 (0%)
SURROUNDINGS (42 forms)	4 (10%	1 (2%)	4 (10%)	0 (0%)
Totals for Level II	29 (7%)	9 (2%)	27 (7%)	6 (1%)
CHANGE (85 forms)	5 (6%)	1 (1%)	9 (11%)	0 (0%)
FEELING FIT (100 forms)	3 (3%)	3 (3%)	4 (4%)	0 (0%)
INVENTION (47 forms)	7 (15%)	2 (4%)	8 (17%)	1 (2%)
KNOWING (70 forms)	3 (4%	0 (0%)	1 (1%)	0 (0%)
Totals for Level III	18 (6%)	6 (2%)	22 (7%)	1 (0%)
TOTAL—All Activities	80 (8%)	23 (2%)	75 (7%)	15 (1%)

the two "need extensive work" columns, it appears as if from 15 to 23 or only one to two percent of the activities need extensive content revision, as evaluated by content reviewers. The actual decisions regarding content revision were made for each individual activity when all evaluation data for the activity were evaluated.

The general reactions by psychologists, as to whether or not the cognitive level of activities were congruent with the developmental levels of 11- to 14-year-olds, were positive. The reviewers also felt that activities were structured to move systematically from concrete to formal operations.

The following comments and suggestions reflect the type of feedback information at the module level that was provided by some of the content reviewers:

LEARNING Module

"I strongly recommend that the basic, almost 'simplistic' science content of this module is one of its strengths. Developmentally, the vast majority of sixth graders are not ready for concrete, let alone formal, operations."

GROWING Module

"I think the unit is weak. The purpose is not clear. More emphasis (is needed) on what growth is and on animal growth."

CHANGE Module

"There is a strong movement among children and adjults that research with animals is fundamentally an inhumane and unethical practice. The positive aspects of the exercise should be stressed."

Several of the reviewers offered suggestions relative to their specific content area. The recommendations varied from inclusion of "missing themes" to a reorganization of the program to reflect a "continuity of content." A capsulization of some of their suggestions follows:

"Include more experiences that focus on the United States as an aging society."

"Include more activities having to do with the future."

"Reorganize the environmental biology activities to stress ecological concepts. Include experiences at all three grade levels."

"Add more activities about animal behavior, e.g., mimicry, pheromones, imprinting. Relate these to human organisms."

In addition to the previous feedback, an outline of ideas was presented for physical science activities.

Despite critical reactions which were encouraged, the general atmosphere of the two conferences was highly commendable of the program--"A general comment overall is BRAVO! You are doing a pioneering job, and apparently doing it well. I was impressed by the breadth and scope of what you have covered, and it is a step function beyond the state of the art."

In summary, the content review conferences provided activity-specific information which was used to help ensure the content integrity of the program. The reviewers' reactions to and comments about each activity, entire modules, and the total Human Sciences Program served as important and useful input in the final revision of the experimental materials.

The Public Review Conference

The public review conference was held on April 14 and 15, 1977 at the BSCS building. Seventeen reviewers and five staff members participated in the two-day conference. Prior to the conference, staff members identified several types of activities that they felt should be reviewed at the conference. Any activity which wholly or in part fell into one of the following categories was included on the review list:

1. activities in which students work with potentially harmful materials, e.g., chemicals, electrical tools, fragile glassware;
2. interview activities that require students to leave the classroom and/or school building;
3. activities that deal with issues that might be considered potentially offensive or controversial by parents or community groups, e.g., reproduction, drugs;
4. activities in which morals, values, ethics, or related issues are addressed;
5. activities in which students are required to participate in role-playing.

Names of parents whose son or daughter had been involved in an HSP class were requested from the field test sites and from the dissemination center personnel. At least two names were requested from each site. An attempt was made to have one parent reviewer from each of the sites. Nine parents representing eight test site and dissemination center schools were invited to participate in the review conference.

Reviewers from National Educational Groups

Ten organizations with special interests in education were asked to recommend competent individuals, preferably parents of children of middle school age who might act as reviewers of the experiemental HSP materials. A total of eighty-five individuals were recommended from the seven organizations that responded to the request for names. To produce a balance among the participants, this list was reduced to seventeen based on criteria such as sex, race, geographic area, type of involvement in education, and recommending organization. Of the 17 non-HSP parents contacted, eight accepted an invitation and participated in the review conference.

The reviewers worked in paired groups; a reviewer who was familiar with HSP was paired with one not familiar with the program. The specific activities to be reviewed by each group were designated on charts listing the titles of all activities in each of the modules. When a pair finished working through an activity, each member of the pair completed an evaluation form, independent of the other member (see Human Sciences Evaluation Materials for the evaluation form). All reviewers were encouraged to respond and evaluate as honestly and critically as possible. When a group finished reviewing the assigned activities in a module they reacted to the module as a whole before moving on to the next module. Upon the completion of all assigned activities, group members worked alone or in pairs and reviewed activities of their own choosing from any module they desired. They were also encouraged to write specific comments about the total program, particular modules, or individual activities.

During the last few hours of the conference, the reviewers discussed their overall reaction to the program in a group situation with staff members. Some of the highlights of the discussion are outlined in the next section.

Summary of Reviewers' Feedback

In Table 90, page 160, the responses to each of four statements on the "Activity Evaluation Form" have been condensed by level of activity. The responses, "strongly agree" and "agree" were considered as positive reactions to the statement. The responses "disagree" and "strongly disagree" were considered as negative reactions. The percentage of reactions which were positive was obtained by adding the "strongly agree" and "agree" responses, dividing by the total responses to the question, and then multiplying by 100 (Total Responses x 100).

Generally, the responses of the public reviewers to the activities were very positive. Out of 260 activities reviewed (508 total reviews) approximately 8% or 21 activities, were reacted to negatively. If we assume that the activities not reviewed were neutral or positive then we can conclude that only 4% (21 out of 530) of all HSP activities might be considered potentially controversial. These particular activities received careful attention during the final revision.

In addition to completing the "Activity Evaluation Forms," some of the reviewers suggested activity topics which might be added to specific modules. Some reviewers also wrote their general reactions to the Human Sciences Program as a whole. The discussion between staff members and reviewers also provided some valuable information and suggestions. The major criticisms, suggestions, and positive features of the program, as perceived by those who participated in the discussion, are highlighted below. A complete transcript of the discussion is included in EP 7,704-18.

Table 90. Summary of Responses by Public Reviewers

Level	Statement	SA	A	SAA	D	SD	DSD
I	1	93	80	78	17	6	10
222 Reviews	2	90	78	76	22	7	13
	3	80	81	73	21	5	12
	4	89	69	71	24	13	17
II	1	96	74	85	7	0	3
201 Reviews	2	100	65	82	10	2	6
	3	65	87	76	6	4	5
	4	93	63	78	8	5	6
III	1	47	30	91	1	0	1
85 Reviews	2	53	20	86	3	0	4
	3	44	25	81	4	0	5
	4	53	24	91	3	0	4
Total of All Levels							
508 Reviews	1	236	184	83	25	6	6
(About 260	2	243	163	80	35	9	9
activities)	3	189	193	75	31	9	8
	4	235	156	77	35	18	10
		903	696	79	126	42	8

Features of the program that the reviewers considered positive were the structure or format, variety of experiences, student independence, and community involvement. The following comments emphasize these features.

- "The nonbook format "invites you to participate and is easy to manipulate."

"Another bonus I find in the program is the community resources that they (students) are able to utilize."

"I like the independence it (HSP) gives the student to do something all alone without that much guidance."

"...They (students) were able to be responsible and not take advantage of this freedom."

The program "is for the intellectual...if you are intellectual...and it's for the person who is interested in being active, working with his hands and doing these other things, and from that point this boy (reviewer's son) is now turning out to be intellectual. In the process it brings back the interest in education to get on..."

"My children had absolutely no problems with your program or with other similar programs-going into, for instance, a police station and asking them to help... or going into a hospital. But to ask children to go to neighbors, which a number of activities do, I think is a big imposition on the neighbors." (There were mixed feelings regarding this point.)

A need was stressed for teacher training and parent involvement.

"There are a number of activities where the parent would not feel comfortable in having their children totally unsupervised, conducting that portion of the activity."

"There is so much in it (the program) that I feel it is impossible to expect, particularly an elementary teacher who has other disciplines to teach, to be able to do an excellent job with this program."

"I am concerned about the proper training (of teachers) especially in the sex and the drug areas and in areas which involve some kinds of personal prejudices."

Regarding the "back to basics" movement, the reviewers felt that the HSP program provided a good foundation in the basic skills as well as a wide variety of interdisciplinary experiences.

"That is part of what makes it so exciting is that so many of the modules incorporate math and writing, literature and imagination, and you know, verbalizing--so amny skills that can be tied into one module. The benefit of that is that it provides the student an opportunity to see there is benefit in knowing math. I can use math to find out other things. There is benefit in knowing how to read. I can use it to discover other kinds of things. I think it gives a sense of purpose to the kinds of skills we want students to learn."

In summary, the public reviewers reacted positively to the Human Sciences Program and materials. The public reviewers' reactions and comments served as important and useful information in the evaluation of the Human Sciences Program.

Summary

Adults who were associated with Human Sciences were asked to participate in the evaluation of the experimental student and teacher materials. Those adults included teachers, school and school district administrators, academic specialists, representives of national organizations concerned with education, and parents of students in test schools.

The overall result of the evaluation responses reported here was positive; both parents and educators were generally supportive of what the Human Sciences was attempting to do and how it was being done. This information was encouraging and pleasant to developers, but it is not the important data to attend to in formative evaluations. Minority criticisms and questions raised that bear on the adoption of innovative materials were what needed to be considered. These issues had the potential of representing obstacles to adoption of the materials when the prestige of participating in field testing would no longer be available.

One issue raised was the potential for interdisciplinary curricula. Some adults could not see a place for "psychology" or "sociology" in a science course. This viewpoint came out in some parents' concerns that students learn science content, in which social and behavioral sciences were excluded. This was also expressed in the concern of some teachers for more attention to science content. This concern could be translated, in revision of the materials, into increasing the science terminology of the natural science activities, to increasing the number of natural science activities, thereby, decreasing the interdisciplinary character of the program.

A second issue raised in this part of the evaluation program was the concern for including physical sciences in the Human Sciences Program. Many schools follow a general science program in the middle school or junior high school years. This is done either by three-year programs of general science or by a one-year or semester course or courses in life science, physical science, and earth science. The issue raised a significant problem for the revision of Levels I, II, and III.

A third issue raised was in regard to the goal of advancing the development of logical thinking and moral reasoning. It was unfortunate that both of these goals were included with one goal statement, for it cannot be ascertained whether parents placed low priority on one or both of these areas of development. If some parents were more concerned with knowledge transmission goals than with developing logical thinking and/or moral reasoning, the latter would not be viewed as valuable by these parents.

A fourth issue raised was that of the goal conflict among the developers, school personnel, and parents with regard to decision-making about controversial social issues. The developers included social issues because they believed them to be important, societally, and because they relate curriculum to the lives of students. Yet, less than one-fourth of the parents placed this as a high priority as an educational goal.

As can be seen in the data presented in this chapter, adult evaluation of curriculum materials provided important insights into the potential impact this innovative curriculum might have on schools. The questions raised here were considered by the project staff and influenced the final revision of the Human Sciences Program. The best example of this impact was the decision to include physical science activities into all modules, and to develop a physical sciences oriented module, MOTION, which would be included in Level I.

CHAPTER 9

ADDITIONAL STUDIES

Several studies that were independent of the module evaluation studies were conducted following the formative evaluation of Human Sciences. Readability studies were conducted on a sample of activities from each level. A study of student attitudes at the end of the three-year field test was conducted with eighth-grade Human Sciences students and a comparison group of eighth-graders in each test school. A partial replication of this study was made using the same attitude scale, but with students in Human Sciences classes who were testing KNOWING, a Level III module. This new group of eighth-grade students was asked to rate Human Sciences and their regular science course prior to testing the KNOWING module.

The 1974 edition of "How Is Your Logic?" was given to sixth graders at the end of the field test of Level I. The 1976 edition was given to all eighth-graders at the end of the three-year field test in May, 1976. These instruments were scored in an identical manner and a study of changes in logical competence over the two-year interval is reported here.

Finally, the National Science Foundation conducted an external evaluation of Human Sciences as part of their program review in 1975, the results of which are reported here.

