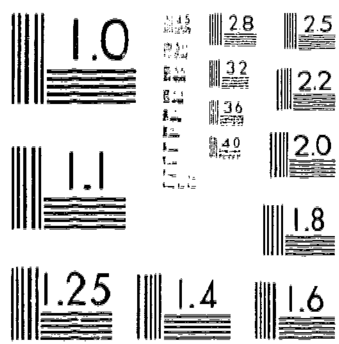


MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

DOCUMENT RESUME

ED 1?8 225 SE 021 330

AUTHOR Pouler, Chris Aemil
TITLE The Effect of Intensive Instruction in Hypothesis Generation Upon the Quantity and Quality of Hypotheses and the Quantity and Diversity of Information Search Questions Contributed by Ninth Grade Students.
PUB DATE 76
NOTE 157p.; Ph.D. Dissertation, University of Maryland; Contains light and broken type throughout

EDRS PRICE MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS Doctoral Theses; *Educational Research; General Science; Grade 9; Information Seeking; *Instruction; *Problem Solving; Research Methodology; Science Education; Secondary Education; *Secondary School Science
IDENTIFIERS Research Reports

ABSTRACT

Subjects selected from all the ninth grade intact science classes of a suburban junior high school were assigned to either a control group or one of four intensive instruction groups. The procedures for each instructional group included both watching a discrepant event until six acceptable hypotheses were written and individual discussion during which the investigator evaluated each of the six hypotheses according to standards that reflected the type of reinforcement and instruction the student received. The seven dependent variables determined by this study were: both the quantity and quality of written hypotheses following intensive instruction, both the quantity and diversity of written information search questions following intensive instruction, the diversity of information search questions during the group discussion, and both the quantity and quality of written hypotheses after the group discussion. The conclusions included: (1) participants who received intensive instruction which emphasized either differentiated reinforcement, criteria, or both, generated a higher quality of written hypotheses, following intensive instruction, than participants who received undifferentiated reinforcement or no intensive instruction; and (2) no form of hypothesis generation intensive instruction improves the participants' ability to generate a greater quantity or diversity of written information search questions following intensive instruction. (Author/MH)

THE EFFECT OF INTENSIVE INSTRUCTION IN HYPOTHESIS GENERATION UPON THE QUANTITY AND QUALITY OF HYPOTHESES AND THE QUANTITY AND DIVERSITY OF INFORMATION SEARCH QUESTIONS CONTRIBUTED BY NINTH GRADE STUDENTS

by
Chris Aemil Pouler

Dissertation submitted to the Faculty of the Graduate School
of the University of Maryland in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
1976

APPROVAL SHEET

Title of Thesis: The Effect of Intensive Instruction in Hypothesis Generation Upon the Quantity and Quality of Hypotheses and the Quantity and Diversity of Information Search Questions Contributed by Ninth Grade Students

Name of Candidate: Chris Aemil Pouler
Doctor of Philosophy, 1976

Thesis and Abstract Approved: Emmett L. Wright

Emmett L. Wright, Ph.D.
Assistant Professor
Secondary Education

Date Approved: July 20, 1976

CURRICULUM VITAE

Name: Chris Aemil Pouler.

Permanent address: 5409 19th Avenue
Hyattsville, Maryland 20782.

Degree and date to be conferred: Ph.D., 1976.

Date of birth: March 15, 1948.

Place of birth: Washington, D. C.

Secondary education: Northwestern Senior High School
Hyattsville, Maryland
June 1966.

Collegiate institutions attended:	Dates	Degree	Date of Degree
University of Maryland	1966-1970	B.S.	1970
The Catholic University of America	1972-1974	M.A.	1974
University of Maryland	1974-1976	Ph.D.	1976.

Major: Science Education.

Minor: Biology.

Professional publications:

Pouler, Chris A. "A DYNAMIC EARTH WEEK: Impetus for Action." Paper presented at the 22nd annual meeting of the National Science Teachers Association, Chicago, Ill., March 18, 1974.

Pouler, Chris A. "Are You Willing?" American Society for Microbiology News, 41 (October, 1975), 733-34.

Professional positions held:

Science Teacher:	Prince George's County Public Schools Upper Marlboro, Maryland October 1971-June 1974 Currently on Educational Leave to Pursue Doctorate
Instructor:	In-service course (Maryland State Department of Education Reference Number 1084) for Prince George's County Teachers September 1973-December 1973
Instructor:	In-service course (Maryland State Department of Education Reference Number 74-14) for Prince George's County Teachers January 1974-May 1974
Instructor:	"Curriculum Development K-12" for The Catholic University of America June 1974-July 1974 (Summer Session).

ABSTRACT

Title of Dissertation: The Effect of Intensive Instruction in Hypothesis Generation Upon the Quantity and Quality of Hypotheses and the Quantity and Diversity of Information Search Questions Contributed by Ninth Grade Students

Chris Aemil Pouler, Doctor of Philosophy, 1976

Dissertation directed by: Emmett L. Wright, Ph.D.
Assistant Professor
Secondary Education

The purpose of this investigation was to:

a. compare the effects of various hypothesis generation intensive instruction procedures on the ability of ninth grade students to generate written hypotheses and information search questions about a discrepant event, and

b. determine if students, who have received hypothesis generation intensive instruction in a setting free of peer interactions, exhibit a greater diversity of questions during a group discussion and greater written hypothesis generation behaviors after the discussion.

The subjects--selected from all the ninth grade intact science classes of a suburban junior high school--were assigned to either a control group or one of four intensive instruction groups. The procedures for each instructional group included:

a. watching the intensive instruction discrepant event until six acceptable hypotheses were written, and

b. individual discussions during which the investigator evaluated each of the six hypotheses by one of the following predetermined standards:

1. differentiated reinforcement and criteria group--depending on the level of acceptability for each hypothesis, the student was both positively reinforced (good, very good, excellent) and told the criteria for good hypothesis formation,

2. undifferentiated reinforcement and criteria group--after each acceptable hypothesis, the student was only told the criteria for good hypothesis formation,

3. differentiated reinforcement only group--depending on the level of acceptability for each hypothesis, the student was positively reinforced (good, very good, excellent), and

4. undifferentiated reinforcement only group--had to generate six hypotheses (of any quality) each of which was accepted without positive reinforcement.

During the intensive instruction sessions, the subjects were shown discrepant events selected from the set of Inquiry Development Program Films. Upon completion of the instructional sessions, all the experimental groups were shown another discrepant event and were requested to write as many hypotheses as possible. Then an additional film was shown and the participants were requested to write as many questions as possible which solicited information to help explain the discrepancy. Five days later each experimental group participated in a discussion where they observed another discrepant event film and, then had the opportunity to voluntarily ask questions of the investigator about the discrepant event. All questions were only answered with yes or no. After twenty questions, the discussion was terminated and the students were requested to individually write hypotheses that might explain the discrepancy.

The seven dependent variables determined by this study were:

a. both the quantity and quality of written hypotheses following intensive instruction,

b. both the quantity and diversity of written information search questions following intensive instruction,

c. the diversity of information search questions during the group discussion, and

d. both the quantity and quality of written hypotheses after the group discussion.

The conclusions included the following:

a. Differentiated reinforcement--as an intensive instruction method--is responsible for a higher quantity of written hypotheses after intensive instruction than the instruction method which involves no intensive instruction.

b. Participants who received intensive instruction which emphasized either differentiated reinforcement, criteria or both generated a higher quality of written hypotheses, following intensive instruction, than participants who received undifferentiated reinforcement or no intensive instruction.

c. No form of hypothesis generation intensive instruction improves the participants' ability to generate a greater quantity or diversity of written information search questions following intensive instruction.

d. Differentiated reinforcement only--as an intensive instruction condition--is responsible for a greater quantity of written hypotheses than no intensive instruction following the group discussion.

e. Criteria as an intensive instruction method is responsible for a higher quality of written hypotheses, after the group discussion, than the instruction method of undifferentiated reinforcement or no intensive instruction.

f. And, in the presence of the group discussion, diversity of oral information search questions is not significantly improved by hypothesis generation intensive instruction.

ACKNOWLEDGMENTS

This author wishes to acknowledge the following individuals who have directly contributed to the completion of this research study.

1. Dr. Emmett L. Wright for spending many hours as a consistently excellent advisor.

2. Drs. John C. Carr, Harris H. Linder, J. David Lockard and William D. Schafer for serving not only as members of the final examining committee but for taking the time to offer useful suggestions and worthwhile advice.

3. Dr. Ann C. Gross, Director of Research and Evaluation for the Prince George's County Public Schools, and her enthusiastic assistant, Dr. Elwood Loh, for their friendly and sincere desire to help implement this study in the Prince George's County Public Schools.

4. Mr. George W. Anderson, Jr., Principal, and Mr. Karl Taschenberger, Vice Principal, of Frederick Sasscer Junior High School for allowing this study to utilize their excellent school.

5. Miss Nancy Douglas and Mrs. Jo Norris not only for allowing me to utilize their classes but also for their genuine interest in promoting science education.

6. And, finally the students who participated in this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

AN INTRODUCTION

> I am myself a great lover of these processes on division and generalization; they help me to speak and to think.
>
> Socrates, 470-399, B.C.
> From Dialogue to Phaedrus

Both knowledge and application of problem solving skills became important objectives of science education for the development and, subsequent, introduction of inquiry curricula during the late fifties and early sixties. The impact of these curricula was so great that massive science curricular reform resulted by modifying instruction in such a way that students were encouraged to acquire an understanding of the attitudes and processes of scientific inquiry (Gagné, 1963, p. 152). Consequently, the science classroom has become the center for activities which encourage students to learn the principles of the various academic disciplines while simultaneously acquiring and applying problem solving skills. As the popularity and widespread use of the inquiry approach increased, so did the number and diversity of curricula available for both elementary and secondary instruction. In fact, the trend has been especially apparent at the junior high level which has experienced an increasing emphasis on discovery type (inquiry) activities (Trowbirdge, 1974, p. 13). The future direction for inquiry curricula is reflected by recent curricula developments which include:

a. the adaptation of the original physics inquiry program (PSSC) into an individualized format (Friedman, 1976, p. 15), and

b. the development of new inquiry curricula emphasizing individualization which permits content selection to be influenced by student interest, local relevance, and existing facilities (Burkman, 1974, p. 30).

Since activities which require problem solving skills comprise an important aspect of contemporary and future science instruction, it appears worthwhile to investigate methods which may improve the acquisition and utilization of the basic inquiry skills.

Inquiry instruction encourages students to discover facts of causality through their own initiative and not to be dependent on explanations from teachers or other knowledgeable persons (Pugliese, 1973, p. 26). Essential among the basic inquiry skills is the ability to isolate variables leading to the generation of relevant hypotheses. In an attempt to improve the ability to generate hypotheses through questioning, Suchman (1961, p. 159) devised a schema which identified the stages that students must follow in order to adequately explain an inconsistent scientific event. These stages include:

a. episode analysis--where the facts of the event are verified,

b. determination or relevance--where the necessary conditions of relevance of variables are isolated, and

c. induction of relational constructs--which allows the formation of explanations about the event.

From the schema, Suchman (1960, 1961, 1962b) developed an entire inquiry training program which was based upon inconsistent (or discrepant) events of physics causality. During each inquiry session, students observed a film and then volunteered individual questions, as part of a group discussion effort, to isolate the variables of the event and explain the causality. A scale, based on the schema, was then utilized by Suchman to classify the questions into categories. Suchman (1962b) discovered upon completion of the training program, the participants attained a greater acquisition of questioning skills than students not participating. Suchman's work illustrated that students can through classroom training improve their abil.ty to gather data by asking questions about the variables inherent within a problem.

More recent research, which considered various extensions of Suchman's work, has indicated that students can be instructed to improve their ability to generate hypotheses. Quinn (1971, 1972), working with students from intact sixth grade classrooms, found a greater quality of written hypotheses from those subjects who received classroom instruction describing the components of a good hypothesis. The instruction consisted of a series of class discussions emphasizing analysis of observations (made from a Suchman Inquiry Development Program Film) in terms of relevant variables and degrees of acceptable hypotheses as determined by a quality scale (see Table 5, p. 72).

Working with students individually in a setting removed from the classroom, Wright (1974) found that ninth grade students improved their verbal hypothesizing ability when exposed to either of two forms of instruction. The instruction consisted of observing a Suchman film and then, depending on the treatment group, the student was required to describe seventy-five details or generate five hypotheses which corresponded to a predetermined level of acceptability. When a student submitted an acceptable hypothesis, he or she was reinforced by an approving term such as good. It is important to note that the reinforcement was differentiated in the sense that each approval comment was intended to promote only one kind of response--the formation of an acceptable hypothesis. No attempt was made to offer (a) undifferentiated reinforcement which would merely encourage the formation of hypotheses and (b) differentiated reinforcement (e.g. good, very good, excellent) which would encourage the formation of a predetermined level of acceptable hypotheses. In this sense, undifferentiated and differentiated reinforcement could be utilized to develop various types of student responses.

In summary, Wright's work reinforced each student response and differed from Quinn's work in four ways:

1. the instruction was individualized so each student worked with the investigator and controlled the amount of time that was necessary and the number of times the film was observed,

2. the criteria were not given for an acceptable hypothesis,

3. the effect of two forms of instruction (observing details and hypothesis generation) upon hypothesis generation skills were compared, and

4. the intensive instruction model, which was employed, included a predetermined minimum of hypotheses or observations that had to be attained by each student.

Suchman's original research, together with the work of Quinn and Wright, raises three questions which are worthy of further research:

1. What is the difference between intensive instruction in hypothesis formation when only the criteria are given to the student and intensive instruction in hypothesis formation when only differentiated reinforcement or both criteria and differentiated reinforcement are given?

2. What effect will previous hypothesis generation intensive instruction have on a group discussion about a specific discrepant event?

3. And, what effect will the group discussion have on the members of the instructed groups to subsequently generate hypotheses?

The first question is an attempt to define suitable means to improve hypothesis generating abilities in classroom settings. Further, the answer can offer a viable addition to already existing science instruction. The second and third questions are important since it is advantageous to couple specific instruction with group interactions because:

a. discussions are important as a learning approach in secondary schools, and

b. most problems of communities and society are solved in a social setting and it appears worthwhile to offer students the opportunity to practice group problem solving behavior.

Also, group discussion may enhance previous hypothesis generation intensive instruction and, therefore, serve as an additional teaching strategy. Research that attempts to answer these questions would offer useful strategies for the teaching of science.

The purpose of this study was to compare various forms of intensive instruction in hypothesis generation and, then, determine if students who have received instruction to improve their ability to formulate hypotheses, in a setting free of peer interactions, were able to exhibit an improved ability to ask a greater diversity of questions about a discrepant event during a group discussion and, then, write more hypotheses of a higher quality to explain the event.

Need for the Study

The original innovators of the science inquiry curricula based their work upon the premise that an understanding of the fundamental principles was sufficient to allow extrapolation of the necessary specifics (Bruner, 1971, p. 18). Consequently, the instructional strategy of the resulting curricula emphasized an independent, systematic, empirical, and inductive approach to science (Pugliese, 1973, p. 26). Therefore, it appears logical to assume that students developed the cognitive skills of problem solving at the expense of developing the ability to express and share the skills they acquired. Along these lines, Babikian (1971), in a comparison of expository, laboratory and discovery methods of instruction, found that students taught by inquiry were significantly less effective in the verbalization of scientific concepts. Thus, an important concern for curriculum designers of the future will be to incorporate group discussion activities which utilize acquired inquiry skills. This research study was an attempt to provide evidence from which to design relevant activities in future curricula.

Although Suchman (1961), Quinn (1971), and Wright (1974) illustrated that discussion is a feasible instruction method to increase problem solving skills, each researcher utilized a different approach. While Suchman and Quinn employed group techniques, Wright worked with each student individually. Since a major direction of science curriculum is presently individualization, there is need for research data which result from:

a. a comparison of different methods of intensive instruction for improving process skills which allow students to work individually, and

b. an assessment of the effect of the intensive instruction on group discussion behavior.

In this way, the obtained findings could serve as a means to help students improve their ability to interact while utilizing previously learned skills. Along these lines, Glasser (1969, p. 36), as a result of success with discussion groups in elementary schools, has recommended the teaching of critical thinking through group discussions from elementary to high school. Specifically dealing with science instruction, Jerrigan (1972) illustrated that specific questions from readings could be employed to create worthwhile discussion when the class members attempted, as a group, to obtain an answer. Therefore, a practical need exists for research that provides evidence that problem solving skills can be taught efficiently through individualization and, later utilized by students during a group discussion.

Another important area of support for this study involves the level of mental development of junior high students. Since the inquiry approach emphasized the acquisition of problem solving skills, it seems only logical that students must possess an appropriate mental development to adequately acquire and apply problem solving skills. Previous works with elementary and junior high students have dealt with discrepant events which were clearly depicted on film, by deliberate actions of an experimenter and close up photography, so students could clearly recognize the problem and several variables (Suchman, 1961, pp. 150-51). By relating the variables, the students were required to perform mental operations which led to the development and utilization of hypothetico-deductive reasoning as defined by Jean Piaget's level of formal operations (for further discussion see pp. 26-30). Thus, it is important to emphasize that findings from hypothesis generation studies could prove to provide the means to aid in the mental development of the student by either:

a. sharpening inquiry skills which may already be present, or

b. enhancing the development of basic inquiry skills not yet present.

This conclusion is crucial since two recent studies--(a) Sayre and Ball (1975) and (b) Lawson and Renner (1975)--have found that while some secondary students have achieved the level of formal mental operations, many students do not think on this level. Therefore, investigation into techniques that a classroom teacher could use to promote the development of formal reasoning is certainly worthwhile. In fact, the researchers mentioned above--(a) Sayre and Ball and (b) Lawson and Renner--strongly recommended that elementary and junior high curricula be designed to present perceivable problem solving experiences which will help students to develop the processes of formal operations. When the preceeding recommendation is coupled with the findings from previous investigations by Suchman, Quinn and Wright, the contention that basic inquiry skills can be taught and that research is necessary to determine the optimal method of instruction becomes apparent.