Readability Study

A sample of activities from each level of the Human Sciences modules was chosen for readability analysis. Modules were selected by the Human Sciences staff based on their judgment regarding reading difficulty. The modules were chosen as representing the most difficult reading tasks. A single problem area, including all the activities in the problem area, was selected from each module, except for KNOWING. KNOWING has eight clusters instead of three problem areas as internal structural elements. Two clusters in KNOWING were selected for analysis. The materials were sent to Dr. Milton D. Jacobson, University of Virginia, Charlottesville. Dr. Jacobson entered the entire manuscripts into a computerized reading analysis program.

In SENSE...OR NONSENSE?, the problem area studied was "Identifying," with nine activities; in PERCEPTION, the problem area was "Perceiving," with nineteen activites. The two clusters analyzed in KNOWING were "Materials and Shapes," with four activities, and "Knowing the Past," with seven activities.

Dr. Jacobson's computer analysis computed Fry, Dale, and Dale-Chall readability scores. The results of the study are shown in Table 91.

TABLE 91
Readability Scores Calculated by M. Jacobson

Level	Module	Fry Score	Dale Score	Dale-Chall Grade Equivalent	Number of Words
I	SENSE...OR NONSENSE?	6.61	5.80	6.61	7,366
II	PERCEPTION	6.21	6.11	7.22	12,537
III	KNOWING				
	Materials and Shapes	7.80	7.21	9.41	4,784
	The Past	8.08	6.88	8.75	11,857

The study conducted was a preliminary one, without data cleanup. The formulas included use of the 3,000-word list by Dale. Examination of text printouts indicated that many words counted were typographical; that pronunciations of words were counted, as well as the words themselves; that names of materials used in activity construction, such as hand saw, battery, and beaker were inserted; that tables of activities were included; and that cities, continents, and names of organisms were also counted. These inclusions, not on the Dale list of 3,000 words, increased the reading difficulty in the Dale-Chall computations.

Since the results showed that the materials were on grade level, and were most likely an overestimate of reading difficulty, a final run on revised materials was not made. Rather, steps were taken to eliminate unnecessary technical terms, to provide physical referents for names of objects cited, and to use important technical terms in as many activity contexts as possible. These measures should result in activities that can be read by the majority of the target student population.

Attitude Study

A "Science Questionnaire" was administrered to students in Human Sciences test classes (eighth-grade level) in May, 1976 and also administered to an equal number of students taking regular eighth-grade science courses in the same schools. Approximately 600 students were tested.

This study was designed to test the null hypothesis, "There is no difference in attitudes between students in Human Sciences classes and students in eighth-grade science classes toward their science course as measured by responses to the "Science Questionnaire." This hypothesis was subsequently divided into several subhypotheses. Treatment (Human Sciences and eighth-grade science), sex, and school were the factors in the design. In addition, a discriminant analysis of responses to selected variables was computed. A complete report of the study has been published (Robinson, 1980); hence, only a summary will be presented here.

One-half of the students (experimental group, N about 300) were in Human Sciences test classrooms. About 240 of these students were in their third year of Human Sciences. The remaining students were about equally distributed between their first and second year in test classes.

The other half of the students tested (comparison group, N about 300) were in regular science classes in the test schools. Data were secured from nine teachers' classes, two of whom were also teaching a Human Sciences test class. No data were gathered to determine the particular curriculum materials being used in these classes except to ascertain that no Human Science materials were in use in the classes during the 1975-1976 school year.

Instrument

The major component of the "Science Questionnaire" was a semantic differential instrument asking students to express their feelings toward their science course. Bipolar adjectives from Osgood, Suci, and Tannenbaum (1957) were reviewed, as were the modifications prepared by DiVesta (1969). Choice of bipolar adjectives was made to reduce dependence on metaphorical interpretations and to select adjectives that had been successfully used with thirteen- to fourteen-year-old students.

Four conceptual dimensions were hypothesized for the instrument: evaluation (like or dislike), value (worth), activity (active involvement), and interest. The conceptual dimensions and bipolar adjectives for each dimension were presented in Chapter 7. The results of the analysis (Table 92) indicate that the postulated subscales can be treated as subscales in subsequent analyses.

TABLE 92
Scale Analysis and Scale Scores of the Four Postulated Subscales of the "Science Questionnaire" Semantic Differential

Subscale	Reliability[1]	Homogeneity Ratio	N	Mean	S.D.
1. Evaluation	.87	.57	601	25.34	6.83
2. Value	.80	.44	577	25.56	6.24
3. Activity	.70	.32	584	22.95	6.27
4. Interest	.83	.62	602	13.90	4.86

[1]Cronbach's alpha

The student scores for each of the four scales of the semantic differential were analyzed in the 2x2x5 factorial design using a factorial analysis of variance program, and multiple classification analysis using SPSS subroutines. The three factors of the design were: sex, school, and course (experimental group versus comparison group). These variables were also examined for possible two-way interactions.

Two schools were removed from further analysis because of incompleteness of data and marking errors such as apparently deliberate patterns of 1, 2, 3, 4, 5, 6, 7, repeat; pairs of responses in regular patterns; etc. The comparison groups in these schools were also eliminated for analysis. The deletion reduced the number of students in the study to approximately 400.

Discussion

Four important findings resulted from the analysis of the four scales of the semantic differential instrument. First, mean responses of Human Sciences students in all test schools were all on the high positive side of the scales.

Second, the Human Sciences student ratings were not only positive, but were significantly more positive (p=.001) than the ratings of students in regular science courses in the same schools.

Third, the effects of school were significant (p=.001 to .03) on all four scales. Interaction effects between sex and course were found only on Scale 4, Interest. On this scale, girls rated Human Sciences higher than did boys, and eighth-grade girls in regular science classes rated their course lower than did boys.

Finally, the discriminant analysis indicated that significantly more positive ratings were made by Human Sciences students about their course, especially on five bipolar adjectives. These students characterized the Human Sciences course as significantly more pleasant, full, happy, important, and fun than did the comparison group.

A partial replication of this study was made in 1977 with the KNOWING module test group of eighth-graders. The results of this study were reported in Chapter 7. Students rated Human Sciences significantly more positively than they rated their regular science course prior to testing KNOWING on all four subscales of the "Science Questionnaire." Girls rated Human Sciences significantly higher than boys on all four subscales. They rated their regular science course signficantly lower than boys on the Interest subscale. In this study, the teacher variable was controlled--the same teacher taught the regular science class and then switched to the KNOWING module in mid-semester.

Logical Competence

The design of the Human Sciences Program assumed that students would vary in cognitive development, as conceptualized by Inhelder and Piaget (1958). The assumptions were made that some sixth-grade students would not be capable of solving concrete operational problems, that most sixth graders would be able to solve many concrete operational problems, and that a few would be capable of solving formal operational problems. This assumption had consequences for curriculum design, for it was interpreted to require that activities designed to meet the characteristics of the student population should be designed primarily at the concrete operational level. At the same time, opportunities needed to be provided to enable students to develop and/or consolidate and elaborate concrete operational thought into as many content contexts as possible, at the same time providing for others to develop formal operational thought competencies.

This grounding of the Human Sciences curriculum in cognitive development theory made it imperative to gather data on the logical competence of students, if at all possible. Attempts made to secure the use of one experimental paper-and-pencil test were not successful. Therefore, the Human Sciences staff decided to turn to a consultant who had intitated work on a problem-solving and logical competence test as part of the BSCS Life Sciences for the Educable Mentally Handicapped program (Steele, 1974).

Working with Dr. William M. Gray, then at the University of Dayton (currently at the University of Toledo), a paper-and-pencil measure of logical competence titled "How Is Your Logic?" was developed. This instrument was developed in two parallel forms, Form A and Form B, of fifteen items each. The 1974 edition was administered to all students in Human Sciences test classes in May, 1974, at the end of their sixth-grade year, the first year of the Human Sciences field test.

Analysis of the data from the 1974 administration (Gray, 1974) led to the revision of the instrument and culminated in "How Is Your Logic?", Form A and Form B, 1976 Edition (Gray, 1976). This instrument was administered to all eighth-grade students in Human Sciences test classes in May, 1976.

There were two major questions to be asked from these two assessments. First, what is the cognitive competence of the students at the end of the sixth grade and at the end of the eighth grade, as measured by "How Is Your Logic?"? Second, is there any change in competence in the two-year interval between test administrations?

"How Is Your Logic?" was the result of over ten years of work on the part of Gray. The final scoring criteria, completed in 1979 (Gray, 1979), were used to rescore both the 1974 and 1976 student test booklets. These scoring criteria enable the evaluator to score each student response into one of eight categories: preoperations (includes "don't know" and no attempt to answer), concrete operations I, concrete operations II, concrete operations III, formal operations I, formal operations II, and formal operations III. These designations follow those of Flavell (1977) and Inhelder and Piaget (1958) and will not be discussed further here.

Both the 1974 and the 1976 editions included concrete and formal operational problems, as simple in content as possible. Teachers read each item orally and explained any difficult terms as needed in order to reduce the confounding effects of readability. Further information about "How Is Your Logic?" is presented in Gray (1981) and Gray and Robinson (in preparation). A preliminary study of the 1976 edition is provided in Robinson and Cobern (1979).

Factor analyses (Rao's canonical, SPSS version 8.0) were computed for the 213 students who had test scores for the 1974 edition and the 1976 edition of "How Is Your Logic?" Using a .40 factor loading as a criterion for selecting items for factors, nine factors were identified in the 1974 data and seven were identified in the 1976 data. These factors accounted for 55.0 and 55.9 percent of the variation, respectively. Three formal item factors and two concrete item factors were selected for analysis on the basis of having comparable logical

TABLE 93

Factors Selected from Rao's Canonical Analysis for Comparison of Potential Changes in Student Competence from May 1974 to May 1976

1974				1976			
Factor	Item	Logic Description	Factor Loading	Factor	Item	Logic Description	Factor Loading
2	A3	Make correct implication	.54	7	A2	Make correct implication	.54
	A3X	Make correct implication	.76		A3	Make correct implication	.44
	A5	Make correct implication	.65		A9	Make correct implication	.50
	A5X	Make correct implication	.81		B9	Make correct implication	.41
	B8	Make correct implication	.35		B10	Make correct implication	.45
7	B14	Deny correct implication	.86	2	B3	Deny correct implication	.59
	B14X	Deny correct implication	.87		B4	Deny correct implication	.64
5	A15	Permutations	.77	1	A6	Permutations	.55
	B15	Permutations	.78		B7	Permutations	.59
3	B1	Increasing series	.82	3	A4	Decreasing series	.44
	B2	Increasing series	.82		B5	Decreasing series	.62
4	B9	Decreasing/decreasing series	.82	4	A8	Increasing/decreasing series	.75
4	B10	Increasing/decreasing series	.85		B8	Decreasing/decreasing series	.74

problems. The factors selected, their logical structure, and factor loadings are shown in Table 93. One item with a factor loading of .35 was included for two reasons: its logical relationships and its match in number of items with the comparable factor. Two items were omitted from factor 1 and one from factor 3 of the 1976 test to reduce the item numbers to be comparable to those of the 1974 edition. These nearly identical factors were used for further analyses.

The same individuals were tested, with two years between tests. Therefore, the paired samples computations of the subprogram t-test (SPSS version 8.0) were used to test the null hypothesis that there would be no significant difference on the five comparable factors between the 1974 and 1976 administrations. A one-tailed test with significance preset at the .05 level was selected for testing the null hypothesis, as it was estimated that any difference would be directional, in favor of the 1976 administration. The output two-tailed probability was divided by two to give the appropriate one-tailed probability. Table 94 shows the comparison of the two comparable concrete items, separately and combined.

TABLE 94

Comparison of Mean Scores for Two Concrete Operational Factors from the 1974 and 1976 Administrations of "How Is Your Logic?", May 1974 and May 1976, N = 213

Subscale	No. of Items	Mean	S.D.	Range	Skew-ness	Reli-abil-ity	T-value	df	1-Tailed Proba-bility
4-1974	2	6.09	2.81	0-8	-1.25	.80	8.36	212	.000
4-1976	2	7.76	.57	5-8	-2.59	.34			
5-1974	2	5.05	1.79	0-8	- .81	.82	5.11	212	.000
5-1976	2	5.92	1.86	2-8	- .41	.79			
4,5-1974	4	11.14	4.02	0-16	-1.29	-	8.61	212	.000
4,5-1976	4	13.68	1.97	9-16	- .36	-			

The null hypothesis for the concrete item subscales was rejected. Students performed significantly better in 1976 then in 1974 on both subscales and on the combined subscales. The lower reliability and high negative skewness of subscale 4-1976 indicated that students were almost all capable of solving these concrete operational problems.

These parallel formal operational factors were used to compare the 1974 and 1976 performance of students. Table 95, page 170, shows the results of these comparisons. Both the individual subscale and combined subscale data show significantly higher competence in 1976 than in 1974.

TABLE 95

Comparison of Mean Scores for Three Formal Operational Factors from the 1974 and 1976 Administrations of "How Is Your Logic?", May 1974 and May 1976, N = 213

Subscale	No. of Items	Mean	S.D.	Range	Skew-ness	Reli-abil-ity	T-value	df	1-Tailed Proba-bility
1-1974	5	20.70	7.58	4-35	.59	.82	2.81	212	.002
1-1976	5	22.54	7.51	5-36	.09	.72			
2-1974	2	6.79	4.02	0-16	.68	.73	2.54	212	.006
2-1976	2	7.80	4.04	2-16	.50	.72			
3-1974	2	4.82	2.64	1-16	2.44	.77	10.95	212	.000
3-1976	2	8.18	4.22	2-16	.55	.78			
1,2,3-1974	9	32.32	11.19	[illegible]-64	.59	-	5.90	212	.000
1,2,3-1976	9	38.51	12.52	12-68	.37	-			

Combining the concrete and formal operational subscales into parallel, thirteen-item scales produced similar results, as shown in Table 96, page 170. The higher level of competence is not only statistically significant, but the effect size of .65 further supports the increased competence attained in 1976.

The results of the data from "How Is Your Logic?" show that students who were in the Human Sciences Program for three years developed increased logical competence in their capabilities to solve both concrete and formal paper-and-pencil problems. Without control groups, which were not possible in the formative evaluation paradigm and funding level for evaluation in the Human Sciences Project, it is not possible to attribute any cause for this effect.

TABLE 96

Comparison of Combined Mean Scores for Three Formal and Two Concrete Operational Factors from the 1974 and 1976 Administrations of "How Is Your Logic?" 1974 and 1976 Editions, N = 213

Date	No. of Items	Mean	S.D.	Range	Skew-ness	T-Value	df	1-Tailed Proba-bility	Reli-abil-ity
1974	13	43.45	13.40	5-77	.14	7.22	212	.000	.86
1976	13	52.19	13.48	24-83	.31	-	-	-	.85

The two questions posed earlier in the paper can now be answered descriptively. At the end of the sixth grade, almost one-third to one-half of the students were giving preoperational responses to concrete operational problems, as shown in Table 97. By the end of the eighth-grade, the percentage of students giving such responses was reduced to about 15 percent. Similarly, the percentage of students giving preoperational responses to formal operational problems was reduced from 39 percent to about 20 percent from the end of the sixth to the end of the eighth-grade (Table 98). Only about 12 percent of the students were consistently giving formal responses to the nine formal problems by the end of the eighth grade.

TABLE 97

Percentage of Students with Correct Logical Responses to Concrete Operational Problems in May 1974 and May 1976, N = 213

	1974			1976		
Logic of Student Responses	Decreasing or Increasing Series	Increasing/ Decreasing Series*	Combined Concrete Problems	Decreasing or Increasing Series	Increasing/ Decreasing Series	Combined Concrete Problems
Preoperations	29.1	55.4	35.7	0.9	42.3	14.6
Concrete I	70.1	44.6	64.3	99.1	57.7	85.4

TABLE 98

Percentage of Students with Correct Logical Responses to Formal Problems in May 1974 and May 1976, N = 213

Year of Testing	Logic of Student Responses	Make Correct Implication	Deny Correct Implication	Permutations	Combined Formal Problems
1974	Preoperations	23.9	32.9	77.9	39.0
	Concrete I	38.0	46.0	13.1	33.3
	Concrete II	14.6	4.2	4.7	12.7
	Concrete II-Formal I	8.7	8.9	1.4	9.4
	Formal I	15.4	0.0	1.4	5.2
	Formal II	0.0	8.0	1.9	0.5
1976	Preoperations	15.5	28.6	33.8	19.7
	Concrete I	30.0	33.3	23.9	29.6
	Concrete II	21.6	17.4	16.0	23.0
	Concrete II-Formal I	21.1	9.9	4.2	15.5
	Formal I	8.0	0.0	8.0	7.0
	Formal II	8.5	10.3	14.1	5.2

These data support the assumptions of the Human Sciences activity design that approached most instruction at the concrete level, and providing opportunities for students to develop formal competencies if they were developmentally capable of this transition. The data also support the provisions of a diverse range of activities, with some, even at the eighth-grade level, that are structured to enable students to develop and consolidate concrete operational thought, avoiding the assumption that all students had already achieved this level of competence.

National Science Foundation Panel Review

During the week of December 8-12, 1975, the National Science Foundation convened a seven-panel review and evaluation of the nineteen precollege curriculum projects then being supported by the NSF. This review was responsive to guidance from the Congress of the United States and the National Science Board.

Forty-two organizations responded to a request to nominate panelists for the review from which seventy-three panelists were selected. The reports produced reflected the views of the panelists alone and not those of the NSF staff. A complete report of all panel evaluations and, in particular, the report of the panel that reviewed the Human Sciences Project is found in NSF (1976).

Nine review questions were presented to the Human Sciences Project staff. Written responses for use by the panel were prepared. The panel responses to the nine review questions are reproduced below from that report.

Question 1: Is there a genuine need for these instructional materials?

The HSP staff conducted a needs assessment as early as 1966 with several additional conferences, checks, and feedback programs designed to seek direction from schools, students, and the public. This needs assessment is considered by the panel an important feature of the project.

Some of the needs to which HSP responds include:

1. learning materials specifically for the middle and junior high schools
2. materials which emphasize societal needs
3. science for the middle-school years which emphasizes the interrelationship of science and society, science discovery and application, and science and other academic disciplines
4. a curriculum with the student and her or his immediate environment as central to the activities
5. a program that considers issues, problems, and values as well as basic content
6. materials appropriate for all students in grades 6, 7, and 8

7. materials that allow choice, sequencing by the student, and individual approaches

These needs are generally identified, discussed, and advocated by current educational leaders, researchers, and philosophers.

Materials are generally appropriate for the diversity of students' maturity during the three-year middle or junior high school years. This program could serve a potential three million students annually. The developers report that the program will initially reach "no less than 10 percent" of this population with an anticipated goal of 50 percent if national implementation programs are conceived and supported.

HSP represents a unique program with the previously listed characteristics. In addition, HSP can be described as a hands-on, student-centered, individualized, and interdisciplinary learning experience. The non-textbook nature of the progam makes it unique. The panel endorses the developer's claim that no viable alternative materials with these characteristics exist.

Question 2: Is there a market for these instructional materials?

At present, many types of middle school/junior high school science programs revolve around the use of a textbook. Thus, the Human Sciences Project materials add a significant dimension to available science curriculum.

Because of the modularized approach, this curriculum, or fractions thereof, can be used within present science courses. The Human Sciences Project could be offered as an alternative to or replacement of the present middle school science program.

Since the curriculum project treats topics not traditionally presented in the targeted grade levels, some in-service teacher education must be included in any implementation program. The training should include discussions of handling potentially sensitive issues (for example, divorce) as well as reviews of basic biological and social sciences.

The panel is not aware of any similar materials which have been produced by commercial publishers for the same audience. Since we believe that this project is a good alternative to present science curricula, we hope that commercial publishers eventually will consider developing materials with characteristics listed in the answers to question 1. However, we do not expect such action until the basic developmental and feasibility studies of the Human Sciences Project are completed.

Question 3: Do these instructional materials possess a clear purpose and rationale?

The instructional materials possess a clear purpose and rationale. The panel agrees with the stated assumptions that, by using the materials, cognitive skills such as problem solving and critical thinking can be learned as can other elements of the curriculum.

The panel agrees with the project staff that "values inherent in the curriculum are critical thinking, autonomous learning, assumption of responsiblity, cooperative efforts in classroom endeavors, shared managerial responsibilities between student and teacher for the classroom environment, decision making, evaluation of data, dealing with problems, self evaluation of individual performance, scientific approaches to problem solving, and value judgments based on evidence."

The instructional materials allow for the fulfillment of the assumptions and the goals since, by design, only those units acceptable to the parents, teachers, administrators, and community members would be used.

The general groupings into which the curriculum is divided are well conceived. In addition it provides for choice on the part of both student and teacher to design a cohesive package to fit the educational needs of the community. The field testing procedures will provide feedback which will allow the staff to revise, add, and/or delete materials. The final product should be a curriculum which is clear and understandable to most students in the target group.

The modules were designed 1) to meet the concerns and interests of the students and 2) to focus upon the interface of the natural and social sciences. The educational effectiveness of the modules is plausible since the materials offer the use of direct experiences as well as phenomena through observation as the learning mode.

Question 4: Is the content of these instructional materials scientifically correct?

The materials are scientifically accurate. The thoroughness with which internal monitoring is performed assures accuracy and currency.

The panel feels that many areas are covered superficially rather than in depth. As a result the program addresses itself toward developing a scientifically literate society. The materials, although human biology oriented, have had input from other natural sciences as well as the social sciences.

In telecommunciations with review panelists, the project staff indicated that some of the materials which do not fulfill the stated objectives are being removed or revised.

Question 5: Is the content of these instructional materials educationally sound?

The fact that HSP addresses itself to current problems suggests that portions of the curriculum could possibly lead to controversy and difference of opinion. The availability of modules on these issues, however, is an attractive feature of the program.

The materials are educationally sound. They are individualized in a manner which gives students the freedom to respond, the freedom to choose, and the freedom to proceed at a chosen rate. HSP is not a course or a course sequence in the traditional sense. Its modular nature provides opportunity for schools, teachers, students, and communities to structure the kind of program that is meaningful to them while providing a resource of suggestions that are concerned with meaningful and significant topics in today's society.

Compared to other available programs, HSP is less dependent upon student reading ability, interest of all students, and general ability, motivation and level of maturity. The approach to HSP content, if handled in the manner recommended by the developers, insures that students at different levels can be accommodated. The materials and the approach, largely because of the nonprescriptive characteristics, appear to be equally appropriate for all students.

The Human Science Program includes a large number of supplementary instructional aids to help schools, communities, and teachers use the materials effectively. In addition, there is an impressive Teacher's Guide for each module. Teachers are provided information to assist with student self-evaluation, with facilitation of further student interaction, and with assurance of appropriate handling of issues with individual students.

Although there has been no attempt to include all science topics, the HSP content is educationally sound in the scope, content, and methods utilized for considering it. The panel is convinced that the HSP materials completed to date and the results of the field tests to date indicate that the materials are indeed educationally sound.

Question 6: Are the proposed and anticipated outcomes of the instructional materials desirable?

The anticipated impacts outlined would provide expanded opportunities to middle school students in science education with the focal point being human sciences. The panel expects students to find the program interesting and stimulating.

Teachers may need to develop additional skills to handle the subject material effectively. In particular, teachers would need to be prepared to handle reactions of students to the social issues that are treated in certain modules.

School administrators and boards of education may be placed in a position of defending the adoption of the materials because of the explicit nature of certain sections on development and reproduction. The target population, because of its diverse level of mental and physical development, may need to be selectively screened to provide alternate activities. This same problem will be experienced with certain non-sensitive materials because some of the activities seem to be simplistic and would not challenge the more mature students in the middle school age group. However, because of the module approach to the learning activities, selective assignment based on the need, abilities, maturity, and interest of the user can be easily arranged.

The panel recognizes the need for educational activity with socially sensitive material. Potential users of HSP should be cautioned that there is some of this type of instructional activity included. The panel suggests that there may be parental reaction to the introduction of a few module topics presented for review.

There is no discernible sexual, racial, or ethnic bias in the material provided the committee. There may be some selective use of modules depending on the sex of the student. However, this is not because of sexual bias but rather as a result of the varied maturity levels of student in the targeted age group.

Question 7: Do these instructional materials present implementation problems for the schools?

To use these instructional materials effectively teachers need skills in the use of individualized instruction procedures, self-paced learning, and the inquiry approach to science. Depending on their background and experience, teachers may need special training in order to guide students of various levels of maturity through those activities which deal with human growth and behavior.

If a school official does not recognize the validity of individualization and self-pacing, the traditional structure can be adapted to this program. The freedom and responsibility to be assumed by the students can be introduced gradually with a minimum of schedule changes and classroom reorganization. It should be recognized that teachers who are to direct more than twenty-five students at a time in this type of program will need some type of classroom assistance (school aides, student monitors, or peer facilitators). The panel would like to emphasize the importance of the teacher commitment to this learning approach.