In summary, Suchman (1962b) developed an effective method to help participants better assimilate data and generate hypotheses about a discrepant event. Further, Quinn (1971) and Wright (1974) devoted their efforts to the analysis of hypothesis training. Their findings have indicated:

a. instruction in hypothesis generation yields a higher number of hypotheses about a discrepant event,

b. instruction in hypothesis generation is effective in both intact classroom and individualized situations, and

c. there exist more than one effective method for hypothesis generation instruction.

Unfortunately, none of the previous studies have included: (a) comparisons of different forms of individualized intensive instruction, (b) assessments of the value of prior individualized intensive instruction on group discussion behavior, or (c) assessments of both previous individualized intensive instruction and group discussion on hypothesis generation. In essence, this study attempted to bridge the gap between research findings and classroom feasibility by further defining the methods to promote problem solving capabilities.

The Research Problem

One of the key elements in the process of scientific inquiry is the generation of hypotheses that offer possible explanations for the occurrence of a discrepant event. Basically, this mental operation involves: (a) exposure to a problem, (b) isolation of the relevant variables, and (c) generation of hypotheses.

This study was designed to:

a. compare the effect of different forms of hypothesis generation intensive instruction on the ability of students to generate hypotheses and ask questions about a discrepant event, and

b. determine if students, who have received hypothesis generation intensive instruction in a setting free of peer interactions, are able to exhibit a greater diversity of questions during a group discussion and, then, a greater ability to generate hypotheses about a discrepant event.

Hypotheses

Quantity of Written Hypothesis Generation Following Intensive Instruction

1. There is a difference in the quantity of written hypotheses between the following groups which receive a form of instruction and the control group which receives no instruction:

   a. the differentiated reinforcement and criteria group versus the control group,

   b. the undifferentiated reinforcement and criteria group versus the control group,

   c. the differentiated reinforcement only group versus the control group, and

   d. the undifferentiated reinforcement only group versus the control group.

2. The effects of differentiated reinforcement and criteria are non-additive on the quantity of written hypotheses.

3. There is a difference in the quantity of written hypotheses between the following groups which receive differentiated reinforcement as an instruction condition and the following groups which receive undifferentiated reinforcement:

    a. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement and criteria group,

    b. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement only group,

    c. the differentiated reinforcement only group versus the undifferentiated reinforcement and criteria group, and

    d. the differentiated reinforcement only group versus the undifferentiated reinforcement only group.

4. There is a difference in the quantity of written hypotheses between the following groups which are told the criteria for good hypothesis formation and the following groups which are not told the criteria for good hypothesis formation:

    a. the differentiated reinforcement and criteria group versus the differentiated reinforcement only group,

    b. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement only group,

    c. the undifferentiated reinforcement and criteria group versus the differentiated reinforcement only group, and

    d. the undifferentiated reinforcement and criteria group versus the undifferentiated reinforcement only group.

Quality of Written Hypothesis Generation
Following Intensive Instruction

5. There is a difference in the quality of written hypotheses between the following groups which receive a form of instruction and the control group which receives no instruction:

    a. the differentiated reinforcement and criteria group versus the control group,

    b. the undifferentiated reinforcement and criteria group versus the control group,

c. the differentiated reinforcement only group versus the control group, and

d. the undifferentiated reinforcement only group versus the control group.

6. The effects of differentiated reinforcement and criteria are non-additive on the quality of written hypotheses.

7. There is a difference in the quality of written hypotheses between the following groups which receive differentiated reinforcement as an instruction condition and the following groups which receive undifferentiated reinforcement:

   a. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement and criteria group,

   b. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement only group,

   c. the differentiated reinforcement only group versus the undifferentiated reinforcement and criteria group, and

   d. the differentiated reinforcement only group versus the undifferentiated reinforcement only group.

8. There is a difference in the quality of written hypotheses between the following groups which are told the criteria for good hypothesis formation and the following groups which are not told the criteria for good hypothesis formation:

   a. the differentiated reinforcement and criteria group versus the differentiated reinforcement only group,

   b. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement only group,

   c. the undifferentiated reinforcement and criteria group versus the differentiated reinforcement only group, and

   d. the undifferentiated reinforcement and criteria group versus the undifferentiated reinforcement only group.

Quantity of Written Information Search Questions
Following Intensive Instruction

9. There is a difference in the quantity of written information search questions between the following groups which receive a form of instruction and the control group which receives no instruction:

   a. the differentiated reinforcement and criteria group versus the control group,

   b. the undifferentiated reinforcement and criteria group versus the control group,

   c. the differentiated reinforcement only group versus the control group, and

   d. the undifferentiated reinforcement only group versus the control group.

10. The effects of differentiated reinforcement and criteria are non-additive on the quantity of written information search questions.

11. There is a difference in the quantity of written information search questions between the following groups which receive differentiated reinforcement as an instruction condition and the following groups which received undifferentiated reinforcement:

   a. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement and criteria group,

   b. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement only group,

   c. the differentiated reinforcement only group versus the undifferentiated reinforcement and criteria group, and

   d. the differentiated reinforcement only group versus the undifferentiated reinforcement only group.

12. There is a difference in the quantity of written information search questions between the following groups which are told the criteria for good hypothesis formation and the following groups which are not told the criteria for good hypothesis formation:

   a. the differentiated reinforcement and criteria group versus the differentiated reinforcement only group,

   b. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement only group,

c. the undifferentiated reinforcement and criteria group versus the differentiated reinforcement only group, and

d. the undifferentiated reinforcement and criteria group versus the undifferentiated reinforcement only group.

Diversity of Written Information Search Questions Following Intensive Instruction

13. There is a difference in the diversity of written information search questions between the following groups which receive a form of instruction and the control group which receives no instruction:

    a. the differentiated reinforcement and criteria group versus the control group,

    b. the undifferentiated reinforcement and criteria group versus the control group,

    c. the differentiated reinforcement only group versus the control group, and

    d. the undifferentiated reinforcement only group versus the control group.

14. The effects of differentiated reinforcement and criteria are non-additive on the diversity of written information search questions.

15. There is a difference in the diversity of written information search questions between the following groups which receive differentiated reinforcement as an instruction condition and the following groups which received undifferentiated reinforcement:

    a. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement and criteria group,

    b. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement only group,

    c. the differentiated reinforcement only group versus the undifferentiated reinforcement and criteria group, and

    d. the differentiated reinforcement only group versus the undifferentiated reinforcement only group.

16. There is a difference in the diversity of written information search questions between the following groups which are told the criteria for good hypothesis formation and the following groups which are not told the criteria for good hypothesis formation:

a. the differentiated reinforcement and criteria group versus the differentiated reinforcement only group,

b. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement only group,

c. the undifferentiated reinforcement and criteria group versus the differentiated reinforcement only group, and

d. the undifferentiated reinforcement and criteria group versus the undifferentiated reinforcement only group.

Quantity of Written Hypothesis
Generation After the Group Discussion

17. There is a difference in the quantity of written hypotheses between the following groups which receive a form of instruction and the control group which receives no instruction:

    a. the differentiated reinforcement and criteria group versus the control group,

    b. the undifferentiated reinforcement and criteria group versus the control group,

    c. the differentiated reinforcement only group versus the control group, and

    d. the undifferentiated reinforcement only group versus the control group.

18. The effects of differentiated reinforcement and criteria are non-additive on the quantity of written hypotheses.

19. There is a difference in the quantity of written hypotheses between the following groups which receive differentiated reinforcement as an instruction condition and the following groups which receive undifferentiated reinforcement:

    a. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement and criteria group,

    b. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement only group,

    c. the differenitated reinforcement only group versus the undifferentiated reinforcement and criteria group, and

d. the differentiated reinforcement only group versus the undifferentiated reinforcement only group.

20. There is a difference in the quantity of written hypotheses between the following groups which are told the criteria for good hypothesis formation and the following groups which are not told the criteria for good hypothesis formation:

   a. the differentiated reinforcement and criteria group versus the differentiated reinforcement only group,

   b. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement only group,

   c. the undifferentiated reinforcement and criteria group versus the differentiated reinforcement only group, and

   d. the undifferentiated reinforcement and criteria group versus the undifferentiated reinforcement only group.

Quality of Written Hypothesis
Generation After the Group Discussion

21. There is a difference in the quality of written hypotheses between the following groups which receive a form of instruction and the control group which receives no instruction:

   a. the differentiated reinforcement and criteria group versus the control group,

   b. the undifferentiated reinforcement and criteria group versus the control group,

   c. the differentiated reinforcement only group versus the control group, and

   d. the undifferentiated reinforcement only group versus the control group.

22. The effects of differentiated reinforcement and criteria are non-additive on the quality of written hypotheses.

23. There is a difference in the quality of written hypotheses between the following groups which receive differentiated reinforcement as an instruction condition and the following groups which receive undifferentiated reinforcement:

   a. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement and criteria group,

b. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement only group,

c. the differentiated reinforcement only group versus the undifferentiated reinforcement and criteria group, and

d. the differentiated reinforcement only group versus the undifferentiated reinforcement only group.

24. There is a difference in the quality of written hypotheses between the following groups which are told the criteria for good hypothesis formation and the following groups which are not told the criteria for good hypothesis formation:

a. the differentiated reinforcement and criteria group versus the differentiated reinforcement only group,

b. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement only group,

c. the undifferentiated reinforcement and criteria group versus the differentiated reinforcement only group, and

d. the undifferentiated reinforcement and criteria group versus the undifferentiated reinforcement only group.

Diversity of Information Search
Questions During the Group Discussion

25. There is a difference in the diversity of questions between the experimental groups during an information search group discussion about a discrepant event.

Performance in Hypothesis Generation After
the Group Discussion Using Previous Hypothesis
Generation Experience as a Covariable

26. Using the results of the quantity of written hypotheses immediately following instruction as a covariable for the quantity of written hypotheses after the group discussion, there is a difference between the following planned comparisons:

a. instruction groups versus the control group,

b. interaction between instruction groups,

c. differentiated reinforcement groups versus undifferentiated reinforcement groups, and

d. criteria groups versus non-criteria groups.

27. Using the results of the quality of written hypotheses immediately following instruction as a covariable for the quality of written hypotheses after the group discussion, there is a difference between the following planned comparisons:

   a. instruction groups versus the control group,

   b. interaction between instruction groups,

   c. differentiated reinforcement groups versus undifferentiated reinforcement groups, and

   d. criteria groups versus non-criteria groups.

Definitions of Terms

Major Definitions

1. Acceptable Hypothesis Generation: When a hypothesis that was written and discussed by a participant rated at least a three on the Hypothesis Quality Scale (see p. 72). Hypotheses in this category were scientific explanations relating at least two variables in general or nonspecific terms.

2. Differentiated Reinforcement: A form of reinforcement which offers verbal encouragement for each hypothesis contributed by a participant during intensive instruction which meets a predetermined level of quality. The specific terms of reinforcement--good, very good and excellent--correspond to hypotheses which are described by levels three, four, and five of the Hypothesis Quality Scale (see p. 72).

3. Discrepant Event: An occurrence that illustrated a scientific principle of causality that could have been logically explained by more than one acceptable hypothesis. In addition, the discrepant events contained several variables and presented (within reason) a new experience to the subjects. This study utilized several Inquiry Development Program Films (Suchman, 1962b) which met these conditions.

4. Discrepant Event for Hypothesis Generation Following Instruction: The discrepant event, "The Sailboat and the Fan," was the source for the hypotheses which were analyzed for the comparisons between experimental groups to assess the effects of the various methods of instruction. The instructions to the participants for this activity were audiotaped and the script appears on pages 68-69.

5. Discrepant Event for Information Search Questions Following Instruction: The discrepant event, "The Ice Cubes," was the source for the questions which were analyzed in the comparisons between the various experimental groups to assess the impact of intensive instruction. The instructions to the participants were audiotaped and the script appears on pages 68-69.

6. Diversity of Information Search Questions: After all the questions for each participant had been placed into the appropriate category (see diversity of information search questions classification, below), a diversity value utilizing the Shannon Index (see p. 74) was determined. All the individual diversity values for all the participants in each experimental group, were averaged and, then, used in the statistical analyses.

7. Diversity of Information Search Question Classification: The extent to which generated questions for each individual corresponded to one of sixteen predetermined categories. To determine the appropriate category for a question, it was first placed into one of the following major groupings on the basis of intent:

   a. verification: if the question sought to identify one aspect about the total event or sequence about the discrepant event,

   b. experimentation: if the question sought to ascertain the consequences of some hypothetical change (would the same have happened if) in the experiment presented by the discrepant event,

   c. necessity: if the question sought to determine whether a specific aspect of the phenomenon was necessary for the outcome (cause and effect), and

   d. synthesis: if the question sought to determine if a particular idea of theory about causation was valid and explained the experiment totally or in part.

   Then each question was classified by one of the following sub-groupings:

   a. events: referred to the occurrence of events,

   b. objects: referred to the nature of objects,

   c. conditions: referred to the states of objects, and

   d. properties: referred to the properties.

   By following the above description, each question fell into one of sixteen categories; for a further description see pages 72-74.

8. Diversity of Information Search Scale: A means to classify questions contributed by participants during an information search discussion. The scale is presented on page 73.

9. Experimental Group E (Control Group): The control group where subjects were only exposed to the introductory discrepant event.

10. Group Discussion: This activity was intended to assess the impact of the various methods of hypothesis generation intensive instruction on the participants of each experimental group to ask a diversity of questions about the group discussion discrepant event. Basically, during the discussion, the students watch the discrepant event and, then volunteered questions which had to be specific enough for a yes or no response. As a question was asked, it was repeated by the investigator, written and projected on an overhead transparency, and, finally, answered. Since the largest group was thirteen students, each student had the opportunity to contribute one question. The group discussion was terminated after twenty questions.

11. Group Discussion Discrepant Event: The discrepant event, "Drinking Boiling Coffee," was used as the problem for the group discussion and subsequent hypothesis generation session. The instructions were audiotaped and the script appears on pages 70-71.

12. Hypothesis: A logical, reasonable, and testable explanation for the occurrence of a discrepant event.

13. Hypothesis Generation: A behavior that involves writing multiple hypotheses which attempt to explain a discrepant event.

14. Hypothesis Generation Intensive Instruction: An instruction session, of one class period (fifty minutes), given to a small group of students (six to twelve) where each individual worked alone. Specifically, the students performed the following: (a) listened to an explanation of hypotheses and observed an introductory discrepant event (see pp. 64-66), (b) wrote between one and six hypotheses about an intensive instruction discrepant event, (c) discussed each hypothesis with the investigator, and (d) completed the session when the investigator accepted six hypotheses. Variations in instruction occurred when the participants of each instruction group (see p. 20) talked with the investigator. To insure individual participation, all instructions were audiotaped and transmitted to each student via headphones. Also, the discussion with the experimenter took place in a discussion area (isolated) so ideas were not overheard.

15. Information Search Following Intensive Instruction: An activity where each individual of each experimental group wrote as many questions as possible in an effort to isolate and define the relevant variables presented by the discrepant event for information search questions following intensive instruction.

16. Intensive Instruction Group: Refers to one of the four groups of participants which received a form of hypothesis generation intensive instruction.

17. Intensive Instruction: Term used to refer to hypothesis generation intensive instruction. Specifically, students had to write and discuss six hypotheses with the investigator and, depending on the intensive instruction group, the hypotheses had to meet predetermined standards.

18. Intensive Instruction Discrepant Event: This discrepant event, "The Knife," provided the problem to the various intensive instruction groups during intensive instruction sessions. Instructions to the groups were audiotaped and are included on pages 64-66.

19. Intensive Instruction Group A (Differentiated Reinforcement and Criteria): Intensive instruction group where each subject wrote six acceptable hypotheses which attempted to explain the intensive instruction discrepant event. At any time a participant could request to come to the discussion area and present one or more hypotheses to the investigator. During the discussion, each acceptable hypothesis (which met a predetermined standard) was accepted and the subject was: (a) positively reinforced (e.g. good) and (b) told the criteria for acceptable hypothesis generation.

20. Intensive Instruction Group B (Undifferentiated Reinforcement and Criteria): Intensive instruction group where each subject wrote six acceptable hypotheses which attempted to explain the intensive instruction discrepant event. At any time a participant could request to come to the discussion area and present one or more hypotheses to the investigator. During the discussion, each acceptable hypothesis (which met a predetermined standard) was accepted and the subject was told the criteria for acceptable hypothesis generation.

21. Intensive Instruction Group C (Differentiated Reinforcement Only): Intensive instruction group where each subject wrote six acceptable hypotheses which attempted to explain the intensive instruction discrepant event. At any time a participant could request to come to the discussion area and present one or more hypotheses to the investigator. During the discussion, each acceptable hypothesis (which met a predetermined standard) was accepted and the subject was positively reinforced (e.g. good).

22. Intensive Instruction Group D (Undifferentiated Reinforcement Only): Intensive instruction group where each subject wrote six hypotheses which attempted to explain the intensive instruction discrepant event. At any time a participant could request to come to the discussion area and present one or more hypotheses to the investigator. During the discussion all hypotheses were accepted and the subject was: (a) not positively reinforced and (b) not told the criteria for acceptable hypothesis generation.

23. Introductory Discrepant Event: This discrepant event, "The Restaurant," served as the means to illustrate the concept of hypotheses as explanations that attempt to solve a problem of causality. Students from all the experimental groups were exposed to the film and instructional audio ape. The script for the audiotape is included on pages 64-66.

24. Quality of Hypothesis Generation: The average value of the numerical quality of hypotheses for each individual in each group. To obtain the value, each hypothesis of an individual was placed into the appropriate category of the previously validated Quinn Hypothesis Quality Scale (see p. 72). The values were then averaged for each individual and, then, for each experimental group.

25. Quantity of Hypothesis Generation: The number of non-repeated hypotheses generated by subjects as determined by count.

26. Undifferentiated Reinforcement: A form of reinforcement which only accepts a hypothesis contributed by a participant during intensive instruction without any form of verbal encouragement. (For example, part were told "I can accept this hypothesis.").

Minor Definitions

1. Basic Inquiry Skills: The ability to isolate the variables that account for a problem of causality and, then, to generate an explanation for the problem which accounts for the variables.

2. Experimental Group: Refers to one of the four hypothesis generation intensive instruction groups or the control group.