The modules are multimediated and cost-competitive. Most of the resources needed are normal budget items or available within the schools.

If a school district chooses to adopt the complete set of modules, the administration may wish to conduct the parent orientation program recommended by the project directors. It should be made clear to the parent that there is provision for students to choose topics within a given area as well as to choose areas in which to work. Procedures should be established which provide parents, if they so desire, with the opportunity to participate in their child's choices.

Question 8: Are the costs for implementing these instructional materials reasonable?

The materials costs for the program are no greater, and possibly less, than the costs of other junior high school science programs. Similarly, refill costs should not exceed current costs for middle school/junior high school science course materials.

The present form of the HSP curriculum does introduce subjects which have potential psychological and/or social impacts. For example, modules on the topics of divorce, death, and reproduction are being tested. The panel believes that these impacts may be minimized in three ways:

1. Because the materials are modular, any topics can be eliminated from the program.
2. The project staff and trial centers are undertaking a careful study of all materials. This study includes reaction of parents as well as students. The panel believes these tests will result in an identification and revision of potentially sensitive areas.

3. Orientation programs can help teachers predict and cope with individual student reactions.

The panel believes that a school need not eliminate a topic from the curriculum solely because it is socially or psychologically sensitive. Programs which will better prepare teachers to teach such topics should be funded.

Question 9: Is the management/organization plan adequate for producing these instructional materials?

Answers to questions addressed to the directors of the project disclosed a well-organized plan of consultation with educational administrators, teachers, parents, and scientific writers. Monitoring, feedback, and materials modification take place and are observed directly by administrators of the project. Evaluation procedures are currently underway. The administrators appear to be well informed about all phases of the project, which indicates that there is neither a cumbersome excess nor a shortage of administrative direction. The mangement/organizational flow chart includes job descriptions. NSF has been adequately informed through periodic reports, correspondence, and open communication.

On the face of the materials seen by the panel, management/organization plans are demonstrated to be excellent and well-executed.

Follow-Up Study

Corley (1978), one of the field test teachers, identified seventy-five students who had participated in the testing of the Human Sciences Program at Lansdowne Middle School, 1973-1976, and seventy-five students who were randomly selected from the same graduating class (eighth-grade). Both groups were stratified into the three ability-level sections in accordance with the school's grouping policy.

Fifty-eight Human Sciences students and fifty-three non-HSP students were located at Lansdowne High School. Achievement grades in science for grades nine and ten were compared. A questionnaire was distributed to the former Human Sciences students asking them to rate their Human Sciences experience with their "regular" science experience in six areas. Questionnaire responses were returned by 68 percent of the students.

Corley found that both HSP and non-HSP students had similar patterns of achievement as determined by grades. Although grades were slightly lower for HSP students in accelerated ninth-grade biology, there was no corroboration of an HSP effect in student responses to the questionnaire.

Student ratings of Human Sciences were very positive, with mean values between four and five on a five-point scale for all six items of the questionnaire: amount of science learned, ability to read directions, opportunity for problem solving, knowledge of science equipment, attitude toward science class, and preparation for high school science.

Summary

This chapter has presented data supporting the rationale of the Human Sciences Program. The rationale assumed that students should be provided science materials that would be both interesting and educational for the full range of students in the public schools. The theoretical base, grounded in developmental psychology, proposed that early adolescents were generally not capable of formal operational thought. The curriculum consequence was to initiate activities at various levels of concrete operational logic, where logic was required, and to provide thought-provoking experiences that would enable students to develop and consolidate concrete operational processes and to make the transition to formal operational thought thereafter.

The curriculum rationale also proposed that if science were presented in a social context, with attention to designing experiences that were meaningful to early adolescents, the curriculum would be interesting to them. Clearly, the results of these additional studies lend further support to confirm the success of the curriculum rationale and the materials developed to implement it.

CHAPTER 10

HUMAN SCIENCES, A NEW CURRICULUM DESIGN

The Human Sciences Program was developed and evaluated as an educational product that was adapted to the unique nature of the emerging adolescent learner. The primary goal of the project from its beginning was to relate the curriculum materials and learning approaches as closely as possible to the characteristics of ten- to fourteen-year-olds.

The developers assumed that emphasis on formal, high-level concepts would result in a program that would be too difficult for most students in grades six to eight. Instead, the developers sought to emphasize the "precursors" to science concepts that could be developed experientially. The result of this major development and testing effort is a three-year inter-disciplinary science curriculum for use with ten- to fourteen-~~year-olds in science~~ classes.

Formative evaluation of the experimental materials was conducted in five major phases, as shown in Table 99.

TABLE 99
Summary Chart of the Five Major Phases of the Formative Evaluation of the Human Sciences Program

Materials	Academic Year	Grade Level(s)	No. of Schools	No. of Teachers	No. of Students
3 Pilot Modules	1972-1973	6,7,8	19	18	540
5 Level I Modules	1973-1974	6	7	19	672
4 Level II Modules	1974-1975	7	7	13	490
4 Level III Modules	1975-1976	8	7	10	335
1 Level III Module	Spring 1977	8	12	14	519

The Human Sciences Program divided each school year into sections of from six to nine weeks for which a particular material, each designating a module, was provided. Each module contained everything needed in classes of thirty to forty students for two class groups each day. "Everything needed" included all unique materials, but did not include ordinary laboratory equipment. Experimental modules did include library resources where these were appropriate.

Choice

A key characteristic of the program was the provision of a bounded free-choice environment for students. "Bounded" means that students were asked to stay within the boundaries of the module and activity design of the program and to remain essentially within the activities or activity extensions provided in each module. Each module contained from thirty to over fifty individual activities. Each activity consisted of several pages of printed paterial plus all of the equipment, supplies, and other materials needed to conduct the activity successfully. There were more activities in each module than any single student could complete within the allotted time period. No activities were prescribed.

Students could choose the activities they would do. In some instances the opportunity was provided for students to devise their own activities. The choice of activities made it possible to include many things for students to do that would not be considered feasible in classes where every student is required to do every laboratory or every one of some other kind of activity. Not only could students choose the activities they would do but they could choose whether they wished to work alone, with a partner, or with several other students. This, then, is what is meant by a "bounded, free-choice environment." Every activity in every module was designed to have educational value for some students. Choice was not from the whole world but from the activities in a particular module and usually only from a segment of a module--a problem area--at any one time. Each module was designed around a particular theme. Subdivisions within modules--problem areas or clusters--provided internal organization for closely related activities.

The three pilot modules were deliberately constructed to be as different as possible from contemporary science materials. The rationale for this approach was to challenge staff and writers to be as inventive as possible. The major constraints were the pre-established characteristics of early adolescents. Evaluation studies were to find the program elements that needed revision and to find what would need to be done with the materials to make them "teachable." This was a critical concern sinch the focus of the activities was on students and their needs and concerns, not on teachers' needs and concerns. The assumption was made that as modifications of materials were made, each being influenced by immediate feedback from testing the precursors, it would be easy to adapt materials toward those that were already in the marketplace, but probably very difficult to stretch them in radically new ways.

Evaluation Outcomes

Formative evaluation of Human Sciences was a small part of the total curriculum development effort. It accounted for the expenditure of a small part of the resources allocated to the project. The tension between development of curriculum materials and evaluation developed in this project as it has in many others, with development taking limited resources when cut-backs were ordered.

The summative-formative evaluation distinotion had been adopted by the science education community before the Human Sciences Project was initiated. Similarly, research and evaluation had been completely separated, with most science educators conceptualizing these as unrelated, and in practice, unrelatable events. The need to rethink these separations has been presented elsewhere (Robinson, in press), but the point needs to be emphasized here. Formative evaluation is inadequate for use with innovative curricula. Critics of anything new wish to know how well students achieved in terms of existing achievement criteria. Similarly, school personnel concerned with program effects wish to know how students achieved in terms of criteria they know. Innovative materials must show that they can produce achievement in terms of exisitng criteria, and also produce added value. This requirement makes it imperative that summative and formative evaluation be coupled in the evaluation plans for innovative curricula.

Innovative curricula usually raise new questions about teaching and learning. Many research opportunities have been lost by the disengagement of curriculum development and evaluation from research. This does not mean that those engaged in curriculum development should do research. It does mean that interaction of developers, evaluators, and researchers, with the coupling of their efforts for some development projects, would lead to greater understanding of the processes involved in teaching and learning science.

A second finding of the formative evaluation of Human Sciences was the need to provide resources for student testing and grading. Development and testing of such materials is especially important for curricula based upon the developmental characteristics of learners.

The more general findings from the Human Sciences formative evaluation are listed below. Other more specific findings are presented in sections devoted to those specific topics.

Data from the evaluation studies showed that:

1. Human Sciences was equally effective in different parts of the country, with a variety of teachers, and with students of a wide range of backgrounds and abilities.
2. interdisciplinary studies, selecting content and methodologies from the biological, physical, earth, social, and behavioral sciences, could be accommodated in self-contained classrooms in elementary schools, in departmentalized elementary and secondary schools, and in team-teaching contexts in middle and junior high schools.
3. activity choices of ten- to fourteen-year-olds were not clustered

by content or difficulty, nor were they influenced significantly by the sex of the students.

4. students could learn to manage an environment (the classroom) even with scarce resources and a student overpopulation.
5. reading, writing, and arithmetic skills were used meaningfully in contexts where students needed them to solve problems of their choice.
6. students can improve their skills in self-direction, with decreasing need for continuous supervision, in a bounded free-choice environment.
7. activities which were potentially controversial drew parent support more consistently than administrative support, or support by department chairpersons.
8. the assumption that activities should be introduced in concrete ways, not yet requiring formal operational competence, was a correct assumption in terms of the competence of early adolescents in logical thinking.

Parents, teachers, administrators, and students validated:

1. the necessity for students at the middle school/junior high school level to develop skills such as observing, inferring, managing, and other such skills judged to be collectively as important as reading, writing, and arithmetic.
2. the study of human beings as a legitimate subject of study in middle school/junior high school science classes.
3. the importance of providing students with opportunities to discuss, question, interview, observe, and in other ways interact with adults and young children, as well as with the students' peers, as part of a science program for middle schools/junior high schools.
4. the role of student experience with choice and evaluation of activities in contributing to decision-making skills.

Testing a variety of options in the Human Sciences Program showed that:

1. activity-specific evaluation problems were most effective for evaluating student performance when they included both student self-evaluation and tasks for external evaluation.
2. activities with structure and explicit directions were more useful to, and more used by, students than other activities which provided only suggestions for what might be done.
3. introductory prose should be limited so that students can get actively engaged in <u>doing</u> something early in an activity. Exposition, clarification, and elaboration can follow as needed.
4. activities selected as most valuable by students usually required action and thought together. Where the "doing" of an activity could be done without thinking, postponing reasoning until later, student ratings were lower.
5. activities had limited appeal for most students where reading was the <u>only</u> form of "doing."
6. students worked best with activities they could do alone or with one partner. Other activities requiring a group of three or more students were generally beyond the group skills of early adolescents.

7. activities with living things were popular and valuable for students except when they required formal logical competence or a well-developed conceptual scheme for their accomplishment.
8. student achievement on multiple-choice and essay test items, given the classroom environment of options in activity choices, was at the 50 percent level of success--the predicted level. This success level was desirable for evaluation purposes, but was considered by the students to be too difficult in relation to their feelings of success or failure.

Student Attitudes

Data from three administrations of attitude measures, from activity evaluation data, and from one study using a comparison group showed that Human Sciences resulted in student attitudes that were:

1. highly positive toward the Human Sciences Program at the end of field testing of a single module.
2. highly positive toward the Human Sciences Program at the end of field testing of five modules in one year.
3. positive toward the Human Sciences Program at the end of all field testing for three years, and more positive toward Human Sciences than a comparison group's attitudes were toward regular science classes in the same schools.
4. more positive for girls than for boys toward Human Sciences, by contrast with the comparison group in which the attitudes of girls were more negative than for boys toward regular science classes.

Logical Competence

The assumption that most early adolescents would not be capable of formal operational thought, and indeed, that there would be many who had yet to develop concrete operational thought was confirmed empirically by the results of the 1974 and 1976 testing with "How Is Your Logic?" In 1974 over one-third of the sixth-graders gave pre-operational responses to concrete operational problems. At the same time, over 85 percent of the eighth graders were giving concrete responses to formal operational problems. By the end of the eighth grade, 82 percent of the students who had been in Human Sciences for the three-year test period were giving concrete responses to concrete problems and nearly 28 percent were giving formal responses to formal operational problems. These data are consistent with other research that finds most students enter ninth grade with concrete, but not formal, operational competence.

Teachers and Human Sciences

The field tests showed that with limited orientation and few resource management materials, teachers could learn to:

1. teach science classes in which many different activities are going on at the same time.
2. treat students as individuals with unique needs and skills, and assist them in selecting the most effective combination of activities for their growth and development.
3. devise cooperative evaluation and grading programs with students.
4. solve classroom management problems and work with students to help them make individualized curricular choices in a complex multimedia program; the end result of which is a highly valuable learning experience for both students and teachers.

Activity Structure

Activities were constructed with both explicit instructions and with rather general instructions as to what was expected of the student. For example, activities suggested, "Here are some things you might want to keep track of." Suggestions were given within activities for data collection tables, but none were provided, nor were complete models of tables, charts, or other organizational aids included for data gathering and display. Questions were imbedded in the activity, but were left to the student as to their use. The guiding rationale for the early activity design was to test a variety of forms, but to provide opportunities to handle activities in a variety of ways, and to encourage investigation, curiosity, and creativity. Through many incremental steps and with much discussion, classroom observation, and evaluation, the final activity structure was developed. This structure is exemplified into many forms in the experimental edition of KNOWING, the last module to be developed and tested.

First, short introductory paragraphs designed to enable students to decide whether or not they wanted to take time to do the activity were found to be an effective way to begin. Many activities were designed with introductory paragraphs only about five lines long. The task of writers was to attempt to interest as many students in choosing the activity as possible.

Second, action--preferably activities action, dealing with the manipulation of objects and/or events--was found to be more attractive to more students than other alternatives. Action could not follow long pages of exposition. If exposition was necessary in an activity, it should follow rather than precede the action. When exposition followed action, students had good experiential reasons for giving their attention to exposition.

Third, directions needed to be illustrated wherever possible to help students understand what they were to do in the activity.

Many kinds of action, or "doing," were provided within activities in order to accommodate the wide variety of styles that students preferred. Table 100 shows the major kinds of things students did in activities. Many activities contained several of these "doings."

TABLE 100
What Students Are Doing in Human Sciences

Appreciating	Information gathering	Nurturing
Constructing	Interviewing	Reading
Creating	Listening	Valuing
Deciding	Listening/Watching	Watching/Viewing
Experimenting		

They are essentially self-explanatory. Construction--building pieces of equipment or constructing objects--is one kind of manipulation of physical objects. The care of plants and animals became a necessary environmental management skill for students. Just caring for organisms, keeping them alive and healthy, was considered a legitimate science activity for students, rather than for teachers. "Information gathering" was distinguished from "experimenting" in that in many activities information was gathered in naturalistic settings without attention to sampling or control of variables. Activities included worksheets, which in many cases contained additional content and procedures. For most activities, question had specific, correct responses. Open-ended problems were also retained in some activities to enable students to develop their own explanations, to compare them with other students, to represent their ideas in language as much as possible, and to initiate challenges by other students so that they had to give reasons for their interpretations. The student-student interchange has become an important part of the design of Human Sciences activities.

Figure 31, page 186, shows three kinds of activity emphases that were found to be important for early adolescents. Cognitive purposes are common to every science material. The means used to achieve cognitive purposes will yield meta-learnings (unplanned learnings) that can be taken into account, such as feelings about the object, feelings about science itself, feelings about oneself, and what one can and cannot do. Science activities for this age group should legitimately have craftsmanship as a major purpose. This does not mean that cognition is excluded, but in some cases the cognitive outcome may not be realized. The building--the production of a model, a piece of equipment, or another object--is sufficient for some students. Finally, affective purpose--developing empathy and appreciations--is also considered legitimate as a primary purpose for activities in Human Sciences. Activities that met all of these purposes were the most effective activities, but they were difficult to invent. Evaluation is also an integral part of the structure of each activity. Each activity has a specific set of evaluation problems. With most activities, both essay and objective problems are included in the evaluation problems set to ensure that early adolescents have as much opportunity as possible to provide explanations in written language.

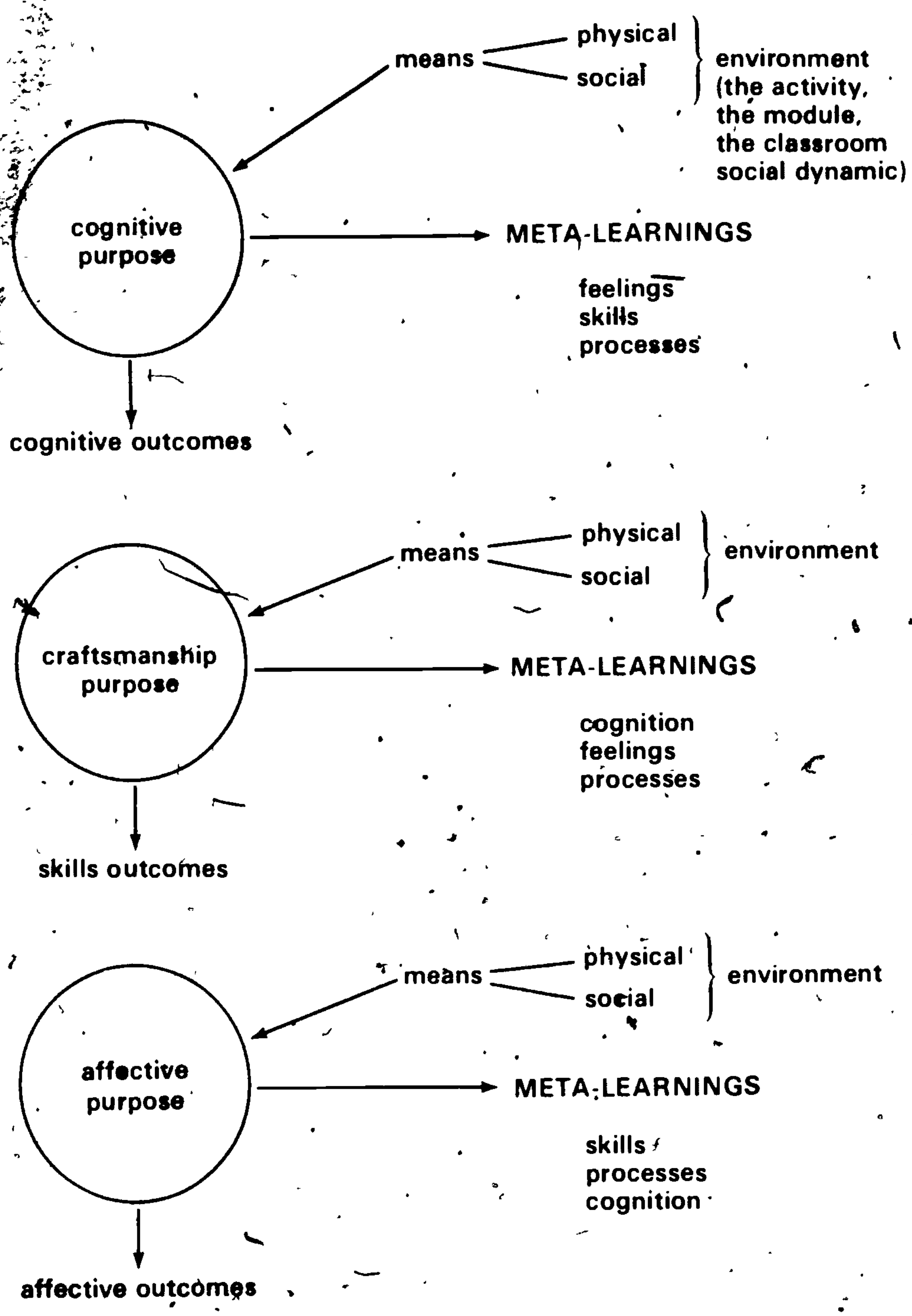


Figure 31. Three kinds of Human Sciences activities.

The final characteristic of activities relates to responsibility and management. Students choose an activity; are responsible for keeping records of activities done; and make choices as to whether they should choose another activity and start on it while waiting for plants to develop. This requires the development of management skills not usually well developed in early adolescence. Such skills are a necessary part of coping in the contemporary world and should be part of a science program. When the activity is completed, the student is responsible for securing a copy of the evaluation problems, completing them, returning them to the teacher, and taking apart and returning all parts of the activity that can be recycled back into the container. This ensures that the next student choosing that activity will have similar opportunities to learn. Evaluation data gathered throughout the field test period indicated that these management skills were improved in test classes but did not reach the desired levels.

Summary

The Human Sciences Project was given the challenge of producing an entirely new science curriculum that would take into account the developmental characteristics of early adolescents. The Human Sciences staff accepted this challenge, asking the question, "How can the sciences contribute to the development--cognitive, psycho-social, and personal--of early adolescents?" This question turned out to be a more profound question than when it was initially proposed.

The resulting curriculum rationale broke with traditional materials in many significant ways. The materials selected content from both the natural and social sciences to produce an interdisciplinary product. The materials produced were modular, with individual activities that were not assigned, but that students chose. A major intent of the design of the materials was to enable students to develop resource management skills in an over-populated, limited resource environment in the Human Sciences classroom. Activities were provided in limited quantities so that all students could not choose the same activity at the same time. Students were responsible for record-keeping and evaluation of their own work and of the materials.

The program continually used evaluation data in constructing the materials so that the structure of modules, activities, evaluation materials, and procedures was in continuous evolution over the life of the project. The final module, KNOWING, was markedly different than the first module produced.

Achievement, cognitive, attitudinal, and observational data gathered during the formative evaluation illustrated problems, raised new questions, and confirmed a priori assumptions of the rationale. This self-correcting process has demonstrated that science materials can be produced that instruct early adolescents and contribute to their achievement in science, their cognitive development, and their positive attitude toward science courses. The self-correcting process also demonstrated that teachers can learn to teach with such materials. For some teachers, the materials were natural, something they had hoped would come along. For others, the diversity of activities and choice was confusing, sometimes threatening. Success varied a great deal. Although teacher differences have yet to be studied, a casual review of the data shows the teacher variable to be important. Clearly, this curriculum program is not for all teachers. But for those teachers who came to understand the rationale and the facilitative role of the teacher, this program was both challenging and valuable.

Reports of a few parents whose children had moved on into high school indicated that Human Sciences was a most valuable experience. They reported that their children had learned how to learn science and wanted to learn more.

One three-year student appeared at the BSCS headquarters in Boulder in the fall of 1981. She was a sophomore at the University of Colorado, had been an honor student with four years of mathematics and science in high school, and was earning A and B grades in college. Her comment was, "Human Sciences was the most valuable experience of my life."

BIBLIOGRAPHY OF PUBLICATIONS FROM THE HUMAN SCIENCES PROJECT

Published Materials

SP 7409-12. Robinson, J. T. Integrative curricula, a response to Peter Dow's "Science, schooling and society: The search for an integrated curriculum. In Current activities and the task ahead. Report of the National Science Foundation PES/MIDS Project Directors Meeting, Airlie House, Warfenton, Va., September 13-15, 1974.

SP 7211-55. Eckenrod, J. E. Looking ahead. Social Education, 1972, 36(7), 772.

SP 7400-97. Robinson, J. T., Rasmussen, F. A., Eckenrod, J. E., McConnell, M. C., and Seitz, A. C. Resource book for BSCS Human Sciences Dissemination & Implementation team members. Boulder, Co.: BSCS, 1974.

SP 7408-99. Robinson, J. T., Rasmussen, F. A., Eckenrod, J. E., McConnell, M. C., and Seitz, A. C. Excerpt from Teachers guide for RULES, a module of the BSCS Human Sciences Program, Experimental Edition, 1974. Boulder, Co.: BSCS, 1974.

SP 7408-107. (1) Rasmussen, F. A. Human Sciences seventh-grade writing conference, summer 1974. Newsletter article, August 5, 1974. (2) Eckenrod, J. E. Human Sciences implementation project. Newsletter articles, August 1974.

SP 7302-111. Robinson, J. T., Rasmussen, F. A., and Eckenrod, J. E. Human Sciences: A developmental approach to adolescent education. Boulder, Co.: BSCS, 1973.

SP 8111-120. Mayer, W. V., and Hurd, P. D. The middle school. BSCS Newsletter, 1969, 34, 1-9.

SP 8111-121. Clark, G. M. (Ed.) Life sciences in the middle school. BSCS Special Publication #7. Boulder, Co.: BSCS, 1969.

SP 8111-122. Robinson, J. T., and Rasmussen, F. A. Human Sciences for the middle school. BSCS Newsletter, 1971, 44, 13-16.