3. Group Discussion Information Search: Term used synonymously with group discussion (see p. 19).

4. Individualized Instruction: Type of hypothesis generation intensive instruction where each student interacts individually with the investigator in a setting removed from peer interactions.

5. Inquiry Approach to Science Instruction: Instructional approach which utlized inquiry curricula and problem solving activities.

6. Inquiry Curricula: Type of science curricula which emphasize activities that allow students to learn the principles of a specific discipline while developing and utilizing basic inquiry skills. Generally, these curricula are associated with activity oriented classrooms rather than lecture oriented classrooms.

7. Instruction Group: Refers to one of the four groups of participants which received a form of hypothesis generation intensive instruction.

8. Problem Solving Skills: The ability to solve a problem with basic inquiry skills.

9. Treatment: Term which refers to one of the four hypothesis generation intensive instruction procedures (see pp. 19-20).

Assumptions

1. It was assumed that all the subjects could generate hypotheses and ask questions about the variables inherent within a problem of causality. The assumption appears reasonable since the average age of the participants during the study was 14 years and 8 months.

2. Since all subjects were selected from the same school, it was assumed that prior sensitivity to hypothesis generation and information search questioning was either nonexistent or if present, evenly distributed among the subjects.

3. It was also assumed that each discrepant event presented an original problem of causality to all the subjects. While this may not have been the case, the random distribution of students in the treatment groups minimized this threat.

Limitations of the Study

1. Although the study involved most of the ninth grade students in a single junior high school, participation was voluntary and dependent upon parental approval. As a result, a portion of the students did not participate. To compensate, however, there was random placement of all the participants into experimental groups.

2. The study was only conducted in a suburban school that was part of a large school system.

3. There may have been some interaction of personality factors between the investigator and some participants that either encouraged or discouraged performance.

4. Students were not screened for visual or hearing problems.

5. Only one visual medium (film) was used to expose subjects to the discrepant events.

6. Students were only permitted to write hypotheses and questions for six of the dependent variables.

Summary

The development of thinking and speaking skills is an essential objective for future science curricula. While previous research has indicated how the development of thinking skills can be improved, little effort has been devoted to comparing the effect of various hypothesis generation intensive instruction methods on the ability to isolate and relate variables in attempts to explain problems of causality. Further, previous studies in hypothesis formation have not assessed the impact of prior intensive instruction on group discussion questioning behavior or hypothesis generation behavior following the group discussion. The purpose of this study was to first compare various procedures for hypothesis generation intensive instruction and, then, to assess the impact of each instructional form on the information search behavior during a group discussion and the hypothesis generating behavior after the discussion about a discrepant event.

Chapter II provides support for this study based on previous findings of studies which attempted to improve hypothesis formation abilities. The review of literature is, then, followed by Chapters III, IV, and V which discuss in detail the (a) procedures, (b) findings, and (c) conclusions, implications and recommendations.

CHAPTER II

A REVIEW OF THE LITERATURE

But such is the irresistible nature of truth,
that all it asks, and all it wants, is the liberty of appearing.

Thomas Paine, 1737-1809

The scientific method, as an "active" state of mental inquiry, begins with casual observations and continues with qeustions, hypotheses, and experiments which lead to the establishment of theories, principles, and laws about causality. Perhaps, the most crucial step in the entire process is the formation of hypotheses since they provide the testable base from which to verify ideas. Because an understanding of the scientific method--as reflected by the scientific advances that affect society--is an essential part of education, contemporary science curricula have been developed that emphasize both the concepts and processes of science. Thus, students discover the principles of biology, earth science, chemistry, and physics by utilizing specific process skills embedded in the scientific method. This study attempted to determine additional effective teaching strategies that will promote the acquisition and development of science process skills. Specifically, the effects of various forms of hypothesis generation intensive instruction on hypothesis generating and information search questioning behaviors of ninth grade students were investigated.

The review of the literature reports and analyzes the important hypothesis generation investigations over the last twenty years. The major sources consulted for this presentation include:

Current Index to Journals in Education
CCM
Information Sciences, Inc.
New York, New York

Dissertation Abstracts
University Microfilms, Inc.
Ann Arbor, Michigan

Education Index
The H. W. Wilson Company
New York, New York

ERIC Research in Education
U. S. Office of Education
Washington, D. C.

Encyclopedia of Educational Research
Robert E. Ebel, Editor
American Educational Research Association
New York, New York

Handbook of Research on Teaching
N. L. Gage, Editor
American Educational Research Association
Rand McNally and Company
Chicago, Illinois 1963

Journal of Research in Science Teaching

The Second Handbook of Research on Teaching
Robert Travers, Editor
Rand McNally and Company
Chicago, Illinois 1973.

The purpose of this review of the literature is to:

a. establish support, direction and a theoretical base for a study on methods to improve hypothesis generation, and

b. verify this study as original and not a duplication of previous investigations.

Hypothesis Generation

Student Development and Hypothesis Generation

The formation of hypotheses which are reasonable, verifiable by testing, and precise is a primary goal in the development of the problem solving skills associated with scientific inquiry. As a result, previous studies in hypothesis formation have been devoted to developing strategies which would help students determine more efficiently the relationships between variables. While these major hypothesis generation investigations (Suchman, 1962b; Quinn, 1971; Wright, 1974) have differed in purpose, each has presented the participants discrepant events which have been especially developed and, subsequently, proved effective. Such caution has been necessary because a clear understanding of the problem is essential for the formation of hypotheses which will be of educative value (Renner and Stafford, 1972, p. 30). Although obvious, it should be emphasized that students can only participate in the process of isolating variables if a problem is perceived. Logically, there are two conditions which must be met for a problem to be appropriate for hypothesis formation studies:

1. the problem must present an original discrepancy in the mind of the student, and

2. the student must possess the appropriate level of mental development in order to relate the variables so to form hypotheses.

While both of the previous concerns are crucial for optimal results, the latter is especially important since it applies to cognitive development. Since the purpose of all hypothesis formation investigations is to ultimately enhance cognitive development, it is crucial that: (a) the selected discrepant event will challenge students in such a way that they will build their skills and (b) the experimental measures account for student improvement regardless of the degree of cognitive development possessed initially. A discussion of mental development in relation to hypothesis formation is, therefore, worthwhile.

The work of Jean Piaget strongly supports the role of cognitive development and the resulting ability to understand a problem and generate hypotheses. Basically, Piaget's explanation for problem recognition involves the concept of adaptation which is further composed of the processes of assimilation and accommodation. As an individual is exposed to various stimuli, assimilation (or the taking in) of the information occurs. As the information is assimilated, it is digested and integrated into the individual's existing cognitive framework. If, however, no similar information has been previously integrated, the individual must accommodate, or adjust, existing behavior and cognitive structures. As a result, accommodation to new situations means intellectual development. Through activities associated with hypothesis formation, the student is presented a problem which is not part of the already existing cognitive structure so subsequent accommodation by hypothesis generation increases mental development. For an individual to first visualize a problem, it is necessary to be at a stage of mental development where the problem can be detected.

There are four stages of development which Piaget feels relate to cognitive growth: sensory-motor (birth to about 2 years), preoperational (2 years to 6 years), concrete operational (6 years to 12 years) and formal operational (12 years to 15 years). As an individual develops those patterns of mental structure and behavior which are characteristic of a specific stage, succession to the next higher stage occurs. As a result, the individual develops into a more complex person by integrating the processes of the previous stages. The major point is that individuals must be presented with a problem that is consistent with the developmental stage for a clear understanding to occur.

An elaboration of hypothesis formation is the ability to think in a hypothetico-deductive mode. The importance of this skill is that it manifests itself during the period of formal operations and involves the ability to use inferential thinking that leads a hypothesis through all possible logical conclusions (Pulaski, 1971, p. 70). Simply, this suggests an ability to hypothesize abstractly.

Thus far the importance for students who have achieved the level of formal operations has been emphasized for optimal hypothesis formation behavior. Certainly, this assertion is relevant since students can only grasp a principle when their cognitive development can deal with the components regardless of the type of instruction used to explain it (Mallon, 1976, p. 32). However, it does not follow that all secondary students operate at the level of formal operations. Sayre and Ball (1975) found, in a comparison of 214 junior high and 205 senior high students, that the number of students who have reached the level of formal operations increases with age and is more closely associated with scholastic grade. This implies that there are more formal operational thinkers in senior rather than junior high and that formal operational thinkers achieve higher scholastic grades. A sizable number of average students are, therefore, not formal operational thinkers. Further, Lawson and Renner (1975, p. 355) found that 85 per cent of a group of 134 high school students had not yet reached the highest level of formal operations. Both these studies provide evidence that many students exposed to inquiry ideas may not be able to benefit. These findings indicate a disparity between the cognitive ability of the secondary students and the purpose of hypothesis formation studies. Fortunately, however, the difference is more apparent than real since previous hypothesis generation instruction studies have been successful due to the emphasis placed on an explanation about a problem in a filmed setting where several variables were easily recognizable (Suchman, 1961; Quinn, 1971; Wright, 1974). Consequently, students were encouraged to observe and relate the obvious variables and not necessarily manipulate variables in their efforts to generate hypotheses. In this respect, hypothesis generation studies have sought to determine the optimal methods by which students could improve already existing cognitive skills.

In an effort to qualify the hypotheses formed by students, Quinn (1971) developed a Hypothesis Quality Scale (see p. 72) that rated hypotheses along the continuum from no explanation (the lowest level) to an explicit statement of a test for a hypothesis. Thus, the highest level of the scale is analogous to the highest levels of formal operations. Further, Quinn (1971) and Wright (1974) found that students hypotheses, after instruction in hypothesis formation, improved in quality. Hypothesis generation studies, therefore, serve the dual function of promoting science process instruction and enhancing higher levels of cognitive development. This is certainly important since Piaget found social transmission an important factor to promote acquisition of higher levels of mental operations (Pulaski, 1971, pp. 11, 33).

In summary, the significance of the preceeding discussion has been to examine and support the idea that worthwhile hypothesis generation investigations must:

a. involve students who have the capacity to eventually function on a hypothetico-deductive level, and

b. present problems to the students that will be visualized as real and solvable.

Fortunately, previous studies have included both of the above conditions. In fact, specific studies utilizing sixth grade participants (Suchman, 1961; Quinn, 1971) and ninth grade participants (Wright, 1974) have indicated that students of these ages can be effectively instructed to improve their ability to identify and relate variables to generate better hypotheses about specific discrepant events of physics causality. These studies are treated in greater depth later in this review.

Single or Multiple Hypothesis Generation

Although a student may possess a clear understanding of the problem and the ability to think hypothetico-deductively, methods of instruction which emphasize the formation of only a single hypothesis may not be of optimal educative value since students become more attached to a single hypothesis and, then, lose perspective about the problem and possible solutions (Woodburn, 1969, p. 333). The alternative is to utilize activities that emphasize the formation of multiple hypotheses.

The concept of multiple hypotheses is certainly congruous with scientific inquiry since it allows for greater flexibility in isolating all aspects of a problem. Justifiably so, many of the previous hypothesis generation studies have included the number of hypotheses contributed as a measure of effectiveness.

Specific Approaches Which Encourage Hypothesis Generation

Method of Problem Presentation

A student must be confronted with a recognizable problem before the process of generating hypotheses can occur. Therefore, the important concern for researchers is the most effective method to present a problem or discrepant event to students during an instructional session. Naturally, an obvious approach is a live demonstration of the discrepant event and, then, continuation of the instruction session. Live demonstrations, however, present several difficulties:

a. can all students easily view the demonstration,

b. will various distractions from the classroom (or instructional area) detract from the effectiveness,

c. does the demonstrator maintain identical conditions over repeated demonstrations, and

d. are any excellent discrepant events not conducive to a live demonstration?

Suchman (1962b, p. 52), as he began his preliminary work in developing inquiry training techniques, encountered such problems. Since he favored the recording of each discrepant event episode on film, the primary question was the presence or absence of differences in motivation between the students who observed the discrepant event on film and those who observed the identical event in a live setting. After several investigations, the results indicated that the motivation was identical for both groups. Suchman (1962b, p. 53) did find, however, that color film was more effective than black and white film. This was attributed to the live and vivid qualities of color film.

In a somewhat related study, DeTure and Koran (1975) found that videotape recordings, where a few selected members of the class illustrated laboratory techniques, resulted in increased involvement of their classmates with process skills so students exhibited more positive and less negative behaviors during subsequent laboratory exercises. The above studies indicate that demonstrations which are recorded on film or videotape are just as effective methods to illustrate a problem and motivate students as are live demonstrations. Further, it did not appear to matter whether strangers (as in Suchman's films) or classmates were the demonstrators. These findings are basic in a discussion of hypothesis generation studies since many have utilized single topic films to present discrepant events to the participants.

Two studies attempted to determine the effect of single topic films on the ability of students to develop an improved capacity to write hypotheses. Gibbs (1967) found that high school biology students, who received five lessons in hypothesis construction via films, increased their ability at writing hypotheses. The measure of improvement for the dependent variable was relevance. In a similar study, Barker (1969) developed and evaluated four single topic films as the means to promote the construction of hypotheses that explain chemical phenomena. Basically, a problem was introduced to the students by the film, after which they were requested to write only one hypothesis. Then more of the film was shown that illustrated some experimental evidence about the original problem. Students were then permitted to revise their original hypothesis or construct a new one. The findings indicated that after the fourth film there was an improvement in the ability to construct revised hypotheses as measured by three forms of relevance.

These studies illustrated that single topic films can improve the ability of high school students to write hypotheses about a presented problem. The criteria of hypothesis quality was a general term--relevance. But this is important since these studies indicated a means to qualify hypotheses generated by students. Further, the instruction for both studies was viewing a film, receiving information and practice at writing one hypothesis to explain each event. Neither offered a specific method of intensive instruction. Other studies have incorporated the use of film to present the problem and a specific form of intensive instruction.

Specific Instruction Studies in Hypothesis Generation Directly Applicable to the Classroom

Although, over the past twenty years, a number of investigations have been conducted in the area of hypothesis generation, much of the work has dealt specifically with the process of hypothesis formation and not with techniques which can enhance the process as additions to existing classroom instruction. Therefore, the following discussion will begin with the three specific studies which have involved instruction methods incorporating problems or discrepant events which are feasible in a science classroom setting. Then, pertinent works which have investigated general aspects of the hypothesis formation process will be capsulized under the category of related studies.

Inquiry Development Program

Suchman (1961, 1962b) in the late fifties to mid-sixties designed, developed, and distributed a comprehensive Inquiry Development Program (IDP) intended for use in actual classroom settings. The objectives were to help students build those inquiry skills necessary to investigate causal relationships. As a result, the activities emphasized: (a) gathering and organizing data, (b) isolating variables, (c) hypothesizing relationships between variables, and (d) testing hypotheses through verbalized experiments.

Suchman felt the Inquiry Development Program needed no system of rewards or reinforcement since the need to inquire into the cause of a puzzling situation was self-rewarding as information became available. To insure this condition, problems were prepared which were readily apparent since a limited number of variables were involved. Also, in this way, students could focus their concentration on their thinking process as well as the problem. The problems were limited to those of physics causality which could be recorded on color film. (For classroom use the eight millimeter filmloop format was prepared.) In total, thirty-three films of approximately two minutes in length were developed. In order to maintain consistency among the films, each problem was demonstrated using recognizable science equipment.

The procedures for each inquiry instruction session lasted for one to two hours per week for a period of twenty-four weeks. After the sixth grade subjects viewed the film, they were requested to ask the teacher questions about causality that were specific enough for a yes or no response. In this way, open-ended questions were avoided so that the students focused on the variables in such a way to ask pertinent questions. As a result, the students--not the teacher--controlled the amount and quality of information gathered during the session. To maintain order during the session, Suchman limited the number of active participants to groups of ten. Since the rest of the class was present, the non-participants observed the film and questions so they could improve their abilities. Further, once a student had "the floor" a series of questions could be asked. Each session was audiotaped and, after the session, the tape was replayed so the teacher could analyze the usefulness of each question. This critique was intended as feedback to students so they could improve their skills.

To assess the effectiveness of the program, a year long study was conducted in twelve schools across the nation. The inquiry leaders from each school had participated in a summer training session directed by Suchman. The procedure in each school involved one treatment (inquiry) intact class and one control intact class which watched the same films as the inquiry class but received an expository lesson about the concepts. At the end of the study, the findings indicated:

a. equal, if not better, conceptual growth for the inquiry classes, and

b. better process of inquiry acquisition by the inquiry classes as measured by a greater fluency (Suchman, 1962b, pp. 117-23).

The importance of the Inquiry Development Program was that it outlined a feasible method to intensively instruct students in data gathering. There is an assumption, in Suchman's work, that each question is a hypothesis in an interrogatory form. Further, the Inquiry Development Program indicated small group discussions about the problem were essential for good participation and, also, that the critique feedback was a useful portion of the instruction.

Inquiry Instruction and Hypothesis
Generation in the Classroom

Although Suchman indicated that inquiry skills were increased by small group instruction, no attempt was made to actually determine the effect of the instruction on hypothesis generation. Quinn (1971) undertook such a study with sixth grade students from one urban and one suburban school each of which contained homogeneous groupings of students from low and middle socio-economic homes respectively. Quinn developed a validated and reliable scale to measure the quality of hypotheses produced by students.

The procedures for Quinn's study involved a treatment and a control group (both intact classes) in each school. Treatment involved twelve film sessions and six discussion sessions. The film sessions were additions to the Suchman Inquiry Development Program since the discussion sessions were detailed analysis of students' observations in terms of possible inferences, relationship of variables, and quality of hypotheses. It should be noted that the observations were written on the blackboard so they were visible to the students. The results of the study indicated:

a. subjects who received instruction in hypothesis generation provided a greater quality of hypotheses than subjects not so instructed,

b. instructed subjects from middle socio-economic homes did not generate a greater quality of hypotheses than subjects from low socio-economic homes,

c. students with high Otis Test of Mental Ability Scores generated a greater quality of hypotheses than subjects with low Otis Scores,

d. students with high grade point averages generated a greater quality of hypotheses than those with low grade point averages,

e. subjects with high reading levels generated a greater quality of hypotheses than students with low reading levels,

f. girls generated a greater quality of hypotheses than boys, and

g. individual subjects with hypothesis generation training (regardless of socio-economic status) generated a greater quality of hypotheses than individual subjects without hypothesis training.