SP 8111-124. Robinson, J. T., Rasmussen, F. A., and Eckenrod, J. E. Human Sciences for the middle school, progress report #2. BSCS Newsletter, 1972, 48, 7-9,19.

SP 8111-125. Robinson, J. T., Rasmussen, F. A., and Eckenrod, J. E. A preview of Human Sciences: A BSCS multidisciplinary program. BSCS Newsletter, 1973, 50, 1-11.

SP 8111-126. Robinson, J. T., Rasmussen, F. A., and Eckenrod, J. E. Human Sciences. BSCS Newsletter, 1973, 52, 8-9.

SP 8111-127. Robinson, J. T., Rasmussen, F. A., and Eckenrod, J. E. Human Sciences demonstration module ready. BSCS Newsletter, 1973, 53, 1.

SP 8111-129. Ross, N. M., Robinson, J. T., Rasmussen, F. A., Eckenrod, J. E., McConnell, M. C., and Seitz, A. C. The Human Sciences Program: Work in progress. BSCS Newsletter, 1974, 56, 8-9.

SP 8111-130. Ross, N. M., Robinson, J. T., Rasmussen, F. A., Eckenrod, J. E., McConnell, M. C., and Seitz, A. C. Implementing Human Sciences. BSCS Newsletter, 1974, 57, 8-9.

SP 8111-131. Ross, N. M., Robinson, J. T., Rasmussen, F. A., Eckenrod, J. E., McConnell, M. C., and Seitz, A. C. Human Sciences introductory package: Packets 1-5. Boulder, Co.: BSCS, 1974.

SP 8111-132. Ross, N. M., Robinson, J. T., Rasmussen, F. A., Eckenrod, J. E., McConnell, M. C., and Seitz, A. C. With learning in mind: Human Sciences Program teachers guide. Boulder, Co.: BSCS, 1974.

SP 8111-133. Hurd, P. D., Nickelson, A. L., Ross, N. M., Robinson, J. T., Rasmussen, F. A., Eckenrod, J. E., McConnell, M. C., and Seitz, A. C. The Human Sciences Program: Evolution, current status, a sampling of activities, first reported research, and experimental delivery system. BSCS Newsletter, 1976, 63, 1-5,8-15,18-20.

SP 8111-134. Hurd, P. D. The BSCS Human Sciences Curriculum Project. Bioscience, 1978, 28(1), 36-38.

SP 8111-135. Asp, E. Getting parents involved. The Middle School Junior High Science Bulletin, 1978, 1(1), 4.

SP 8111-136. Ross, N. M., Hurd, P. D., Robinson, J. T., Rasmussen, F. A., Eckenrod, J. E., McConnell, M. C., and Seitz, A. C. An update: The Human Sciences Program. BSCS Journal, 1978, 1(2), 1-23.

SP 8111-137. McConnell, M. C. Interdisciplinary curriculum. Middle School Journal, 1979, 10(2), 10-11.

SP 8111-138. McConnell, M. C. Human Sciences Program: Speaking to where students "live." The Science Teacher, September 1978, 20-22.

SP 8111-139. Ross, N. M., Robinson, J. T., Rasmussen, F. A., Eckenrod, J. E., McConnell, M. C., and Seitz, A. C. The Human Sciences Program: A focus on Level III. BSCS Journal, 1980, 3(2), 1-20.

SP 8112-140. Ross, N. M., McConnell, M. C., and Robinson, J. T. Teaching Human Sciences. Boulder, Co.: BSCS, 1979.

SP 8112-145. Ross, N. M. The Human Sciences Program: A history. Boulder, Co.: BSCS, 1981. (Available from the ERIC Clearinghouse for Science, Mathematics, and Environmental Education.)

Unpublished Papers

SP 7410-1. Robinson, J. T. Clarification in communications. October 22, 1974.

SP 7408-6. Robinson, J. T., Rasmussen, F. A., Eckenrod, J. E., McConnell, M. C., and Seitz, A. C. Planning for Human Sciences: A resource paper for seventh-grade teachers. September 1974.

SP 7310-8. Rasmussen, F. A. Some preliminary thinking on helping teachers: Implementing Human Sciences. October 1, 1973.

SP 7310-10. Robinson, J. T. How do students learn how to learn? October 4, 1973.

SP 7111-11. Robinson, J. T., Rasmussen, F. A., and Eckenrod, J. E. Human Sciences: Student questions. Working papers and notes from HSP writers of questions asked by early adolescents, Spring 1973.

SP 7303-13. Robinson, J. T. Grading. Preliminary drafts, September 1972.

SP 7503-14. Robinson, J. T. Research priorities for science-related social issues. Paper presented at the AERA Annual Meeting Symposium on Teaching of Science-Related Social Issues: Research Priorities, Washington, D. C., March 1975.

SP 7311-15. Robinson, J. T. Inservice workshop. Memo from B. P. Ebersole to State Superintendent of Schools, Balton / Baltimore Counties, May 1975.

SP 7503-16. Robinson, J. T. Considerations for including science-related social issues in science classes. Paper presented at the NARST Panel, "Do science-related social issues belong in science classes?", Los Angeles, Ca., March 17, 1975.

SP 7503-17. Robinson, J. T. Teaching of science-related social issues. Paper presented at the AERA Annual Meeting Symposium, Washington, D. C., March 31-April 4, 1975.

SP 7504-18. Flexer, A. S., and Robinson, J. T. Teaching of science-related social issues. Paper presented at the AERA/SIG Symposium on Research Priorities, Washington D. C., April 1, 1975.

SP 7503-19. Robinson, J. T. Opportunities for research in BSCS curriculum development projects. Paper presented at the NARST 48th Annual Meeting Symposium on Natural laboratories: Opportunities for research in classrooms using new curricula, Los Angeles, Ca., March 19, 1975.

SP 7506-20. Robinson, J. T. Comments regarding "National assessment of science, 1969 and 1973." No date.

SP 7406-21. Robinson, J. T. A theoretical framework for interdisciplinary science curricula for early adolescents: The BSCS Human Sciences Program. September 5, 1974.

SP 7410-22. McConnell, M. C., and Robinson, J. T. Some thoughts on the Level III modules. Draft III, October 30, 1974.

SP 7404-23. Robinson, J. T., Values in curriculum decisions. Invited paper, University of Colorado School and College Conference, Boulder, Co., April 18, 1974.

SP 7308-24. Various authors. Reader reactions to Human Sciences: A developmental approach to adolescent education. 1973.

SP 7203-27. Robinson, J. T. Eighth-grade modules. March 14, 1972.

SP 7203-28. Robinson, J. T. Seventh-grade modules. March 9, 1972.

SP 7109-29. Robinson, J. T., Rasmussen, F. A., and Eckenrod, J. E. Content charts for the Human Sciences Program. June 1972.

SP 7412-31. Eckenrod, J. S., Hurd, P. D., Rasmussen, F. A., and Robinson, J. T. A multidisciplinary Human Sciences Program for middle schools: A preliminary statement. September 1971.

SP 7511-32. Robinson, J. T. Video-taping ideas for the Queens D&I demonstration classroom. November 7, 1975.

SP 7407-33. Robinson, J. T., Rasmussen, F. A., Eckenrod, J. E., McConnell, M. C., and Seitz, A. C. HSP teachers guide. November 19, 1974.

SP 7203-34. Rasmussen, F. A. Papers that might be helpful in thinking about multidisciplinary education. March 21, 1972.

SP 7203-35. Robinson, J. T. Piaget's developmental psychology: Readings. March 21, 1972.

SP 7501-36. Robinson, J. T. The seventh-grade program: A working paper. February 1974.

SP 7501-37. Robinson, J. T., Rasmussen, F. A., Eckenrod, J. E., McConnell, M. C., and Seitz, A. C. HSP--Conservation of volume, task 1. January 1975.

SP 7405-39. Robinson, J. T., Rasmussen, F. A., and Eckenrod, J. E. Sample items for workshops: "How is your logic?" (1974 Edition). 1974.

SP 7503-42. Eckenrod, J. E. D&I: Piaget's stages of intellectual development. A summary for conference handout, June 1974.

SP 7503[illegible]rod, J. E. D&I: Loevinger's stages of ego development. A su[illegible]nference handout, June 1974.

SP 7503-44[illegible]od, J. E. D&I: Kohlberg's stages of moral growth. A summary for c[illegible]nce handout, June 1974.

SP 7602-47. Flexer, A. S., and Robinson, J. T. A working outline of the Level III module SURROUNDINGS. February 1976.

SP 7409-49. Robinson, J. T. Can you live wi[illegible] child's explanation? June 1974.

SP 7301-51. Schlitt, D. M. How do students see Me?: An activity for teachers. January 1973.

SP 7310-52. Schlitt, D. M. What do teachers do? I A working paper, October 1973.

SP 7407-53. Eckenrod, J. S. The Human Sciences Program is for kids. Final draft of article for Learning magazine, September 1974.

SP 7605-54. McConnell, M. C. Human Sciences Program objectives. May 24, 1976.

SP 7209-56. Robinson, J. T., Rasmussen, F. A., and Eckenrod, J. E. Human Sciences for the middle school progress report. September 1972.

SP 7407-57. Thompson, C., and Puckett, D. Commentary on the activity "Boy or girl?" Interoffice memo, July 27, 1975.

SP 7503-58. Eckenrod, J. S. Report of the model of curriculum diffusion for the BSCS Human Sciences Program. Paper presented at the Annual Meeting, American Education Research Association, Washington, D. C., April 1975.

SP 7303-59. Robinson, J. T. The early adolescent. March 15, 1973.

SP 7310-60. Robinson, J. T., Rasmussen, F. A., and Eckenrod, J. E. For parents: A working paper about Human Sciences. October 1973.

SP 7704-61. Ross, N. M. Questions and answers about Human Sciences: A parent guide. April 1977.

SP 7410-62. Robinson, J. T., Rasmussen, F. A., Eckenrod, J. E., McConnell, M. C., and Seitz, A. C. GROWING module. October 4, 1974.

SP 7410-63. Robinson, J. T., Rasmussen, F. A., Eckenrod, J. E., McConnell, M. C., and Seitz, A. C. LEARNING module. October 4, 1974.

SP 7410-64. Robinson, J. T., Rasmussen, F. A., Eckenrod, J. E., McConnell, M. C., and Seitz, A. C. SURVIVAL module. October 4, 1974.

SP 7410-65. Robinson, J. T., Rasmussen, F. A., Eckenrod, J. E., McConnell, M. C., and Seitz, A. C. BEHAVIOR module. October 4, 1974.

SP 7410-66. Robinson, J. T., Rasmussen, F. A., Eckenrod, J. E., McConnell, M. C., and Seitz, A. C. SENSE...OR NONSENSE? module. October 4, 1974.

SP 0000-68. Robinson, J. T., Rasmussen, F. A., and Eckenrod, J. E. Inservice training for curriculum dissemination. March 1973.

SP 7210-69. Robinson, J. T., Rasmussen, F. A., and Eckenrod, J. E. An explanatory note for teachers. A working paper, October 1972.

SP 7302-70. Robinson, J. T., Rasmussen, F. A., and Eckenrod, J. E. Rationale and program development plans for Human Sciences. March 1972.

SP 7204-71. Robinson, J. T., and Rasmussen, F. A. Scripts for Piaget (Belanger) kinescopes made in Jefferson County from Hal Pratt. April 13, 1972.

SP 0000-72. Belanger, M. An introduction to Piaget's theory of mental development. Edited transcript of lecture, February 1972.

SP 7410-73. Flexer, A., McConnell, M. C., and Robinson, J. T. Descriptions of Level III modules. October 30, 1974.

SP 7402-74. Robinson, J. T. Community involvement in curriculum change. Drafts of article for journal or BSCS Newsletter, December 21 and 31, 1973.

SP 7405-75. McConnell, M. C. An annotated bibliography related to a multidisciplinary developmental approach to adolescent education. May 8, 1974.

SP 7511-76. Eckenrod, J. S. Human Sciences implementation activities. November 19, 1975.

SP 7411-77. Eckenrod, J. S. Developing tomorrow's curricula (and making sure they stay off the shelf). Paper presented at the College and University Faculty Assembly Symposium on Disseminating Tomorrow's Curricula Today, National Council for the Social Studies, Chicago, Il., November 27, 1974.

SP 0000-78. Ross, N. M. Human Sciences Program (HSP). June 1976.

SP 7308-79. Robinson, J. T. A theoretical framework for multidisciplinary science curricula for early adolescents. August 13, 1973.