The importance of Quinn's research to the present study is that it illustrated intact classroom instruction can improve the ability of students to generate a higher quality of hypotheses as determined by the analysis of written hypotheses utilizing a validated scale.

Intensive Instruction and Hypotheses
Generation in an Individualized Setting

Since Quinn's method of classroom instruction was proved successful, a worthwhile next step was a comparison of different instruction methods in an individualized setting. Wright (1974) conducted such a study utilizing ninth grade subjects who were randomly placed into one of three treatment groups--control, hypothesis generation intensive instruction, or cue attendance (number of details observed) intensive instruction. The hypothesis generation intensive instruction consisted of watching an Inquiry Development Program Film and, then, generating five acceptable hypotheses; the cue attendance instruction consisted of viewing a film and, then, generating seventy-five details. As the student would offer an acceptable response, it was reinforced with "ok, good, very good." After treatments, comparisons were made between each intensive instruction group and the control as well as between treatment groups.

The dependent variables were the quantity and quality of hypotheses generated, the quantity of cue attendance, and the quantity and diversity of information search. In addition, aptitude treatment interactions were determined using scores from Verbal Reasoning and Hidden Figures tests as variables. The major findings indicated:

a. each intensive instruction group performed better on each dependent variable,

b. the treatment groups performed equally on all dependent variables but cue attendance where the intensive instruction cue attendance group did better,

c. no differences existed between the three high and three low ability groups, and

d. no significant aptitude treatment interactions existed between intensive instruction in cue attendance or hypothesis generation for Verbal Reasoning or Hidden Figures levels.

Fifteen months later a follow-up study was conducted to determine if difference results for the dependent variables would occur (Wright, 1975). For cue attendance quantity as well as quantity and quality of hypothesis generation, the treatment groups were significantly better than the control groups. Further, there were no significant differences between the treatment groups. As for information search, the cue attendance group was significantly better in performance than the hypothesis generation and control groups. The major implication is that hypothesis generation behavior is improved by either cue attendance or hypothesis generating instruction.

In both studies, Wright indicated that intensive instruction, which involved predetermined minimum standards for hypothesis generation was effective in an individualized setting to promote immediate and long-term inquiry skills. The importance of Wright's work was that it verified:

a. more than one form of hypothesis generation instruction is applicable in a secondary school setting,

b. predetermined minimum standards can be set for instruction and acceptable responses can be reinforced,

c. Inquiry Development Program Films present adequate discrepant events to ninth graders for hypothesis generation, and

d. intensive instruction yields persistent retention of basic inquiry skills.

In summary, the three preceeding investigations are important to the present study since the findings of each provide support. Suchman (1962b) provided evidence that discrepant events can be filmed and, then, utilized as part of a training program for the process skill of isolating variables by questioning. Quinn (1971) continued inquiry training by first developing a scale to measure the quality of hypothesis generation and, then, by exposing students to classroom instruction in hypothesis generation. The results indicated that students can be instructed to improve hypothesis generating ability. Wright (1974), further, found that hypothesis generating behaviors can be improved by individualized reinforced instruction in either hypothesis generation or observing details.

While the work of Quinn and Wright have been valuable extensions of Suchman's original research, together they raise three questions which are worthy of further research:

1. What is the difference between individualized instruction in hypothesis formation when only the criteria are given to the student and individualized instruction in hypothesis formation when only reinforcement or both criteria and reinforcement are given?

2. What effect will individualized intensive instruction have on a group discussion about a specific discrepant event? And,

3. what effect will the group discussion have on the members of the instructed groups to subsequently generate hypotheses?

The present study attempted to provide answers to these questions in an effort to further the information available about hypothesis generation instruction techniques. Thus, this study further narrowed the gap between research findings and classroom application.

Related Studies

General Ideas

As previously mentioned, a number of investigations have been conducted which deal specifically with the process of hypothesis formation. In addition, investigators have studied general aspects of hypothesis generation. The discussion of related studies will, therefore, begin with general aspects and continue with specific findings about the process of hypothesis formation.

In 1930, Tyler (1930) studied the inference processes of college students in an elementary zoology class. The subjects were presented with zoological facts from which inferences could be drawn. Since care was taken to make the facts as "original" as possible, the inferences were a result of thinking and not memory. First, the subjects were presented with a series of items for which they were required to write an inference. Later, in the same day, the subjects were given the same items with multiple choice responses and required to select the best inference. Although the findings indicated little correlation (.38) between the ability to formulate an inference and the ability to select the best inference from a group, Tyler was able to obtain a method for determining the quality of generated inferences. Using a system of rating each inference by a panel of judges, he found a correlation between judges of .96. The significance was that given a group of written inferences about a discrepant topic a panel of judges could objectively rate the hypotheses of individuals in general (good, bad or moderate) or specific (point system) ways.

Along similar lines, Atkin (1958) investigated: (a) the nature of elementary students' development and their ability to formulate hypotheses and (b) the relationship between type of classroom and the development of problem solving abilities. The results indicated that older children tended to rely on a greater use of authority figures as the basis for their hypotheses and this effect was most pronounced in less permissive classrooms. Naturally, there were more original responses in more permissive classrooms. Another important point was that the investigation substantiated the appropriateness of using a predetermined standard to determine the quality of generated hypotheses.

A more recent correlational study of hypothesis generation and student traits was conducted by Brown, D. (1973). The subjects were 108 female undergraduates enrolled in an independent study introductory biology class. Further, each participant had no previous college science instruction. The findings indicated:

a. as independent study was increased so was precision in hypothesis formation, however after the fourth session there was a decrease which was explained by a fatigue factor,

b. as independent study was increased a greater number of deductive hypotheses occurred, and

c. while there was a significant inverse correlation between anxiety and object visualization, there was no correlation between anxiety and the type of hypothesis formed (precise and deductive).

Brown indicated that hypothesis generation was positively correlated to independent study but after a certain point a fatigue factor affected hypothesis generation. Further, deductive hypotheses were stated more precisely. Similarly, Atkin showed that elementary students offered different types of hypotheses but these were determined by age and degree of permissiveness in the classroom. There appears to be an environmental factor suggested by both Brown and Atkin. Tyler's work, while it did generally describe student traits and inference drawing ability, emphasized the need for a problem to be original so the subsequent hypotheses could be qualitatively evaluated.

In addition, mention of a finding from a recent investigation by Sprafka (1973) may be worthwhile. Utilizing medical students, the study determined if individuals who were constrained to verbalize during an individualized hypothesis generating session would generate a greater number of hypotheses. The findings indicated this was only the case on one of three problems. Thinking out-loud was, therefore, not an effective technique in utilizing inquiry skills.

Instruction in Hypothesis Generation

Many studies, conducted at the college level, offer many points which are worthwhile to consider. One such study was reported by Klein, Frederiksen, and Evans (Klein, et. al., 1969) who utilized 127 paid male freshman and sophomore "volunteers" from Rutgers. Basically, the subjects were randomly placed into experimental and control groups that received treatments (and observations) simultaneously on three consecutive evenings. The dependent variables were: (a) subject anxiety, (b) the quality (number corresponding to a list of acceptable) of hypotheses, and (c) the quantity (number reported) of hypotheses on the Formulating Hypothesis Test. The independent variables were illustrations of acceptable hypotheses from a predetermined list. The findings indicated a greater quantity of hypotheses from the treatment group but no significant difference between the control and treatment groups for quality of hypothesis generation due to the amount of anxiety.

Several years later the preceeding study was modified and repeated (Frederiksen and Evans, 1974). The dependent variables were clearly delineated as:

a. quantity of hypothesis generation--those responses not previously mentioned,

b. quality of hypothesis generation--those responses that corresponded to ideas previously agreed upon by a panel of judges,

c. quality of hypothesis generation--the average number of points for the quality of hypotheses as rated by two judges,

d. quality of hypothesis generation--based on length, handwriting, or grammatical correctness, and

e. quality of hypothesis generation--based on the average number of words per response.

The subjects included 395 paid male and female "volunteers" from two Pennsylvania colleges who were placed into three groups--control, quality treatment, and quantity treatment. Similar to the previous study, the independent variables were proper illustrations for either quality or quantity hypothesis formation. The findings were that each treatment group generated a greater quality or quantity of hypotheses with fewer words per response. Also, females generated more responses than males. These two studies indicated that college students can be instructed to generate a greater quality and quantity of hypotheses as determined by a systematic method of rating. It should be noted that these studies were designed to compare the effect of treatment not to compare different treatment methods.

A comparison of three different methods of instruction (attribute block material, pictorial logic, and basic elements) on second and third grade students was conducted by McGinty (1972). The findings indicated that instruction had positive effects on certain logical abilities of the subjects. Further, it was found that the third graders out-performed the second graders when the same instruction method was compared. The significant point of this study is the comparison of all three independent variables on the outcome of each dependent variable.

Similarly, Salomon (1970) employed two training procedures (cue attendance and hypothesis generation) and two levels of training (whether criteria were met) and two kinds of training design (structure and unstructured) in a study on response uncertainty in teacher interns. Although subjects were not specifically instructed in how to determine quality hypotheses, they were positively reinforced when an acceptable hypothesis (or detail) was submitted. To present the problem to the subjects, a segment of a motion picture of approximately four minutes was either presented as it was or in a randomly spliced arrangement. Subjects were required to meet criteria of 150 details or seven hypotheses during instruction that utilized the problem. In addition, a subject could observe the film as often as desired.

In order to determine the effects of the instruction methods, the subjects were shown slide sequences and, then, asked to generate either hypotheses or details. Again the subject could view the problem (slides) as often as desired. The subsequent findings included:

a. immediately after intensive instruction, the treatment groups displayed greater hypothesis generating and cue attending behaviors,

b. after seven days, the treatment groups displayed greater information search behavior,

c. intensive instruction utlizing unstructured films was more effective for hypothesis generation and utilizing structured films was more effective for information search, and

d. intensive instruction participants displayed more subjective uncertainty.

The importance of this study was the reinforcement used to illustrate acceptable hypothesis generation and the comparison of immediate and delayed (week later) improvements.

To what extent hypothesis generation occurs was partially determined by Byers (1965) who tested twenty-four college educational psychology students for the relative frequency of hypotheses and the amount of information remaining after the first hypothesis occurred. The procedures, which were complex, involved a question board and card sequence arrangement. The findings were that students delayed overt expressions of hypotheses until they had more information about the concept. The significance, however, is in the discussion where Byers states that positive reinforcement increases hypothesis generation frequency (Byers, 1965, p. 342).

Intensive Instruction Methods for Concept Attainment

The previous studies in the area of hypothesis formation have been devoted primarily to determining the effect of one form of instruction on the ability to generate hypotheses. Since the present study continued the previous work by comparing more than one method of hypothesis generation intensive instruction, it is worthwhile to discuss a few pertinent studies in concept attainment which have similarly compared various methods of instruction about specific concepts.

Kersh (1962) sought to compare the effect of instruction about two novel rules of addition using three methods of instruction--guided discovery, directed learning, and rote learning. The effect of each was determined by a test of recall given three days, two weeks, and six weeks after instruction. Utilizing three groups of high school geometry students, it was found that: (a) the rate of forgetting was constant for all groups and (b) the rote learning group did significantly better than the other two (although the guided discovery group outperformed the directed learning group). These findings were at variance with previous work (Kersh, 1958) which indicated discovery learning was superior to learning with direction. The disparity was attributed to retroactive inhibition which meant the treatment groups retention was inhibited by interpolated learning. The importance of this work was that differences occurred due to different methods of instruction and that giving students too much information may not facilitate transfer of previous learned principles.

As a result of studies which measured to what extent a concept had been established, Gagné and Brown (1961) investigated the effects of variations in the programming of conceptual learning materials on the effectiveness of learning (as measured by performance in a problem solving situation). The participants, who were thirty-three boys from ninth and tenth grades, were placed into three treatment groups--rule and example, discovery, and guided discovery. After the instruction, the guided discovery group was the most effective--while the rule and example group the least effective--in solving a new problem. This study established that three different approaches--which presented the students various amounts of information--did not succeed equally in promoting concept attainment. Further, the results tended to verify that the students needed to discover on their own but with the aid of some direction.

Sechrest and Wallace (1962) compared the assimilation and utilization of information when subjects, who were instructed with different conditions of information transmission, attempted to determine a concept. The premise of this work was that less than perfect performance may be attributed to failure to use all available information. Among the four groups of college psychology (introductory) students, there were no significant differences and, therefore, it was concluded all subjects assimilated information efficiently. Each group differed by the number of clues they received during instruction. The importance of this study was not the findings but, rather, the discussion. The investigator found in addition to the results:

a. the earlier subjects ventured hypotheses the fewer instances were required for a solution, and

b. some students used the experimenter's invalidations of their hypotheses as a source of information while others used them as a source of punishment. The latter group, naturally, was reluctant to engage in quantity hypothesis generation (Sechrest and Wallace, 1962, p. 163).

Working from an assumption that subjects can only discover how to apply a rule, not discover it, Wittrock (1963) sought to determine the effectiveness of giving subjects the rule, the answer, neither the rule or answer, or both. The results indicated that the college psychology juniors who were given the rule and answers or just the rule did better than the other two groups. The no-rule no-answer group was the least effective. The point of this study was that activities which emphasized discovering the application of a principle were superior means of instruction.

The evidence indicates that various forms of a treatment in concept attainment studies do not yield equal results. Since previous intensive instruction studies for hypothesis generation have not generally considered various methods of instruction comparisons, this is a logical area to investigate. The data from the preceeding investigations indicate subjects with guided discovery did better in applying concepts. Similarly, this study attempted to determine if one specific method of hypothesis generation intensive instruction was superior in promoting student application of the concept of hypothesis formation. It appears that the studies discussed in this review definitely support the need for the present investigation.

A further point of interest is the question--would all students react identically to varying methods of intensive instruction? A recent study indicates a negative answer to the preceeding question. Graybill (1975) investigated sex differences in problem solving of selected science problems taken directly from the work of Piaget and Inhelder. The results indicated that males ages nine, eleven, thirteen and fifteen outperformed females and, also, were more confident in handling equipment and less aware of the presence of the experimenter. These results, when considered with previous studies where female participants formulated better hypotheses than males (Quinn, 1971; Frederiksen and Evans, 1974), leave the questions about the role of intensive instruction beyond the scope of this study. However, it appears that intensive instruction may offer all students the flexibility to react differently but ultimately the ability to master a specific performance objective. Therefore, investigations which attempt to illustrate feasible activities for developing concept attainment in the classroom are worthwhile.

Summary of Related Studies

The related studies dealt primarily with three areas of hypothesis formation:

1. general ideas,
2. instruction in hypothesis generation, and
3. intensive instruction for concept attainment.

The general ideas studies have indicated that the quality of hypotheses generated by students can be assessed. Further, a few of these studies have indicated that the type of learning environment can affect the type of hypothesis formation by elementary students (Atkin, 1958) and college introductory biology students (Brown, D., 1973). The major importance of the instruction in hypothesis generation investigations is that each illustrated, in the presence of one form of instruction and, also, a measurement instrument, hypothesis formation abilities can be increased. Further, Salomon (1970) illustrated that positive reinforcement was a useful instructional condition for students when a predetermined standard was attained. Byers (1965) similarly found positive reinforcement increased the quantity of hypotheses formed by students. Finally, the concept attainment studies have illustrated that guided discovery is a superior instructional method for students to understand and apply a concept than rote learning or total discovery.

Relationship Between Hypothesis Generation and Creativity

Inquiry Training and Creativity

Naturally, any addition to the curriculum which emphasized an infinite view of knowledge was considered to develop creative thinking or creativity. In fact, Suchman (1965, 1962a) indicated that the Inquiry Development Program enhanced creative thinking abilities of the participants. This is logical when the nature of the discrepant events (as the promotors of enlarged conceptual system) is considered. Further, inquiry is self-motivating and as such maintains a high interest level which results in more fluent, precise analysis of the possible causes of the problem. Since critical thinking is an autonomous process which is self-directed and aims toward the production of a new form, the Inquiry Development Program developed critical thinking.

There have been two studies on the impact of inquiry curricula and activities on creativity. Brown, T. (1973), in a comparison of the inquiry curriculum Science Curriculum Improvement Study (grades one to six) and a conventional science curriculum, found the inquiry students were significantly more creative. In addition, it is important to mention that the participants had only inquiry or conventional instruction for six years. Contrasting results were obtained by Bills (1971) in an assessment of a weekly inquiry lesson on the creativity productions of students. It was found that 306 eighth grade students (taught by six teachers) did not make significant gains due to the inquiry instruction. The investigator, however, admitted creativity was a difficult term to accurately assess (Bills, 1971, p. 420).

Creativity--In General

Guilford (1967, pp. 108-09) listed three basic traits of creativity:

a. fluency as measured by quantity,

b. flexibility as measured by a change, and

c. originality as measured by the presence of the unusual or somewhat removed.

Further, Guilford included mention of practicality in the sense of feasibility. Guilford's distinctions are important since the Hypothesis Quality Scale of Quinn (see page 72), although not intended as a measure of creativity, certainly accounts for creative responses. For example, a student who performs well will have isolated variables and manipulated them into various hypothetical situations--flexibility and originality.

In addition, a creative person must have a high sensitivity to the problems so to sense imperfections (Guilford, 1967, p. 118). In this way, complacency with a single hypothesis is not creativity! The Inquiry Development Program definitely is based on a philosophy that encourages the creative student.

Creativity and Classroom Activities

Brainstorming has the advantages of quantity but is it at the loss of quality? This is an impossible question to answer! However, the ideas gained from a group discussion setting are of a higher quality when the fear of being wrong is not reinforced (Guilford, 1967, p. 114). Other conditions which further stimulate group creativity are: preparation, attitude, open-mindedness, receptivity, enthusiasm, concentration, and expression (Rapp, 1967). The group inquiry session of Suchman contained each of these criteria. For this reason, inquiry probably encourages group creativity and, consequently, individual creativity. As previously mentioned, however, creativity is difficult to assess.

Although Bills (1971) and Brown (1973) attempted to assess the impact of science inquiry skills on creativity, they really only measured a small aspect of a broad topic. Since creativity is a self-initiating trait, the chances of developing creative behavior is maximal when classroom activities are interesting and challenging. Actual practices, naturally, depend on the teacher. To aid teachers in their effort to promote creative activities in the classroom, articles have appeared in secondary science teaching journals over the last fifteen years (Kilburn, 1963; Coleman, 1966; Micciche and Keany, 1969; Ankney and Sayre, 1975). As the teaching of hypothesis generation skills becomes a popular addition for science curricula, more emphasis will probably be placed on the role of creativity as reflected by hypothesis generating behaviors.

Summary of Important Findings

The major findings of previous studies that have been involved with intensive instruction and assessment or hypothesis generating skills include:

a. hypothesis generating behavior in students can be improved by intensive instruction,

b. information search behavior in students can be improved by intensive instruction,

c. intensive instruction for hypothesis generation can be effectively conducted in an individualized, small group or classroom setting,

d. as long as a problem meets the criteria as a discrepant event, it can be filmed and, then, effectively motivate students,

e. quality of hypotheses can be determined,

f. hypothesis generation skills are transferable immediately and over a period of time,

g. various methods of intensive instruction result in different levels of concept attainment,

h. creativity--although related to hypothesis generation--cannot be quantitatively measured,

i. regardless of intensive instruction, the ability to generate hypotheses varies according to sex, and

j. intensive instruction is effective in elementary and secondary schools as well as colleges to promote the formation of hypotheses.

None of the previous studies have compared the effect of giving students either criteria or a form of reinforcement as an intensive instruction condition. Further, none of the studies have compared the effect of previous hypothesis formation instruction on group questioning behavior or the effect of group discussion on subsequent hypothesis generation. This study (a) compared four different hypothesis generation intensive instruction methods and (b) determined the effect of each method on information search behavior during a group discussion and subsequent written hypothesis generation behavior. Chapter III outlines the procedures utilized and the design of the study.

CHAPTER III

PROCEDURES

Design of the Study

A 2 x 2 factorial design was employed. There were four independent variables--criteria (given and not given) and a form of reinforcement (differentiated and undifferentiated). In addition, there was a control group which received no treatment. Table 1 (below) illustrates the combinations of the independent variables for each treatment.

TABLE 1.--Combinations of the independent variables and the control

	Differentiated Reinforcement Given:	Undifferentiated Reinforcement Given:
Criteria Given:	Group A	Group B
Criteria Not Given:	Group C	Group D

Control:
Group E

The study was conducted during February and March, 1976 at Frederick Sasscer Junior High School located in Prince George's County, Maryland. The participants were the ninth grade students who attended science class for one period (fifty minutes) each day. In addition, all twelve science classes of the two ninth grade science teachers were involved. In each intact class, the participants were randomly placed into one of the four instruction groups or the control group. The only exception to the randomization was the sex composition of each group which was stratified in the same proportion that existed in the intact class since previous researchers (Graybill, 1975; Quinn, 1971) had found differences in process skill behavior due to the sex of the participant.

To minimize the possible interaction affect between the students of the various groups, during each class period, students from the classes of both teachers were simultaneously exposed to intensive instruction. This procedure was possible by employing an instruction sequence as described by a Greco-Latin square (Dayton, 1970, pp. 149-50). Table 2 (below) illustrates how the instruction sequence was based on (a) the group, (b) the day of the study, and (c) the class period.

TABLE 2.--Group intensive instruction sequence utilizing individuals from both classes by the period and day of the study

Period	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6
2*	group: E_1	A_1	B_1	C_1	D_1	Make up
3**	groups: A_1 A_2	B_1 C_2	C_1 E_2	D_1 B_2	E_1 D_2	day for
4	B B	C D	D A	E C	A E	students
5	C C	D E	E B	A D	B A	who
6	D D	E A	A C	B E	C B	were
7	E E	A B	B D	C A	D C	absent

*only the classes of one teacher participated
**the classes of both teachers (numbers refer to each) participated

Group A: Differentiated Reinforcement and Criteria
Group B: Undifferentiated Reinforcement and Criteria
Group C: Differentiated Reinforcement only
Group D: Undifferentiated Reinforcement only
Group E: Control

Upon completion of the instruction sessions, during each period, students of all the groups were shown an Inquiry Development Program Film ("The Sailboat and the Fan") and asked to write as many hypotheses as they could to explain the discrepant event. Then, the students were shown another Inquiry Development Program Film ("The Ice Cubes") and requested to write as many questions as they could which might help better explain the discrepant event. Five days later, group discussions began with only one of the groups during each class period. Table 3 (see page 57) shows the discussion schedule for each group. During the discussion the students were requested to ask questions that might help them solve the problem. After twenty questions, the students were requested to apply the group information by writing as many hypotheses as possible that might explain the discrepant event. Therefore, there were seven dependent variables:

1. quantity of written hypotheses immediately following treatment,
2. quality of written hypotheses immediately following treatment,
3. quantity of information search questions immediately following treatment,
4. diversity of information search questions immediately following treatment,
5. diversity of oral questions during the group discussion,
6. quantity of written hypotheses after the group discussion, and
7. quality of written hypotheses after the group discussion.

Figure 1 (see page 58) illustrated the relationship between the dependent variables and the experimental groups.

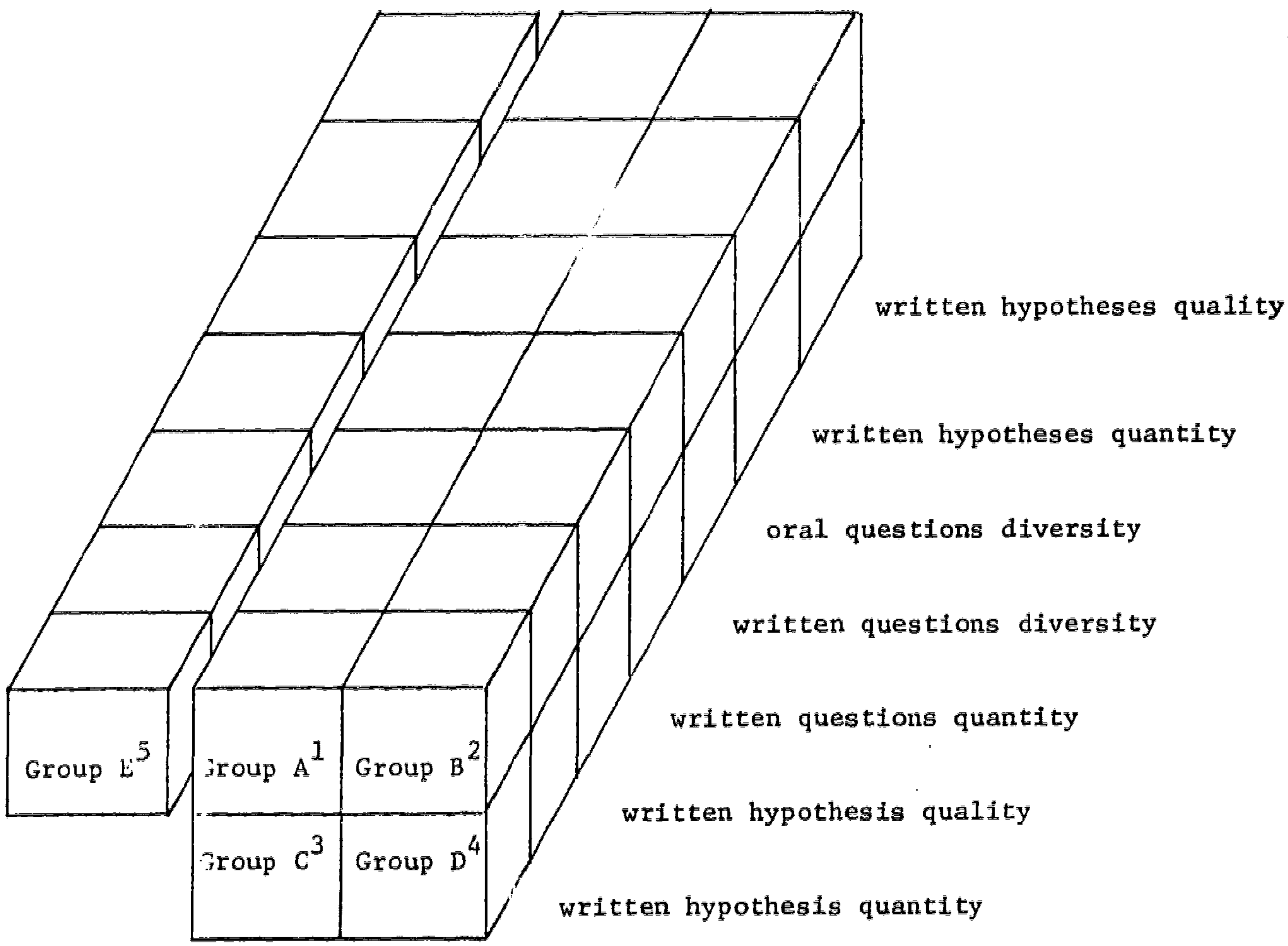
TABLE 3.--Discussion group sequence for information search behavior utilizing intensively instructed individuals from both of the classes by the period and day of the study

Period	Day 11		Day 12	Day 13	Day 14	Day 15
2*	group:	E_1	A_1	B_1	C_1	D_1
3**	groups:	$A_1 A_2$	$B_1 B_2$	$C_1 C_2$	$D_1 D_2$	$E_1 E_2$
4		B B	C C	D D	E E	A A
5		C C	D D	E E	A A	B B
6		D D	E E	A A	B B	C C
7		E E	A A	B B	C C	D D

*only the classes of one teacher participated
**the classes of both teachers (numbers refer to each) participated

Group A: Differentiated Reinforcement and Criteria
Group B: Undifferentiated Reinforcement and Criteria
Group C: Differentiated Reinforcement only
Group D: Undifferentiated Reinforcement only
Group E: Control

In summary, there were four independent variables--criteria (given and not given) and reinforcement (differentiated and undifferentiated). Immediately following the completion of the various methods of intensive instruction, the effect of instruction in hypothesis generation, was determined by having students write hypotheses and questions about discrepant events. Five days later, the students were presented an additional problem in a small group setting where they were allowed to ask questions about the variables in the discrepancy. After twenty questions, each participant was requested to write as many hypotheses as possible to explain the discrepancy. Figure 2 (page 59) illustrates the design of this investigation.



[1]Group A: Differentiated Reinforcement and Criteria
[2]Group B: Undifferentiated Reinforcement and Criteria
[3]Group C: Differentiated Reinforcement only
[4]Group D: Undifferentiated Reinforcement only
[5]Group E: Control

Fig. 1.--Relationship Between Dependent and Independent Variables

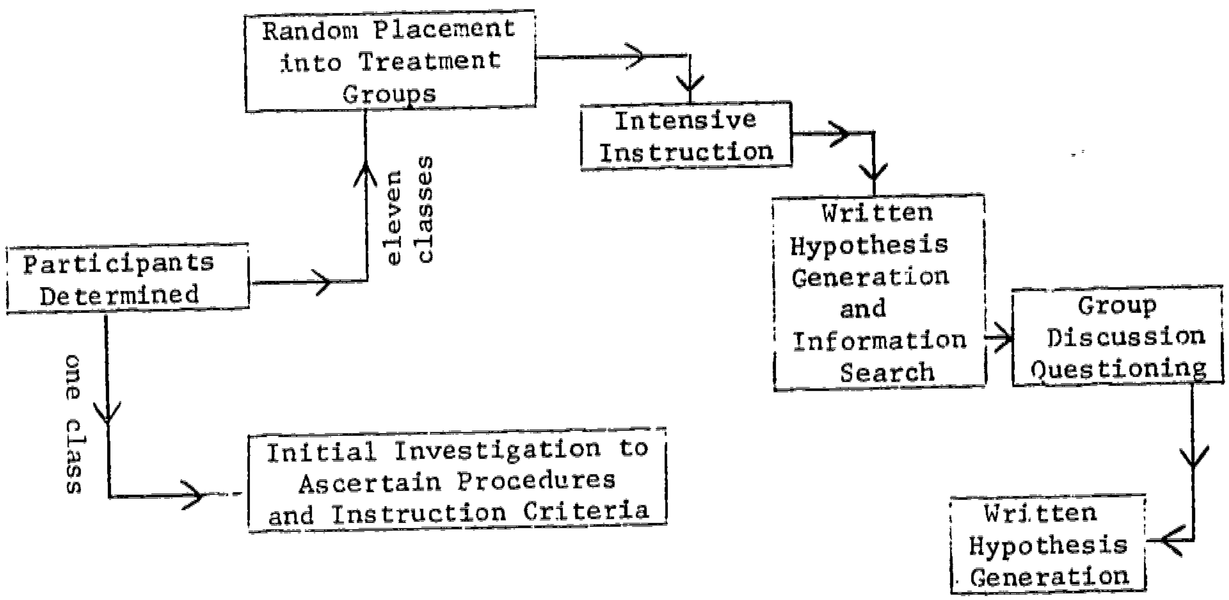


Fig. 2.--Diagram of the research design

Initial Investigation

Three weeks prior to the start of this study, an investigation was conducted with one of the ninth grade science classes (the only first period class) to determine:

a. the criterion level for the number of acceptable hypotheses which could be expected during intensive instruction,

b. the criterion level for the number of questions which could be asked by individuals during the group discussion, and

c. the need for procedural modifications--including visibility of the films, clarity of the tapes, and understanding of directions.

The findings (from the eighteen participants) indicated the suitability of the procedures for widespread application and the number of hypotheses possible and questions practical during the sessions was six and twenty respectively.

Population

The participants were ninth grade students of Frederick Sasscer Junior High School in Upper Marlboro, Maryland. As a school in suburban Washington, D. C., Frederick Sasscer is part of the Prince George's County Public Schools which is one of the largest systems in the nation. The school operates on a seven period day with approximately fifty minute classes. According to assignment procedures, the distribution of the students in each intact science class was heterogeneous.

Participation in both the initial investigation and the actual study was voluntary and, also, required parental consent. Therefore, the study was explained to all potential participants and each was given a parental permission letter (see Appendix I, p. 121). The response for participation included approximately 80 percent of the students from the eleven classes utilized in the main study. (It is important to note that students from twelve intact classes participated; however one of the classes was selected for use in the initial investigation.) The total number of participants available for the investigation was 221. Unfortunately, due to absenteeism on crucial days, only 205 or forty-one students per experimental group were utilized in the final analysis. The average age of the participants--assessed during the discussion phase--was 14 years and 8 months.

The teaching philosophy of the ninth grade science teachers was contract oriented utilizing the Earth Science Curriculum Project. Both teachers had covered the same topics prior to this investigation. Further, Frederick Sasscer serves a relatively stable student body as most of the seventh graders continue and finish the ninth grade. Therefore, most participants had similar science experiences. The participants from each class were randomly placed in one of the experimental groups utilizing their numbers from a sheet containing the alphabetized last names of the students in the class. The random numbers table provided by Dayton and Stunkard (1971, pp. 270-76) was employed.

Stimuli Presented

Inquiry Development Program Films (Suchman, 1966) were used in this investigation because previous research by Wright (1974) had shown them as effective discrepant events for ninth grade students. Since only five films were necessary, it was important to select from the complete set of thirty-three those which would present an interesting discrepancy to the greatest number of students. It was possible to immediately eliminate several because of previous student exposure due to prior science instruction which was known through discussions with science educators and prior knowledge of the investigator. As a result, a group of possible films was shown to two junior high teachers who together with the experimenter selected the six best. Then, these were shown for further verification to:

a. a science specialist in the Prince George's County School System,

b. a university secondary education professor with prior experience as a junior high teacher, and

c. two Prince George's County science teachers.

The members of the film panel viewed the films separately and completed the questionnaire included in Appendix II (see page 123). The unanimous concensus was that all six were appropriate. Therefore, the following Inquiry Development Program Films were selected by the investigator for each of these specific purposes:

a. Introductory Discrepant Event--"The Restaurant,"

b. Intensive Instruction Discrepant Event--"The Knife,"

c. Discrepant Event Immediately Following Treatment for Hypothesis Generation--"The Sailboat and the Fan,"

d. Discrepant Event Immediately Following Treatment for Information Search Questioning--"The Ice Cubes," and

e. Group Discussion Discrepant Event--"Drinking Boiling Coffee."

Hypothesis Generation Intensive Instruction

The intensive instruction was conducted in a large office and preparation room combination located between two science classrooms. There was enough room for fifteen desks and a discussion area. To minimize distractions and the amount of light entering the room, bookshelves and windows were covered with paper. Further, pegboard partitions were fastened onto the desks to create a more individualized atmosphere and to serve as support for the headphones. Each student listened to the instructions--which were recorded on audiotape--via these headphones. This method of presentation was found as excellent since it improved concentration by eliminating occasional noise from the classrooms. The participants were allowed to remove their headphones at any time after the taped instructions ended. Finally, the films were shown continuously on a screen clearly visible by all students. For this reason, the eight millimeter filmloop cartridge and the eight millimeter filmloop projector were the means of presentation.

As previously mentioned, one instruction group from each class received intensive instruction during a normal science class period. Thus, it took five days for all the participants from a class to be instructed. The average session lasted thirty-five minutes. Since the class period lasted about fifty minutes, the time was adequate. Usually, the classroom teachers sent the students--previously requested by the investigator--as a group to the "laboratory" approximately five minutes after class began. Students were allowed to sit where they wished since enough desks and chairs were available.

The instruction among the groups varied in the amount of information presented to the members by the experimenter during one or more individualized conferences. It should be noted that while the investigator could devote "total" attention to one participant, it was possible to observe the others. Students, therefore, were monitored to insure individual work. Table 4 (p. 64) illustrates the differences in information provided to each group.