SP 7505-80. Eckenrod, J. S. Middle grade science program is on the way. Draft of article to be published in Learning magazine, May 20, 1975.

SP 7301-81. Robinson, J. T. What is a module? Drafts of summer conference paper, January 25, 1973 and June 1974.

SP 7406-82. Robinson, J. T. What is an activity? Summer conference paper, June 1974.

SP 7404-83. Seitz, A. C. Checkpoints for a module. Summer conference paper, April 15, 1974.

SP 0000-84. Robinson, J. T. Module descriptions, seventh grade. April 1973.

SP 7406-85. Robinson, J. T., Rasmussen, F. A., and Eckenrod, J. E. Organization for the conference. Summer conference paper, 1974.

SP 7210-86. Hurd, P. D. The developmental psychology of Jean Piaget and the learning of science by children, with an addendum on concrete and formal operations. 1972.

SP 7210-87. Hurd, P. D. The diffusion and implementation of science curricula: A planning and developmental proposal. December 1972.

SP 7207-88. Hurd, P. D. Human Sciences for the middle school: A BSCS project. Paper presented at the NSF Conference for Elementary and Middle School Administrators, Purdue University, Lafayette, In., July 24, 1972.

SP 7010-89. Hurd, P. D. Biology: A study of man and society. Paper presented at the National Association of Biology Teachers Convention, Denver, Co., October 23, 1970.

202

SP 7007-90. Hurd, P. D. The changing philosophy of secondary school science. Paper prepared for the AAAS Conference, Annapolis, Md., July 29-30, 1970.

SP 7207-91. Hurd, P. D. Human Sciences for the middle school: Our changing way of living. July 19, 1972.

SP 7204-92. Hurd, P. D. Notes on problem solving derived from research findings (including copies of 7 related papers). September 1972.

SP 7208-93. Hurd, P. D. Human Sciences for the middle school: Principal's Guide. May 25, 1972.

SP 7201-94. Hurd, P. D. A profile of the emerging adolescent. Unpublished paper, January 1972.

SP 7404-95. Robinson, J. T. A theoretical framework for multidisciplinary science curricula for early adolescents: The BSCS Human Sciences Program. Paper presented at the AERA Convention, Chicago, Il., April 16, 1974.

SP 7212-96. Hurd, P. D. Review of the professional literature on curriculum implementation and diffusion. December 1972.

SP 7409-98. Robinson, J. T., Rasmussen, F. A., Eckenrod, J. E., McConnell, M. C., and Seitz, A. C. Planning for Human Sciences: A resource paper for seventh-grade teachers. September 1974.

SP 7303-100. Robinson, J. T., Rasmussen, F. A., and Eckenrod, J. E. Experimental modules for Human Sciences, 1972. March 15, 1973.

SP 0000-101. Eckenrod, J. E. Modes of learning. Adaptation of an article by J. Coleman, April 1973.

SP 7500-102. Robinson, J. T. Human Sciences Program - projected developmental and dissemination activities - 1975-1977. Flowchart, December 1974.

SP 7402-103. Robinson, J. T. On developing objectives. February 19, 1974.

SP 7310-104. Robinson, J. T., Rasmussen, F. A., and Eckenrod, J. E. Human Sciences, seventh- and eighth-grade module descriptions: A working paper (A developmental approach to adolescent education). October 1973.

SP 7203-105. Robinson, J. T., Rasmussen, F. A., and Eckenrod, J. E. Objectives and modules for Human Sciences. March 1972.

SP 7400-106. Robinson, J. T., Rasmussen, F. A., and Eckenrod, J. E. With learning in mind: A working paper about Human Sciences. Camera copy, 1974.

SP 7200-108. Robinson, J. T., Rasmussen, F. A., and Eckenrod, J. E. Module overview for DEVELOPING. 1972.

SP 7200-109. Robinson, J. T., Rasmussen, F. A., and Eckenrod, J. E. Module overview for HUMANSELF. 1972.

SP 7200-110. Robinson, J. T., Rasmussen, F. A., and Eckenrod, J. E. Module overview for LEARNING. 1972.

SP 7412-112. Robinson, J. T., Rasmussen, F. A., Eckenrod, J. E., McConnell, M. C., and Seitz, A. C. The Human Sciences Program. Special Publication #2. 1974.

SP 7610-114. Robinson, J. T. Analysis of secondary school enrollments in Biology I and BSCS biology, 1972-73. October 27, 1976.

SP 7600-117. Ross, N. M. The Human Sciences module cart: Plans for manufacturing wooden module carts for Human Sciences modules. 1976.

SP 7703-118. Robinson, J. T. Design for middle school science materials. Paper prepared for the NSTA Annual Convention, Cincinnati, Oh. March 28, 1977.

SP 7710-119. Robinson, J. T. Designing science programs to enhance intellectual development of emerging adolescents. Invited paper presented at the Conference on Intellectual Development, Research and Education, Southern Illinois University at Edwardsville, October 11, 1977.

SP 8111-123. Robinson, J. T., Rasmussen, F. A., and Eckenrod, J. E. Human Sciences progress report. April 1972.

SP 8112-141. BSCS. Record and transcript of Middle School Life Sciences Conference I (Human Sciences Project), March 18-21, 1971, Harvest House Motor Hotel, Boulder, Co., March 1971.

SP 8112-142. BSCS. Record and transcript of Middle School Life Sciences Conference II (Human Sciences Project), April 15-18, 1971, Royal Inn of Boulder, Boulder, Co., April 1971.

SP 8112-143. BSCS. Record and transcript of Middle School Life Sciences Conference III (Human Sciences Project), May 13-16, 1971, Harvest House Motor Hotel, Boulder, Co., June 1971.

SP 8112-144. Cleaver, T. J. Toward a preliminary rationale for development of an interdisciplinary life science program for the middle school, a position paper (Human Sciences Project). May 10, 1971.

BIBLIOGRAPHY OF EVALUATION PAPERS FROM THE HUMAN SCIENCES PROJECT

SP 7410-2. Robinson, J. T. Evaluation goals, first test class visits. Unpublished manuscript, Fall 1974.

SP 7502-3. Robinson, J. T. Ego development. Unpublished manuscript, Fall 1974.

SP 7502-4. Robinson, J. T. Formative evaluation of Human Sciences, 1972-73. Unpublished manuscript, February 18, 1975.

SP 7503-5. Robinson, J. T. Hierarchical relationships of items in "How is your logic?" Unpublished manuscript, Spring 1975.

SP 7405-7. Gray, W. M., and Robinson, J. T. How is your logic? (1974 Edition). Unpublished manuscript, May 1974.

SP 7310-9. Steele, J. Speaking of evaluation...alternatives to grading students. Unpublished manuscript, October 4, 1973.

SP 7412-25. Eckenrod, J. S. Human development goals of the Human Sciences Program. Unpublished manuscript, December 20, 1974.

SP 7503-26. Steele, J. M. A longitudinal and comparative look at cognitive development in EMH children. Paper presented at the Annual Meeting of the American Educational Research Association, Washington, D. C., March 31, 1975.

SP 7502-38. Robinson, J. T., Rasmussen, F. A., Eckenrod, J. E., McConnell, M. C., and Seitz, A. C. Analysis of selected items: "How is your logic?" (1974 Edition). Unpublished manuscript, March 17, 1975.

SP 7509-40. Robinson, J. T. Sample protocols for "Choose your problem." Unpublished manuscript, September 25, 1975.

SP 7509-41. Robinson, J. T. Human Sciences evaluation, 1974-75. Unpublished manuscript, December 1974.

SP 7512-45. Robinson, J. T. Evaluation of Human Sciences, Level III. Unpublished manuscript, December 12, 1975.

SP 7601-46. Robinson, J. T. Development of self-evaluation for the Human Sciences Program. Unpublished manuscript, January 23, 1976.

SP 7505-48. Robinson, J. T. Summary report, Human Sciences Program Level III Evaluation Conference. Unpublished manuscript, May 31, 1975.

SP 7602-50. Robinson, J. T. Self-evaluation systems for Human Sciences. Draft of <u>BSCS Newsletter</u> article, December 1975.

SP 7303-67. Robinson, J. T., and Rasmussen, F. A. Grading: A working paper for Human Sciences. Unpublished manuscript, October 1972.

SP 7500-113. Robinson, J. T., Rasmussen, F. A., Eckenrod, J. E., McConnell, M. C., and Seitz, A. C. Evaluation of RULES, a Level II Human Sciences module. Unpublished manuscript, 1975.

SP 7603-115. Robinson, J. T. Evaluation of the Human Sciences Program. Draft of BSCS Newsletter articles, March 2, 1976.

SP 7410-116. Robinson, J. T., Rasmussen, F. A., Eckenrod, J. E., McConnell, M. C., and Seitz, A. C. Summary report of student interviews conducted in October 1974 in Human Sciences test school. Unpublished manuscript, 1974.

EP 7504-01. Gray, W. M. An integration of Piaget and criterion-referenced measurement. Paper presented at the annual meeting of the American Educational Research Association, Washington, D. C., April 1975.

EP 7604-02. Gray, W. M. Confirmation of the Piagetian logic of exclusion and combinations during concrete and formal operations. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April 1976.

EP 7410-03. Gray, W. M. Analysis of "How is your logic?" Consultant report, October 22, 1974.

EP 7404-04. Steele, J. M. The relationship of developmental level with science performance: A case for an alternative to IQ grouping of mildly retarded children? Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, Il., April 17, 1974.

EP 7610-05. Havighurst, R. J. Comments on seven modules from the Human Sciences Program of the BSCS. Unpublished manuscript, October 1976.

EP 7506-06. Hofmann, R. J., and Gray, W. M. Multiple hierarchical analysis: Its application and interpretation in the identification of developmental sequences. Paper presented at the Fifth Annual Symposium of the Jean Piaget Society, Philadelphia, June 1975.

EP 7611-07. Gega, P. C. Directions--elementary school science. Teacher, November 1976, 59-61.

EP 7405-08. Robinson, J. T., and Eckenrod, J. S. Summary report of the data collected by means of the questionnaire evaluation of Human Sciences (incomplete). Unpublished manuscript, December 1976.

EP 7403-09. Schlitt, D. Tabulations and comments from the Human Sciences questionnaire, regarding the student journal. Unpublished manuscript, March 21, 1974.

EP 7412-10. Seitz, A. C., and Rasmussen, F. A. Interviews with and comments from students. Unpublished manuscripts, October-November 1973.

EP 7412-11. Seitz, A. C. Report of a parent meeting and summary of parent questionnaire responses from one test school. Unpublished manuscript, April 1, 1974.

EP 7412-12. Various authors. Letters to BSCS about Human Sciences. Unpublished manuscripts, 1973-74.

EP 7412-13. Robinson, J. T. Reports of classroom observations. Unpublished manuscripts, October 1, 1973.

EP 7706-14. Robinson, J. T. Student attitudes towards science courses in test schools using Human Sciences. Unpublished manuscript, June 10, 1977.

EP 7706-15. Robinson, J. T. Student attitudes towards science courses in test schools using the Human Sciences Program: Summary report. Unpublished manuscript, June 13, 1977.

EP 7706-16. Robinson, J. T., Martin, T. S., Cobern, W. W. Evaluation of CHANGE, a Level III Human Sciences module. Unpublished manuscript, June 29, 1977.

SP 7712-17. Robinson, J. T., and Cobern, W. W. Validity of "How is your logic?" Unpublished manuscript, April 17, 1979.

EP 7704-18. Ross, N. M., Ladd, R., McConnell, M. C., and Robinson, J. T. Summary report: HSP public review conference. Unpublished manuscript, April 1977.

EP 7704-19. Ross, N. M., Ladd, R., McConnell, M. C., Robinson, J. T. Summary report: HSP content review conferences. Unpublished manuscript, April 1977.

EP 7708-20. Robinson, J. T. A preliminary comparison, three Level III modules. Unpublished manuscript, August 29, 1977.

EP 7704-21. Ladd, R. Activity evaluation form. Unpublished manuscript, March 1977.

EP 7805-22. Robinson, J. T. A study of the most highly used activity parts and the less used parts of the same activities in INVENTION, a Level III Human Sciences module. Unpublished manuscript, May 25, 1978.

EP 7110-23. Dexter, L. A. Comments on "A multidisciplinary Human Sciences Program for middle schools: A preliminary statement." Consultant report, October 14, 1971.

EP 7110-24. Bohannan, P. Comments on "A multidisciplinary Human Sciences Program for middle schools: A preliminary statement." Consultant report, October 21, 1971.

EP 7110-25. Shaver, J. P. Evaluation of "A multidisciplinary Human Sciences Program for middle schools: A preliminary statement." Consultant report, October 22, 1971.

EP 7110-26. Markley, O. W. Comments on "A multidisciplinary Human Sciences Program for middle schools: A preliminary statement." Consultant report, October 28, 1971.

EP 7211-27. Scharf, P. A developmental educator's comments on HUMANSELF, DEVELOPING, and LEARNING. Consultant report, November 11, 1972.

SP 7408-28. Smith, T., and Ideation Resource Consultants. Phase I of HSP evaluation study. Unpublished manuscript, August 7, 1974.

EP 7408-29. Robinson, J. T. Ego development of Human Sciences students. Unpublished manuscript, Fall 1974.

EP 7410-30. Lafferty, S. Observation reports of Human Sciences, Level II at Dunstan Junior High School, Lakewood, Colorado, October 4, 1974 to May 1, 1975. Unpublished manuscript, May 1, 1975.

EP 7501-31. Gray, W. M. Handout to accompany "The formal structure of concrete and formal operations: A confirmation of Piaget." Consultant report, January 24, 1975.

EP 7502-32. Smith, T., and Meier, A. Cluster analysis of "What's happening?" Consultant report, February 1975.

EP 7502-33. Smith, T., Meier, A., and Ideation Resource Group. Cluster analysis of "What's your opinion?" Unpublished manuscript, February 5, 1975.

EP 7505-34. Smith, T., and Ideation Resource Group. Completion of data analysis and data file preparation for Human Sciences. Unpublished manuscript, May 16, 1975.

EP 7508-35. Gagnon, J. Comments on REPRODUCTION, a Level II Human Sciences module. Consultant report, August 5, 1975.

EP 7110-36. Gerard, R. Comments on "A multidisciplinary Human Sciences Program for middle schools: A preliminary statement." Consultant report, October 4, 1971.

EP 7807-37. Robinson, J. T. Student attitudes towards science courses in test schools using Human Sciences. A working paper prepared for submission to the Journal for Research in Science Teaching, July 21, 1978.

EP 7904-38. Gray, W. M. Summary of "Standardized scoring criteria for measures of Piagetian logical operations." Published by William M. Gray, 1979.

EP 7904-39. Robinson, J. T. Selection of field test sites for Human Sciences. Unpublished manuscript, April 19, 1979.

EP 7906-40. Robinson, J. T. REPRODUCTION module evaluation: Student understanding of the concepts of sexual reproduction and fertilization. Unpublished manuscript, June 22, 1979.

EP 7906-41. Robinson, J. T. REPRODUCTION module evaluation: The context for evaluation. Unpublished manuscript, June 22, 1979.

EP 7907-42. Robinson, J. T. Evaluation of "Rabbit, Island," an activity in REPRODUCTION, a Level II Human Sciences module. Unpublished manuscript, July 11, 1979.

EP 7908-43. Robinson, J. T. Study of the sensitivity of achievement test items to instruction. Unpublished manuscript, September 29, 1978.

EP 7909-44. Robinson, J. T. Factor structure of "What's Happening?"--An attitude measure administered to Human Sciences test group students, May 1974. Unpublished manuscript, March 15, 1980.

EP 7910-45. McConnell, M. C. Report of the administration of "Goals of education," May 1974. Unpublished manuscript, September 11, 1974.

EP 8009-46. Robinson, J. T. A sample of findings from the evaluation of Human Sciences. Unpublished manuscript, September 11, 1980.

EP 8010-47. Robinson, J. T. Science curriculum and attitudes toward science. Unpublished manuscript, October 1, 1980.

EP 8101-48. Robinson, J. T. Effects of a bounded free-choice science curriculum environment. Paper presented at the Annual Meeting of the American Association for the Advancement of Science, Toronto, Canada, January 5, 1980.

EP 8103-49. Robinson, J. T. Achievement in Human Sciences test classes. Unpublished manuscript, March 3, 1981.

EP 8104-50. Ross, N. M., Robinson, J. T., Rasmussen, F. A., Eckenrod, J. E., McConnell, M. C., and Seitz, A. C. Evaluation of the Human Sciences Program. Unpublished manuscript, July 1980.

EP 7910-51. Robinson, J. T. A brief history of the Human Sciences project. Unpublished manuscript, October 17, 1979.

EP 7911-52. Robinson, J. T. The Human Sciences Program. Unpublished paper, November 6, 1969.

EP 7911-53. Robinson, J. T. Pilot modules: The model materials. Unpublished manuscript, November 8, 1979.

EP 7911-54. Robinson, J. T. Pilot modules/model materials field test. Unpublished manuscript, November 14, 1979.

EP 7912-55. Robinson, J. T. Results and recommendations from testing the pilot modules/model materials. Unpublished manuscript, December 27, 1979.

EP 8001-56. Robinson, J. T. Preparing for Level I field tests. Unpublished manuscript, January 1980.

EP 8002-57. Robinson, J. T. Evaluation of BEHAVIOR. Unpublished manuscript, February 1980.

EP 8003-58. Robinson, J. T. Evaluation of SURVIVAL. Unpublished manuscript, March 1980.

EP 8004-59. Robinson, J. T. Evaluation of SENSE...OR NONSENSE? Unpublished manuscript, April 1980.

EP 8005-60. Robinson, J. T. Evaluation of LEARNING. Unpublished manuscript, May 1980.

EP 8006-61. Robinson, J. T. Evaluation of GROWING. Unpublished manuscript, June 1980.

EP 8007-62. Robinson, J. T. Year ending assessment, Level I. Unpublished manuscript, July 1980.

EP 8008-63. Robinson, J. T. Adult evaluation of Human Sciences, Level I. Unpublished manuscript, August 1980.

EP 8111-64. Steele, J. Speaking of evaluation...alternatives to grading students. BSCS Newsletter, 1974, 54, 20-21.

EP 8111-65. Lorish, R. C. Evaluation of the Human Sciences Program. Consultant report, May 2, 1974.

EP 7408-66. Sullivan, P. Ego development of Human Sciences students. Consultant report, May 1974.

EP 8111-67. Robinson, J. T. Self-evaluation systems for Human Sciences. BSCS Newsletter, 1976, 63, 16-17.

EP 8111-68. Robinson, J. T. Comments on the readability study by Jacobson. Unpublished manuscript, August 5, 1977.

EP 8111-69. Gray, W. M. A comparison of Piagetian theory and criterion-referenced measurement. Review of Educational Research, 1978, 48(2), 223-249.

EP 8111-70. Ross, N. M. Learning and liking: Predicting student interests. BSCS Journal, 1978, 1(2), 22-23.

EP 8111-71. Robinson, J. T. Student attitudes toward science courses in test schools using Human Sciences. Journal of Research in Science Teaching, 1980, 17(3), 231-241.

EP 8112-72. Robinson, J. T. Evaluation of the BSCS Human Sciences Program. Unpublished manuscript, December 31, 1981.

EP 7804-73. Corley, K. A. BSCS Human Sciences follow-up. Unpublished manuscript, April 18, 1978.

EP 8112-74. Robinson, J. T. Human Sciences Evaluation Materials. Center for Educational Research and Evaluation, BSCS. Louisville, CO, 1981. (Available from the ERIC Clearinghouse for Science, Mathematics, and Environmental Education.)

REFERENCES

Clark, G. M. (ed.) Life Sciences in the Middle School. BSCS Special Publication Seven. Biological Sciences Curriculum Study Center, University of Colorado, Boulder, Colorado, 1969.

BSCS. Life Sciences for the Middle School. Proposal to the National Science Foundation. Biological Sciences Curriculum Study, Boulder, Colorado, October, 1970.

BSCS. Record and Transcript of Middle School Life Sciences Conference I, March 18-21, 1971. Harvest House Hotel, Boulder, Colorado. Human Sciences Project. Biological Sciences Curriculum Study, Boulder, Colorado, 1971a. Unpublished manuscript.

BSCS. Record and Transcript of Middle School Life Sciences Conference II, April 15-18, 1971. Royal Inn of Boulder, Boulder, Colorado. Human Sciences Project. Biological Sciences Curriculum Study, Boulder, Colorado, 1971b. Unpublished manuscript.

BSCS. Record and Transcript of Middle School Life Sciences Conference III, May 13-16, 1971. Harvest House Motor Hotel, Boulder, Colorado. Human Sciences Project. Biological Sciences Curriculum Study, Boulder, Colorado, 1971c. Unpublished manuscript.

BSCS. Teaching Human Sciences, Preliminary Edition. Boulder, Colorado: Biological Sciences Curriculum Study, 1979.

Campbell, D. T., and Stanley, J. C. Experimental and quasi-experimental designs for research on Teaching. In Gage, N. L. (ed.) Handbook of Research on Teaching. Chicago: Rand McNally, 1963.

Coleman, James S., et al. The Hopkins Game Program. Educational Researcher, 2:3-7, August, 1973.

Dewey, John. Experience and Education. New York: Macmillan, 1938.

DiVesta, F. J. A developmental study of semantic structures of children. In Snider, J. G. and Osgood, C. E. (eds.) Semantic Differential Techniques, A Sourcebook. Chicago, Illinois: Aldine, 1969.

Erikson, Erik H. Identity and the life cycle. Psychological Monographs, 1(1):1-171, 1959.

Flavell, J. Cognitive Development. Englewood, N.J.: Prentice-Hall, 1977.

Gray, W. M. Analysis of "How Is Your Logic?" Consultant report EP 7410-04 to the Human Sciences Project. Biological Sciences Curriculum Study, Boulder, Colorado, October 22, 1974. Unpublished manuscript.

Gray, W. M. and Robinson, J. T. The Developmental Construction and Validation of "How Is Your Logic?" (In preparation.)

Gray, W. M. How Is Your Logic? Biological Sciences Curriculum Study, Boulder, Colorado, May, 1976. See Human Sciences Archives, BP 8112-74, Robinson, J. T., Human Sciences Evaluation Materials. Center for Educational Research and Evaluation, BSCS, Lousiville, Colorado, 1981.

Gray, W. M. The Structure of Operational Thought: Some Empirical Support. Paper presented at the Eleventh Annual Symposium of the Jean Piaget Society, Philadelphia, Pennsylvania, May, 1981.

Inhelder, D. and Piaget, J. The Growth of Logical Thinking from Childhood to Adolescence: An Essay on the Construction of Formal Operational Structures. A. Parsons and S. Milgram (trans.). New York: Basic Books, 1958.

Kohlberg, L. Stage and Sequence: The Cognitive-Development Approach to Socialization. Handbook of Socialization Theory. Goslin, D. A. (ed.) Chicago: Rand McNally, 1969.

Loevinger, J. The Meaning and Measurement of Ego Development. American Psychologist. 1966, 21, 195-206.

Loevinger, J. and Wessler, R. Measuring Ego Development. Vol. 1 and Vol. 2. San Francisco: Jossey-Bass, 1970.

NSF. Panel Evaluation of 19 Pre-College Curriculum Development Projects. Directorate for Science Education, National Science Foundation, Washington, DC. NSF 76-23, 1976.

Osgood, C. E., Succi, G. J., and Tannenbaum, P. H. The Measurement of Meaning. Urbana, Ill: University of Illinois Press, 1957.

Robinson, J. T. and Cobern, W. W. Validity of "How Is Your Logic?" Human Sciences paper SP 7712-17. Biological Sciences Curriculum Study, Boulder, Colorado, April 17, 1979. Unpublished manuscript.

Robinson, James T. Student Attitudes Toward Science Courses in Test Schools Using Human Sciences. Journal of Research in Science Teaching, 17(3): 231-241, 1980.

Robinson, J. T. Human Sciences Evaluation Materials. Louisville, Colorado: Center for Educational Research and Evaluation, BSCS, December 31, 1981.

Robinson, J. T. The Logical Competencies Project: The Interaction of Curriculum Development and Research. In NSF. SEDR Project Directors Meeting, 1981. National Science Foundation (SEE 82-56). Washington, DC, 1982.

Schwab, Joseph J. The Practical 2: Arts of the Eclectic. School Review, 79(4):493-542, 1971.

Steele, J. M. The Relationship of Developmental Level with Science Performance: A Case for an Alternative to IQ Grouping of Mildly Retarded Children. A paper presented at the 1974 Annual Meeting of the American Educational Research Association, Chicago, Illinois, April, 1974.

Turkey, John W. Exploratory Data Analysis, Reading, Mass: Addison-Wesley, 1977.