TABLE 4.--The information received by each hypothesis generation intensive instruction group and by the control group

Information	Group A	Group B	Group C	Group D	Group E
Audiotaped instructions and the Introductory Discrepant Event	X	X	X	X	X
Continuous showing of the Intensive Instruction Discrepant Event	X	X	X	X	
Acceptance of written hypotheses with undifferentiated reinforcement		X		X	
Acceptance of written hypotheses only when a predetermined level was attained and then verbal differentiated reinforcement was given	X		X		
Acceptance of written hypotheses only when a predetermined level was attained but the subjects were told the criteria for good hypothesis formation	X	X			

The audiotaped instructions and introductory discrepant event comprised a general description about hypotheses presented in such a way so it was comprehensible to all participants. The initial investigation and a panel--a science education specialist and two junior high science teachers--agreed. The presentation consisted of the following:

> Mankind has always attempted to find out why something happens. Such attempts are called hypotheses and they are responsible for the world around you. Even you generate hypotheses--although you do not call them by name. For example, John may be absent from school and you may try to figure out why. Your ideas may include: he is sick, out of town, missed the bus, overslept, needed at home, and some others. As you can see, you have no way of knowing the right answer until you do some research--in this case wait until you see John again and ask him! This is

how the scientists work--observing an event, generating hypotheses, conducting research, and finally, finding an answer that will probably create more questions and more hypotheses, and more research. The process is endless!

For this experiment, you will be asked to approach a problem much like a scientist. First, you will watch a short film and, then, generate some hypotheses to try to explain it. To illustrate the procedure, carefully watch this film that will be followed by five hypotheses which would be considered good if made by a person who participated in this training. (The Inquiry Development Program Film--"The Restaurant"--is shown one time.)

Some hypotheses which may explain this event are:

a. When the waiter touched the tablecloth to remove the spot he put something on it that made the tablecloth slide out,

b. The two tablecloths were made of two different materials and the slippery one could be pulled out,

c. The waiter pulled the tablecloth out horizontally and the man pulled it up at an angle so the plates fell off,

d. If the tablecloth is pulled out fast the force of gravity keeps the plates on the table, and

e. One table had a magnet holding metal plates on the table.

<u>The following instructions were given to Groups A, B, C, and D.</u> During this session, you will generate six hypotheses that you feel explain the events in another film. Since the film will run continuously, you should do the following when you feel ready:

a. Write down between one and six hypotheses on the paper in front of you using the pen provided. And,

b. then raise your hand so you will be called to the discussion area to explain your hypotheses to the experimenter. When he feels you have reached a satisfactory level for each of your six hypotheses, you may return to class. To make this session worthwhile for you, your classmates, and the experimenter, you will be requested to do your own work and not to talk to anyone but the experimenter.

<u>The following instructions were given to Group E (control).</u> You have now finished your participation in this study for today. In a few days, you will take part in phase two. Please remove your headphones and come to the discussion area and the experimenter will give you a pass to return to class. Thank you.

The groups that had to verbally interact with the experimenter (Group A, B, C, D) were given a sheet of paper and a ballpoint pen so they could write complete or partial hypotheses which could be discussed. When ready to discuss a hypothesis, participants raised their hands and were called over to the discussion area one by one. Generally, the students were interested in the problem and appeared to do their best. As for the various degrees of information the following remarks were uniformly stated:

a. for the criteria given: an acceptable hypothesis is a reasonable explanation that could be tested if you were given the equipment,

b. for differentiated reinforcement given:

   1. good--if the hypothesis would have rated a three on the Hypothesis Quality Scale (see page 72),

   2. very good--if the hypothesis would have rated a four on the Hypothesis Quality Scale,

   3. excellent--if the hypothesis would have rated a five on the Hypothesis Quality Scale, and

   4. I cannot accept this one--if the hypothesis would have rated a two or below on the Hypothesis Quality Scale.

c. for participants in the group where there was only undifferentiated reinforcement--I can accept that hypothesis.

The intensive instruction sessions went smoothly and quickly. After the sessions, there appeared to be no "gossip"--in the sense of sharing answers--among students. In fact, the participants were enthusiastic about keeping their specific answers (hypotheses or questions) a secret. The peer pressure to make the investigation a success was evident.

Obtaining Dependent Variables

Hypotheses and Questions Immediately Following Intensive Instruction

Following completion of the intensive instruction sessions, the effects of the various methods were measured by the quality and quantity of hypotheses as well as the quantity and diversity of questions about separate discrepant events presented by Inquiry Development Program Films. To maintain uniformity in testing conditions, all the participants--from both classes--of a given period were simultaneously tested in the larger of the two science classrooms. Additional tables and chairs were added so each participant could be seated. Also, to improve visibility, the films were shown using a regular eight millimeter movie projector which produced a picture of greater intensity than the filmloop projector. Similar to the intensive instruction sessions, students were not assigned seats and paper was distributed for responses (and pens were loaned to students who needed one). Instructions for this session were audiotaped. The only contact of the experimenter or teachers with the students prior to the audiotaped instructions was a b[illegible] time after class began focus the attention of the participan[illegible] on [illegible]e electronic medium [illegible] should be added that the students were most cooperative and caused no disruptions. The instructions and procedures included on the tape were:

> Hi! During one of the last few days, you have participated in a lesson on how to generate more than one hypothesis that could explain a scientific problem. In other words, you found out that there can be more than one reason which explains why something happens. Today, you will also be asked to generate some hypotheses but about a different scientific problem. Before we go on, it should be mentioned that today's work will be written and the next phase will be verbal. Therefore, you are requested to do your own work and to please try to do your best. Since it is difficult in a

group of this size to put your group number on each paper during this session, would you please write your name and science teacher's name on your paper. If you need a pencil or pen or have another problem, please raise your hand.

As previously mentioned you are to watch some films today and then follow the directions. You should carefully watch this film two times after which you will be instructed to write as many hypotheses or reasons as you can that may explain why the problem exists.

(At this point, "The Sailboat and the Fan" is shown two times.)

Now you will have five minutes to write as many hypotheses as you can that may explain the scientific problem. Please do your best since this session is important. Also, do not worry about spelling or awkward grammar--the important thing is to write your hypotheses. (During this time the film is shown one more time.)

(After five minutes.) Five minutes are up--so please stop writing and turn your paper over. If you have written on this side of the paper, raise your hand for another sheet. As you have probably thought, if you could only ask questions about scientific problems which would be answered then you could write better hypotheses. Therefore, you will be shown another film after which you will be asked to write questions which might help you better understand the problem. Therefore, carefully watch this film two times.

(At this point, "The Ice Cubes" is shown twice.)

You will now have five minutes to write as many questions as you want that might help you better understand the problem when they are answered. Remember you are to write questions you want answered not hypotheses. Again, do not worry about spelling or awkward grammar. Raise your hand if you have any difficulties. Please do your best and concentrate on your own work. (During this time the film is shown once.)

(After five minutes.) Your time is up! Please stop writing. You have now finished your participation in this study for today. In a few days, you will take part in phase three. To make this session worthwhile for you, your classmates, and the experimenter, you will be requested to not discuss your hypotheses and questions with anyone but the experimenter. Please pass your papers to the front. Thank you.

Questions During a Group Discussion
and Hypothesis Generation Afterward

Five days after the written measures for the preceeding dependent variables, small group discussions began which were composed of those participants from both classes who were in the same treatment groups (see Table 3, p. 57). Since only one treatment group could have its discussion session each day, the procedure took five days to complete. The questions of the participants were audiotaped and written on an overhead transparency by the investigator so transcription for later analysis was possible. As a participant stated a question, the following applied:

a. the question was repeated by the experimenter,

b. written on the transparency,

c. projected for the members of the group to observe,

d. repeated by the experimenter, and

e. answered with a yes or no response.

If a question could not be answered, the investigator told the participant: "I cannot answer the question the way you stated it; it needs to be more specific so yes or no can be the answer." No other comments were made by the experimenter. To ascertain the relative absence or consistency of nonverbal cues by the investigator, four teachers, on separate occasions made an unannounced visit to one session for five minutes. They were asked to discuss their observations and determine if the procedures were consistent and devoid of nonverbal or verbal cues. The unanimous decision was in favor of uniform behavior by the investigator that neither encouraged nor discouraged student participation. It should be noted that the teachers were able to enter and leave undetected by a the entrance. Further, the time spent observing the session was long enough to obtain a general idea of the procedure yet short enough to not prohibit a student from participating if the teacher's presence was annoying. The audiotaped instructions to the participants were:

> Hi! Remember last week after you watched the film called "The Ice Cubes," you were asked to write questions that you felt might help you write better hypotheses if they were answered. Today, you are going to do the same thing except for two major differences:
>
> first, you will ask questions instead of writing them, and
>
> second, the investigator will answer your questions with a yes or no response.
>
> But, before we can begin you need a scientific problem. Therefore, please watch carefully this film two times after which you will receive further instructions. ("Drinking Boiling Coffee" is projected two times.)
>
> We can now begin a small group question and answer session about the film. The procedure will include:
>
> a. when you have a question about the scientific problem that you feel might help you generate better hypotheses, raise your hand,
>
> b. when you are called upon state your question to the experimenter, (Remember it must be stated so it can be answered with a yes or no!)
>
> c. your question will then be written on an overhead transparency so you and others in the group can use it to develop other specific questions, and
>
> d. finally, your question will be answered with a yes or no.
>
> This procedure will continue until twenty questions have been answered or until twenty minutes have passed. So the questions can be later analyzed, this session will be audiotaped. To insure your anonymity, I will not call anyone by name. (The session now began.)

(After twenty questions or twenty minutes, the audiotape continued.) Now I will distribute paper and turn on the projector so you can view the film again. I would appreciate it if you would write as many hypotheses as you can that could explain the scientific problem. As in the past, you will have five minutes.

(After five minutes.) Your time is up! Please stop writing and pass your paper to the front. Your participation in the study is now completed. But your help is still needed since other members of your class are not finished. So to make the experiment a success, please do not discuss this session with your classmates until next week. Thank you.

Assessment of the Dependent Variables

Quantity of Hypotheses and Questions

The necessary measure for quantity was the number of nonrepeatable hypotheses and questions. In the event of an incomplete sentence, it was counted as part of the total quantity only if the meaning was understood so to be rated as to quality or diversity category.

Quality of Hypotheses

Since the scale designed by Quinn (1971) was validated and, further, proven useful in analyzing the hypotheses of both elementary students (Quinn, 1971) and secondary students (Wright, 1974), it was selected as the quality measure. By using this scale (Table 5, p. 72), each hypothesis of a participant was given the point value which corresponded to its category. These numbers were averaged to determine the quality of hypothesis generation for each individual. These averages were used in the later statistical analyses.

TABLE 5.--Hypothesis Quality Scale of Quinn (1971)

Points	Classification
0	No explanation: for example a nonsense statement, a question, an observation, a single inference about a single concrete object
1	Nonscientific explanation: for example, " . . . because it's magic" or "because the man pushed a button."
2	Partial scientific explanation: for example, incomplete reference to variables, a negative explanation, analogy
3	Scientific explanation relating at least two variables in general or nonspecific terms
4	Precise scientific explanation, a qualification and/or quantification of the variables
5	Explicit statement of a test for a hypothesis (An inference is made here that the child who states a test can also hypothesize adequately and precisely.)

Diversity of Questions

Suchman (1962b) devised a scale to determine the category of questions generated during an information search group discussion. Wright (1974) found it applicable for the analysis of the questions generated by ninth graders. Therefore, this scale (see Table 6, p. 73) was the preference for this study.

The diversity of information search question scale is arranged so a question can be classified into one of sixteen categories which are defined by eight parameters:

1. events--refer to the occurrence of events (e.g. Did he wipe the blade?),

2. objects--refer to the nature of objects (e.g. Was the liquid water?),

3. conditions--refer to the states of an object, in this context conditions can vary and are defined by numbers (e.g. Was the temperature of the water 85° F?),

4. properties--refer to properties, in this context properties do not vary and refer to constant characteristics (e.g. Does an ordinary knife bend when heated?),

5. verification--if the question seeks to identify or verify some aspect of the entire filmloop sequence,

6. experimentation--if the question seeks to ascertain the consequences of some hypothetical change in the experiment presented by the film,

7. necessity--if the question seeks to determine whether a particular aspect of a phenomenon in the film was necessary for the outcome (cause and effect), and

8. synthesis--if the question seeks to determine if a particular idea of theory of causation is valid and explains totally of some aspect of the experiment.

TABLE 6.--Diversity of information search question scale

	Events	Objects	Conditions	Properties
Verification	V_e	V_o	V_c	V_p
Experimentation	E_e	E_o	E_c	E_p
Necessity	N_e	N_o	N_c	N_p
Synthesis	S_e	S_o	S_c	S_p

The simplest way to analyze a question is to first determine the vertical and, then, the horizontal category. As a result, questions fall into one of the sixteen categories. A condition of high diversity is the presence of questions in many categories as compared with low diversity which contains an equal number of questions but only in one category. For this investigation, after all the questions were classified for each individual, mathematical manipulations made it possible to determine diversity.

The Shannon Index was utilized to calculate the diversity of questions because the various functions provided a concise way to express how the participants responses corresponded to maximum diversity. Originally an information scale (currently employed by ecologists), the Shannon Index measures the uncertainty of predicting the specific identity of specific questions when drawing individuals at random. Naturally, the higher diversity values indicate greater uncertainty. For calculations of this study, the following equation applied:

a. $\bar{H} = -\Sigma P_i \log P_i$: where the sum of P_i was the cummulative probability of having a question in each category of the scale,

b. $\bar{H}_{Bits/Sample} = C\,(N \log N - \Sigma n_i \log n_i)$: where N is the total number of questions; C is a conversion factor; and n_i is the number of questions in each category,

c. $\bar{H}_{Bits/Individuals} = \frac{C}{N}\,(N \log N - \Sigma n_i \log n_i)$,

d. $\bar{H}_{Max} = \text{Log}_{10}\, S\,(C)$: maximum diversity where S is the number of questions, and

e. $e = \frac{\bar{H}_{B/I}}{\bar{H}_{Max}}$: the evenness calculation compares the number in each category to the maximum diversity.

Only the evenness value was presented in the data. If the maximum diversity was attained by a student (e.g. nine questions with one in each category), then the evenness value was one. Values, therefore, ranged from zero (if no questions were asked) to one for maximum diversity.

Correlation to the Measurement Scales

To assure the lack of prejudice on the part of the investigator in rating each hypothesis or question, two junior high science teachers were asked to rate fifteen hypotheses and questions which were randomly selected from student responses during the study. Using Kendall's Coefficient of Concordance (Hayes, 1973, pp. 801-03), the results of the teachers and of the experimenter were compared. The coefficient values were 0.83 for the hypotheses and 0.94 for the questions. The selected questions and hypotheses as well as instructions are included in Appendix III (page 125). It is interesting to note that the good agreement on the Hypothesis Quality Scale indicates the applicability of the instrument. Quinn (1971, p. 45) found an interjudge reliability coefficient of 0.94 when she validated the scale.

Statistical Analyses

The six dependent variables for the written questions and hypotheses for each participant of the various experimental groups were analyzed by planned comparisons (contrasts) for the main effect interactions and treatment versus nontreatment comparisons. When a significant difference occurred, either the Newman-Keuls or Dunnett Test was conducted to determine the extent. The level of significance to support the hypothesis was 0.05. Figure 3 (page 78) depicts the sequence of statistical analyses for the preceeding dependent variables.

To determine the impact of the dependent variables that referred to diversity of questions during the group discussion, an analysis of variance and both the Newman-Keuls and Dunnett test were conducted. To further determine the effect of previous exposure to written hypotheses on subsequent hypothesis generation behaviors, the analysis of covariance on the planned comparisons was conducted. The covariables were quantity and quality of hypothesis generation immediately following instruction, and the variables were quantity and quality of hypothesis generation after the group discussion. Again, the level of significance to support the hypotheses was 0.05.

Summary

This study involved a comparison of five methods of hypothesis generation intensive instruction and also, the assessment of the effect of the instruction on group discussion behavior and the, subsequent, ability to generate hypotheses. Basically, students from the intact classes of two ninth grade science teachers were placed into each of the groups which received varying amounts of information. The independent variables were the presence or absence of reinforcement and knowledge of the criteria for an acceptable hypothesis. The seven dependent variables were:

a. both the quantity and quality of written hypotheses immediately following instruction,

b. both the quantity and diversity of written information search questions immediately following instruction,

c. the diversity of information search questions during a group discussion, and

d. both the quantity and quality of written hypotheses after the group discussion.

The discrepant event utilized to stimulate intensive instruction and the dependent variables were selected films from the Inquiry Development Program. The measure of the quality of hypotheses was the Hypothesis

Quality Scale, and the measure for diversity of question was the Information Search Scale. The statistical analyses involved planned comparison between experimental groups. Chapter IV presents the specific findings.

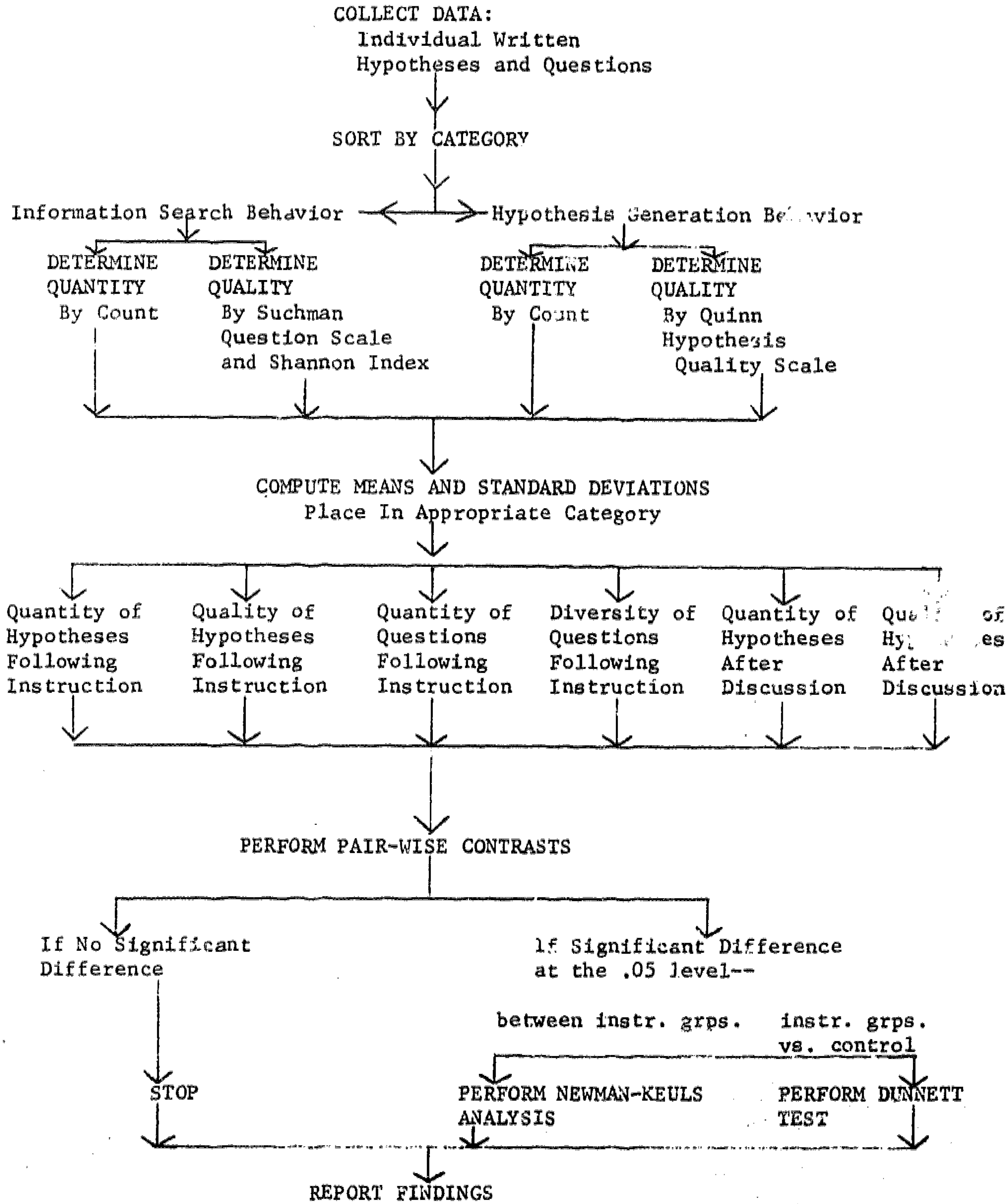


Fig. 3.--Statistical analyses for written hypotheses and questions

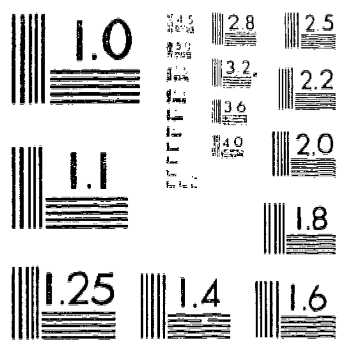
CHAPTER IV

FINDINGS OF THE STUDY

Techniques Utilized

The individual written hypotheses and questions were analyzed using respectively the Hypothesis Quality Scale (see Table 5, p. 72) and the Information Search Scale (see Table 6, p. 73). Further, the diversity of the questions was calculated by the Shannon Diversity Index (see p. 74). The derived values for all participants appear in Appendix IV (see p. 130).

There were five equally populated experimental groups and six dependent written response variables. Pair-wise comparisons were performed on each variable to ascertain if differences between groups existed because of the form of instruction rather than by chance alone. Specifically, contrasts between the presence or absence of reinforcement, criteria, and treatment interactions were conducted. For the analyses, the computer program Multivariate Analysis of Variance (Clyde, *et al.*, 1966, pp. 20-41) was employed because it could determine the specific comparisons as requested. If a significant difference occurred in a comparison of intensive instruction groups, then the Newman-Keuls analysis (Dayton, 1970, p. 147) was conducted (by the researcher) to determine the extent. However, if a significant difference occurred in the comparison of the intensive instruction groups and the control, then the Dunnett test (Dayton, 1970, p. 49) was utilized.



MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

The questions generated during the group discussion were transcribed and analyzed to obtain the diversity value for each group. These values were first submitted to an analysis of variance (Hays, 1973, pp. 457-519) followed by the Newman-Keuls and Dunnett analyses. (These three analyses were determined by the researcher.) To assess the effect of previous exposure to generating hypotheses prior to the group discussion on subsequent hypothesis generation behavior after the group discussion, an analysis of covariance was conducted by the Multivariate Analysis of Variance (Clyde, et al., 1966, pp. 20-41). The covariables were the quantity and quality of written hypotheses immediately following treatment and the variables were the quantity and quality of written hypotheses after the group discussion. The level of significance to support each hypothesis was 0.05.

The Specific Contrasts

The importance of planned pair-wise comparisons is that the univariate analysis is further delineated so the presence of significant differences becomes obvious. The four comparisons of this investigation involved:

1. the four intensive instruction groups versus the control,
2. the additive effect of criteria and reinforcement (interaction between treatment),
3. the reinforced groups versus the non-reinforced groups, and
4. the criteria groups versus the non-criteria groups.

The Research Hypotheses

The hypotheses for this study were listed in detail on pages 9-17. Since each hypothesis for written hypotheses and questions (hypotheses one through twenty-four) corresponds to the planned pair-wise comparisons, the most concise way to present the data is to first outline the six dependent variables in terms of specific findings. Secondly, the data which refer to the hypotheses about the group discussion and, subsequent, effect on hypothesis generation behaviors (hypotheses twenty-five through twenty-seven) will be presented.

Summary of Results--Hypothesis Generation Following Instruction Sessions

Dependent Variable One--
Quantity of Written Hypothesis
Generation Following Intensive Instruction

The four hypotheses which refer to the quantity of written hypotheses following intensive instruction are:

1. There is a difference in the quantity of written hypotheses between the following group which receives a form of instruction and the control group which receives no instruction:

   a. the differentiated reinforcement and criteria group versus the control group,

   b. the undifferentiated reinforcement and criteria group versus the control group,

   c. the differentiated reinforcement only group versus the control group, and

   d. the undifferentiated reinforcement only group versus the control group.

2. The effects of differentiated reinforcement and criteria are non-additive on the quantity of written hypotheses.

3. There is a difference in the quantity of written hypotheses between the following groups which receive differentiated reinforcement as an instruction condition and the following groups which receive undifferentiated reinforcement:

   a. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement and criteria group,

   b. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement only group,

c. the differentiated reinforcement only group versus the undifferentiated reinforcement and criteria group, and

d. the differentiated reinforcement only group versus the undifferentiated reinforcement only group.

4. There is a difference in the quantity of written hypotheses between the following groups which are told the criteria for good hypothesis formation and the following groups which are not told the criteria for good hypothesis formation:

   a. the differentiated reinforcement and criteria group versus the differentiated reinforcement only group,

   b. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement only group,

   c. the undifferentiated reinforcement and criteria group versus the differentiated reinforcement only group, and

   d. the undifferentiated reinforcement and criteria group versus the undifferentiated reinforcement only group.

Table 7 (p. 83) lists the mean value for each experimental group. Although the mean values differ, the specific pair-wise comparisons only indicate a significant difference for intensive instruction (treatment) versus control and differentiated reinforcement versus undifferentiated reinforcement (see Table 8, p. 83). The subsequent Newman-Keuls and Dunnett analyses indicate that significant differences only occur between the differentiated reinforcement groups and the control (see Table 9, p. 84). The implication is, therefore, that differentiated reinforcement (e.g. good) as a treatment condition results in higher quantity of hypotheses only in comparison with the control (and not compared to the other intensive instruction groups). Accordingly, only parts (a) and (c) of hypothesis one are supported while parts (b) and (d) of hypothesis one as well as hypotheses two, three and four are not supported.

TABLE 7.--Means and standard deviations for each intensive instruction group on the quantity of written hypotheses following intensive instruction

Mean Standard Deviation	Group				
	A*	B*	C*	D*	E*
M	4.069	3.390	4.000	3.439	2.463
SD	2.121	1.686	1.844	1.450	1.790

*Group A: Differentiated Reinforcement and Criteria
*Group B: Undifferentiated Reinforcement and Criteria
*Group C: Differentiated Reinforcement only
*Group D: Undifferentiated Reinforcement only
*Group E: Control

TABLE 8.--Pair-wise comparisons of the experimental groups for the quantity of hypothesis generation following intensive instruction

Source	Degrees Freedom	Sum of Squares	Mean Square	F	Probability (less than)
Treatment vs. Control	1	51.751	51.751	16.123	.001*
Interaction	1	.098	.098	.030	.862
Differentiated Reinforcement vs. Undifferentiated Reinf.	1	15.244	15.244	4.749	.030*
Criteria vs. no Criteria	1	.000	.000	.000	1.000
Error	200	641.951	3.21		
Total	204	709.004			

*Significant at 0.05

TABLE 9.--Pair-wise analyses of experimental group data for the quantity of written hypotheses following intensive instruction

		Newman-Keuls*				Dunnett**
Group	Mean	A 4.049	C 4.000	D 3.439	B 3.390	E 2.463
A[1]	4.049	-----	.1751	2.1801	2.3552	4.0080**
C[3]	4.000		-----	2.0049	2.1801	3.8842**
D[4]	3.439			------	.1751	2.4665
B[2]	3.390				-------	2.3425
E[5]	2.463					------

*Significant at 0.05 when t ≅ 4.04
**Significant at 0.05 when t ≅ 3.21

[1]Group A: Differentiated Reinforcement and Criteria
[2]Group B: Undifferentiated Reinforcement and Criteria
[3]Group C: Differentiated Reinforcement only
[4]Group D: Undifferentiated Reinforcement only
[5]Group E: Control

Dependent Variable Two--
Quality of Written Hypothesis
Generation Following Intensive Instruction

The four hypotheses which refer to this dependent variable are:

5. There is a difference in the quality of written hypotheses between the following groups which receive a form of intensive instruction and the control group which receives no instruction:

   a. the differentiated reinforcement and criteria group versus the control group,

   b. the undifferentiated reinforcement and criteria group versus the control group,

   c. the differentiated reinforcement only group versus the control group, and

d. the undifferentiated reinforcement only group versus the control group.

6. The effects of differentiated reinforcement and criteria are non-additive on the quality of written hypotheses.

7. There is a difference in the quality of written hypotheses between the following groups which receive differentiated reinforcement as an instruction condition and the following groups which receive undifferentiated reinforcement:

   a. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement and criteria group,

   b. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement only group,

   c. the differentiated reinforcement only group versus the undifferentiated reinforcement and criteria group, and

   d. the differentiated reinforcement only group versus the undifferentiated reinforcement only group.

8. There is a difference in the quality of written hypotheses between the following groups which are told the criteria for good hypothesis formation and the following groups which are not told the criteria for good hypothesis formation:

   a. the differentiated reinforcement and criteria group versus the differentiated reinforcement only group,

   b. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement only group,

   c. the undifferentiated reinforcement and criteria group versus the differentiated reinforcement only group, and

   d. the undifferentiated reinforcement and criteria group versus the undifferentiated reinforcement only group.

Table 10 (p. 87) lists the mean values for all the experimental groups. Pair-wise comparisons, illustrated in Table 11 (p. 87), indicate the differences to be significant for treatment (intensive instruction) versus the control, interaction of the treatments, and criteria versus no criteria. Further, the Newman-Keuls and Dunnett analyses (Table 12, p. 88) indicate that differentiated reinforcement, criteria or both are responsible for a significantly greater quality of hypotheses than treatment without their presence or no intensive instruction. Figure 4 (below) illustrates the condition implied by the data that when differentiated reinforcement is present the addition of criteria has no effect. Therefore, parts (a), (b), and (c) of hypothesis five, hypothesis six, and parts (b) and (d) of hypotheses seven and eight are supported.

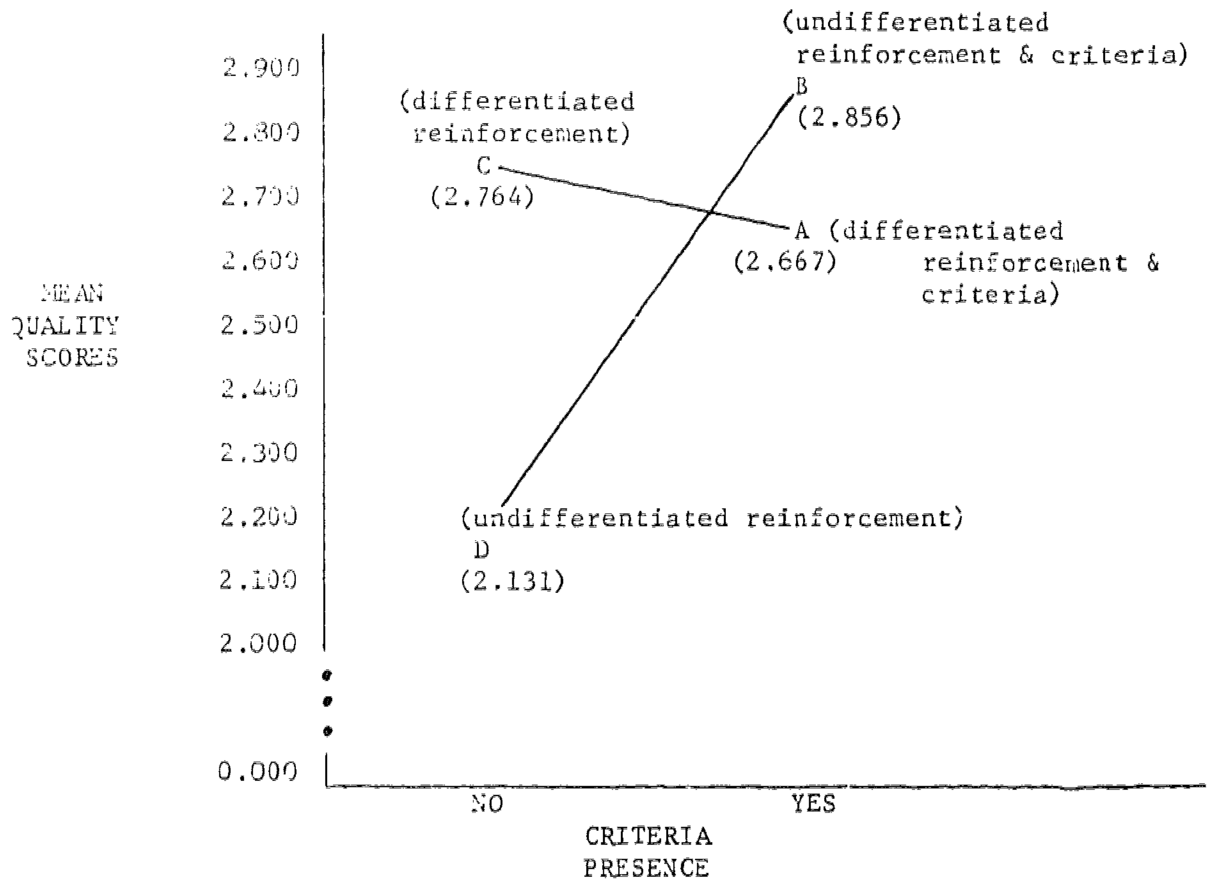


Fig. 4.--Interaction Graph of Treatments

TABLE 10.--Means and standard deviations for each intensive instruction group on the quality of written hypotheses following intensive instruction

Mean Standard Deviation	Group				
	A*	B*	C*	D*	E*
M	2.667	2.856	2.764	2.131	1.853
SD	.757	.640	.567	.669	1.028

*Group A: Differentiated Reinforcement and Criteria
*Group B: Undifferentiated Reinforcement and Criteria
*Group C: Differentiated Reinforcement only
*Group D: Undifferentiated Reinforcement only
*Group E: Control

TABLE 11.--Pair-wise comparisons of the experimental groups for the quality of hypothesis generation following intensive instruction

Source	Degrees of Freedom	Sum of Squares	Mean Square	F	Probability (less than)
Treatment vs. Control	1	18.519	18.519	32.984	.001*
Interaction	1	6.913	6.913	12.312	.001*
Differentiated Reinforcement vs. Undifferentiated Reinf.	1	2.018	2.018	3.593	.059
Criteria vs. no Criteria	1	4.037	4.037	7.190	.008*
Error	200	112.289	.562		
Total	204	143.775			

*Significant at 0.05

TABLE 12.--Pair-wise analyses of experimental group data for the quality of written hypotheses following intensive instruction

Group	Mean	Newman-Keuls*				Dunnett**
		B 2.856	C 2.764	A 2.667	D 2.131	E 1.853
B[2]	2.856	-----	.7858	1.6143	6.1924*	6.0631**
C[3]	2.764		------	.8285	5.4066*	5.5070**
A[1]	2.667			------	4.5781*	4.9206**
D[4]	2.131				------	1.6805
E[5]	1.853					

*Significant at .05 when $t \geqq 4.04$
**Significant at .05 when $t \geqq 3.21$

[1]Group A: Differentiated Reinforcement and Criteria
[2]Group B: Undifferentiated Reinforcement and Criteria
[3]Group C: Differentiated Reinforcement only
[4]Group D: Undifferentiated Reinforcement only
[5]Group E: Control

Summary of Results--Information Search Following Intensive Instruction

Dependent Variable Three-- Quantity of Written Information Search Questions Following Intensive Instruction

Table 13 (p. 90) shows the mean values for the quantity of written information search questions. The pair-wise comparisons indicate the contrast of the intensive instruction groups to the control as significant (see Table 14, p. 91). However, the Dunnett test, which accounts for error associated with multiple individual comparisons, indicates no significant differences as shown in Table 15 (p. 91). The implication is that the intensive instruction does not improve the students' ability to generate a greater quantity of questions. Accordingly, the following hypotheses are not supported.

9. There is a difference in the quantity of written information search questions between the following groups which receive a form of instruction and the control group which receives no instruction:

    a. the differentiated reinforcement and criteria group versus the control group,

    b. the undifferentiated reinforcement and criteria group versus the control group,

    c. the differentiated reinforcement only group versus the control group, and

    d. the undifferentiated reinforcement only group versus the control group.

10. The effects of differentiated reinforcement and criteria are non-additive on the quantity of written information search questions.

11. There is a difference in the quantity of written information search questions between the following groups which receive differentiated reinforcement as an instruction condition and the following groups which receive undifferentiated reinforcement:

    a. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement and criteria group,

    b. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement only group,

    c. the differentiated reinforcement only group versus the undifferentiated reinforcement and criteria group, and

    d. the differentiated reinforcement only group versus the undifferentiated reinforcement only group.

12. There is a difference in the quantity of written information search questions between the following groups which are told the criteria for good hypothesis formation and the following groups which are not told the criteria for good hypothesis formation:

    a. the differentiated reinforcement and criteria group versus the differentiated reinforcement only group,

b. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement only group,

c. the undifferentiated reinforcement and criteria group versus the differentiated reinforcement only group, and

d. the undifferentiated reinforcement and criteria group versus the undifferentiated reinforcement only group.

TABLE 13.--Means and standard deviations for each intensive instruction group on the quantity of written information search questions following intensive instruction

Mean Standard Deviation	Group				
	A*	B*	C*	D*	E*
M	4.878	5.195	5.220	5.098	3.902
SD	2.159	2.442	2.770	2.755	1.960

*Group A: Differentiated Reinforcement and Criteria
*Group B: Undifferentiated Reinforcement and Criteria
*Group C: Differentiated Reinforcement only
*Group D: Undifferentiated Reinforcement only
*Group E: Control

Dependent Variable Four--
Diversity of Written Information
Search Questions Following
Intensive Instruction

Table 16 (p. 93) illustrates the differences in mean values among the experimental groups. Further, the contrasts (see Table 17, p. 93) indicate significant differences between the treatment groups and the control. The Dunnett analysis, however, indicates no significant differences for the various pair-wise comparisons (see Table 18, p. 94). This finding implies, that although the intensive instruction groups have a greater mean diversity of questions than the control, the intensive

TABLE 14.--Pair-wise comparisons of the experimental groups for the quantity of information search questions following intensive instruction

Source	Degrees of Freedom	Sum of Squares	Mean Square	F	Probability (less than)
Treatment vs. Control	1	46.849	46.849	7.880	.005*
Interaction	1	1.976	1.976	.332	.565
Differentiated Reinforcement vs. Undifferentiated Reinf.	1	.390	.390	.066	.798
Criteria vs. no Criteria	1	.610	.610	.103	.749
Error	200	1189.073	5.945		
Total	204	1238.898			

*Significant at 0.05

TABLE 15.--Pair-wise analyses of experimental group data for the quantity of written information search questions following intensive instruction

Group	Mean	E(Dunnett**) 3.902
C[3]	5.220	2.4475
B[2]	5.195	2.4010
D[4]	5.098	2.2209
A[1]	4.878	1.8125
E[5]	3.902	------

**Significant at 0.05 when t ≥ 3.21

[1]Group A: Differentiated Reinforcement and Criteria
[2]Group B: Undifferentiated Reinforcement and Criteria
[3]Group C: Differentiated Reinforcement only
[4]Group D: Undifferentiated Reinforcement only
[5]Group E: Control

instruction does not significantly improve the diversity of questions.

Therefore the following hypotheses are not supported.

13. There is a difference in the diversity of written information search questions between the following groups which receive a form of instruction and the control group which receives no instruction:

a. the differentiated reinforcement and criteria group versus the control group,

b. the undifferentiated reinforcement and criteria group versus the control group,

c. the differentiated reinforcement only group versus the control group, and

d. the undifferentiated reinforcement only group versus the control group.

14. The effects of differentiated reinforcement and criteria are non-additive on the diversity of written information search questions.

15. There is a difference in the diversity of written information search questions between the following groups which receive differentiated reinforcement as an instruction condition and the following group which received undifferentiated reinforcement:

a. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement and criteria group,

b. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement only group,

c. the differentiated reinforcement only group versus the undifferentiated reinforcement and criteria group, and

d. the differentiated reinforcement only group versus the undifferentiated reinforcement only group.

16. There is a difference in the diversity of written information search questions between the following groups which are told the criteria for good hypothesis formation and the following groups which are not told the criteria for good hypothesis formation:

a. the differentiated reinforcement and criteria group versus the differentiated reinforcement only group,

b. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement only group,

c. the undifferentiated reinforcement and criteria group versus the differentiated reinforcement only group, and

d. the undifferentiated reinforcement and criteria group versus the undifferentiated reinforcement only group.

TABLE 16.--Means and standard deviations for each intensive instruction group on the diversity of written information search questions following intensive instruction

Mean Standard Deviation	Group				
	A*	B*	C*	D*	E*
M	.797	.770	.737	.713	.608
SD	.340	.330	.372	.388	.448

*Group A: Differentiated Reinforcement and Criteria
*Group B: Undifferentiated Reinforcement and Criteria
*Group C: Differentiated Reinforcement only
*Group D: Undifferentiated Reinforcement only
*Group E: Control

TABLE 17.--Pair-wise comparisons of the experimental groups for the diversity of information search questions following intensive instruction

Source	Degrees of Freedom	Sum of Squares	Mean Square	F	Probability (less than)
Treatment vs. Control	1	.703	.703	4.919	.028*
Interaction	1	.000	.000	.001	.981
Differentiated Reinforcement vs. Undifferentiated Reinf.	1	.026	.026	.180	.672
Criteria vs. no Criteria	1	.139	.139	.973	.325
Error	200	28.566	.143		
Total	204	29.434			

*Significant at 0.05

TABLE 18.--Pair-wise analyses of the experimental group data for the diversity of information search questions following intensive instruction

Group	Mean	E (Dunnett**) .6079
A[1]	.7966	.7145
B[2]	.7702	.6145
C[3]	.7370	.4888
D[4]	.7134	.3994
E[5]	.6079	-----

*Significant at .05 when t≥3.21

[1]Group A: Differentiated Reinforcement and Criteria
[2]Group B: Undifferentiated Reinforcement and Criteria
[3]Group C: Differentiated Reinforcement only
[4]Group D: Undifferentiated Reinforcement only
[5]Group E: Control

Summary of Results--Hypothesis Generation Following the Group Discussion

Dependent Variable Five-- Quantity of Written Hypothesis Generation After the Group Discussion

The four hypotheses which refer to this dependent variable are:

17. There is a difference in the quantity of written hypotheses between the following groups which receive a form of instruction and the control group which receives no instruction:

    a. the differentiated reinforcement and criteria group versus the control group,

    b. the undifferentiated reinforcement and criteria group versus the control group,

    c. the differentiated reinforcement only group versus the control group, and

d. the undifferentiated reinforcement only group versus the control group.

18. The effects of differentiated reinforcement and criteria are non-additive on the quantity of written hypotheses.

19. There is a difference in the quantity of written hypotheses between the following groups which receive differentiated reinforcement as an instruction condition and the following groups which receive undifferentiated reinforcement:

    a. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement and criteria group,

    b. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement only group,

    c. the differentiated reinforcement only group versus the undifferentiated reinforcement and criteria group, and

    d. the differentiated reinforcement only group versus the undifferentiated reinforcement only group.

20. There is a difference in the quantity of written hypotheses between the following groups which are told the criteria for good hypothesis formation and the following groups which are not told the criteria for good hypothesis formation:

    a. the differentiated reinforcement and criteria group versus the differentiated reinforcement only group,

    b. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement only group,

    c. the undifferentiated reinforcement and criteria group versus the differentiated reinforcement only group, and

    d. the undifferentiated reinforcement and criteria group versus the undifferentiated reinforcement only group.

Table 19 (p. 96) indicates the mean values for each experimental group. Further, Table 20 (p. 96) shows significant differences between the treatment groups versus the control and the interaction between treatment groups. The Newman-Keuls and Dunnett analyses (see Table 21, p. 97) indicate differentiated reinforcement only intensive instruction is responsible for significantly higher quantity of hypotheses than no intensive instruction.

Therefore, part (c) of hypothesis one is supported while the remaining portions of hypothesis seventeen and hypotheses eighteen, nineteen, and twenty are not supported.

TABLE 19.--Means and standard deviations for each experimental group on the quantity of written hypotheses after the group discussion

Mean Standard Deviation	Group				
	A*	B*	C*	D*	E*
M	2.293	2.805	3.293	2.341	1.902
SD	1.736	2.552	1.874	1.493	1.241

*Group A: Differentiated Reinforcement and Criteria
*Group B: Undifferentiated Reinforcement and Criteria
*Group C: Differentiated Reinforcement only
*Group D: Undifferentiated Reinforcement only
*Group E: Control

TABLE 20.--Pair-wise comparisons of the experimental groups for the quantity of written hypotheses after the group discussion

Source	Degrees of Freedom	Sum of Squares	Mean Square	F	Probability (less than)
Treatment vs. Control	1	19.980	19.980	5.944	.016*
Interaction	1	21.951	21.951	6.531	.011*
Differentiated Reinforcement vs. Undifferentiated Reinf.	1	1.976	1.976	.588	.444
Criteria vs. no Criteria	1	2.951	2.951	.878	.350
Error	200	672.244	3.361		
Total	204	719.102			

*Significant at 0.05

TABLE 21.--Pair-wise analyses of experimental group data for the quantity of written hypotheses after the group discussion

		Newman-Keuls*				Dunnett**
Group	Mean	C 3.293	B 2.805	D 2.341	A 2.293	E 1.902
C[3]	3.293	-----	1.7045	3.3250	3.4927	3.4353**
B[1]	2.805		-------	1.6206	1.7882	2.2301
D[4]	2.341			------	.1676	1.0842
A[1]					------	.9636
E[5]	1.902					-------

*Significant at .05 when $t \geq 4.04$
**Significant at .05 when $t \geq 3.21$

[1]Group A: Differentiated Reinforcement and Criteria
[2]Group B: Undifferentiated Reinforcement and Criteria
[3]Group C: Differentiated Reinforcement only
[4]Group D: Undifferentiated Reinforcement only
[5]Group E: Control

Dependent Variable Six--
Quality of Written Hypothesis
Generation After the Group Discussion

The four hypotheses for this dependent variable are:

21. There is a difference in the quality of written hypotheses between the following groups which receive a form of instruction and the control group which receives no instruction:

    a. the differentiated reinforcement and criteria group versus the control group,

    b. the undifferentiated reinforcement and criteria group vers s the control group,

    c. the differentiated reinforcement only group versus the control group, and

    d. the undifferentiated reinforcement only group versus the control group.

22. The effects of differentiated reinforcement and criteria are non-additive on the quality of written hypotheses.

23. There is a difference in the quality of written hypotheses between the following groups which receive differentiated reinforcement as an instruction condition and the following groups which receive undifferentiated reinforcement:

    a. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement and criteria group,

    b. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement only group,

    c. the differentiated reinforcement only group versus the undifferentiated reinforcement and criteria group, and

    d. the differentiated reinforcement only group versus the undifferentiated reinforcement only group.

24. There is a difference in the quality of written hypotheses between the following groups which are told the criteria for good hypothesis formation and the following groups which are not told the criteria for good hypothesis formation:

    a. the differentiated reinforcement and criteria group versus the differentiated reinforcement only group,

    b. the differentiated reinforcement and criteria group versus the undifferentiated reinforcement only group,

    c. the undifferentiated reinforcement and criteria group versus the differentiated reinforcement only group, and

    d. the undifferentiated reinforcement and criteria group versus the undifferentiated reinforcement only group.

Table 22 (p. 99) indicates the mean values of each group. Further, Table 23 (p.100) illustrates comparisons which show a significant difference between the treatment groups versus the control and the criteria groups versus no criteria. As a result of the Newman-Keuls and Dunnett analyses (see Table 24, p.100), significant differences occur between the groups receiving criteria as an intensive instruction condition and those receiving no intensive instruction or undifferentiated reinforcement only. This finding indicates that the intensive instruction method of criteria is better than no instruction or instruction without criteria and undifferentiated reinforcement. Accordingly, parts (a) and (b) of hypothesis twenty-one and parts (b) and (d) of hypothesis twenty-four are supported while the remaining parts of these hypotheses and hypotheses twenty-two and twenty-three are not supported.

TABLE 22.--Means and standard deviations for each experimental group for the quality of written hypotheses after the group discussion

Mean Standard Deviation	Group				
	A*	B*	C*	D*	E*
M	2.798	2.815	2.590	2.234	2.087
SD	.955	.758	.825	.984	.905

*Group A: Differentiated Reinforcement and Criteria
*Group B: Undifferentiated Reinforcement and Criteria
*Group C: Differentiated Reinforcement only
*Group D: Undifferentiated Reinforcement only
*Group E: Control

TABLE 23.--Pair-wise comparisons of the experimental groups for the quality of written hypothesis generation after the group discussion

Source	Degrees of Freedom	Sum of Squares	Mean Square	F	Probability (less than)
Treatment vs. Control	1	8.942	8.942	11.304	.001*
Interaction	1	1.429	1.429	1.807	.180
Differentiated Reinforcement vs. Undifferentiated Reinf.	1	1.183	1.183	1.496	.223
Criteria vs. no Criteria	1	6.381	6.381	8.066	.005*
Error	200	158.216	.791		
Total	204	176.152			

*Significant at 0.05

TABLE 24.--Pair-wise analyses of experimental group data for the quality of written hypotheses after the group discussion

		Newman-Keuls*				Dunnett**
Group	Mean	B 2.815	A 2.798	C 2.590	D 2.234	E 2.087
B[2]	2.815	-----	.1224	1.6199	4.1829*	3.7061**
A[1]	2.798		------	1.4975	4.0605*	3.6196**
C[3]	2.590			------	2.5630	2.5607
D[4]	2.234				------	.7484
E[5]						------

*Significant at 0.05 when $t \geq 4.04$
**Significant at 0.05 when $t \geq 3.21$

[1]Group A: Differentiated Reinforcement and Criteria
[2]Group B: Undifferentiated Reinforcement and Criteria
[3]Group C: Differentiated Reinforcement only
[4]Group D: Undifferentiated Reinforcement only
[5]Group E: Control

Summary of Results--Diversity of Information Search During the Group Discussion

These findings have been treated separately from the written data since they deal only with group results and not the sum of all individuals in each group. This reduction is reflected by a decrease in the total number of participating units--205 (200 df) to 30 (25 df). The hypothesis for this--the seventh dependent variable--is:

25. There is a difference in the diversity of verbal questions contributed by the various treatment groups during a group discussion about a discrepant event.

After the diversity index for the data was compiled for each group, the values were analyzed with the Analysis of Variance and the Newman-Keuls and Dunnett tests. The results, which appear in Tables 25 and 26 (p. 102), indicate no significant differences between any treatment groups or any treatment group and the control. This implies, in the presence of peer interaction, diversity of oral questions is not significantly improved by intensive instruction in hypothesis generation. Therefore, the above hypothesis is not supported.

TABLE 25.--Analysis of variance for diversity of information search questions during the group discussion

Source	Degrees of Freedom	Sum of Squares	Mean Square	F*
Treatments (between groups)	4	0.0251	0.0063	1.9688
Error (within groups)	25	0.0606	0.0032	
Total		0.0857		

*For significance at 0.05, F must be greater than 2.76.

TABLE 26.--Pair-wise analysis of data for diversity of information search during the group discussion

			Newman-Keuls*				Dunnett**
Group	Standard Deviation	Mean	D .8755	A .8736	C .8564	B .8452	E .7802
D[4]	0.0370	.8755	-----	.0823	.8271	1.3120	2.9180
A[1]	0.0684	.8736		-----	.7448	1.2298	2.8598
C[3]	0.0649	.8565			------	.4893	2.3362
B[2]	0.0435	.8452				------	1.9902
E[5]	0.0547	.7802					------

*Significant at 0.05 when t ≥ 4.04
**Significant at 0.05 when t ≥ 3.21

[1]Group A: Differentiated Reinforcement and Criteria
[2]Group B: Undifferentiated Reinforcement and Criteria
[3]Group C: Differentiated Reinforcement only
[4]Group D: Undifferentiated Reinforcement only
[5]Group E: Control

Summary of Results--Performance for Hypothesis Generation After Group Discussion Using Previous Exposure as a Covariable

An interesting aspect of the interpretation of the data was to determine if students from the various groups improved their ability to generate hypotheses (both quality and quantity) because of previous exposure to the process of generating hypotheses. Accordingly, the best measure was to examine the significant differences of the pair-wise comparisons using previous data as a covariable. The two hypotheses for the analysis are:

26. using the results of the quantity of written hypotheses immediately following treatment as a covariable for the quantity of written hypotheses after the group discussion, there is a difference between the following planned comparisons:

    a. treatment vs. control,
    b. interaction between treatments,
    c. differentiated reinforcement vs. undifferentiated reinforcement, and
    d. criteria vs. no criteria, and

27. using the results of the quality of written hypotheses immediately following treatment as a covariable for the quality of written hypotheses after the group discussion, there is a difference between the following planned comparisons:

    a. treatment vs. control,
    b. interaction between treatments,
    c. differentiated reinforcement vs. undifferentiated reinforcement, and
    d. criteria vs. no criteria.

Tables 27 and 28 (p. 104) indicate significant differences for the quantity of written hypotheses with the interaction comparison and with the quality of written hypotheses with the criteria vs. no criteria comparison. However, the Newman-Keuls results for quantity of hypotheses after group discussion (Table 21, p. 97), indicates no further significance. Therefore, hypothesis

TABLE 27.--Pair-wise comparisons of experimental groups for the quantity of written hypothesis generation after the group discussion using previous quantity of written hypotheses as a covariate

Source	Degrees of Freedom	Sum of Squares	Mean Square	F	Probability (less than)
Treatment vs. Control	1	8.102	8.102	2.561	.111
Interaction	1	24.800	24.800	7.838	.006*
Differentiated Reinforcement vs. Undifferentiated Reinf.	1	.301	.301	.095	.752
Criteria vs. no Criteria	1	1.994	1.994	.630	.428
Error	198				
Total	202				

*Significant at 0.05

TABLE 28.--Pair-wise comparisons of experimental groups for the quality of written hypothesis generation after the group discussion using previous quality of written hypotheses as a covariate

Source	Degrees of Freedom	Sum of Squares	Mean Square	F	Probability (less than)
Treatment vs. Control	1	2.899	2.899	3.761	.054
Interaction	1	.515	.515	.668	.415
Differentiated Reinforcement vs. Undifferentiated Reinf.	1	.407	.407	.528	.468
Criteria vs. no Criteria	1	4.529	4.529	5.875	.016*
Error	198				
Total	202				

*Significant at 0.05

twenty-six is not supported as no differences occur between any of the comparisons. The quality of hypotheses after group discussion Newman-Keuls data (see Table 24, p. 100), indicates significant differences do occur between treatment groups. Therefore, hypothesis twenty-seven is supported for the comparison involving criteria vs. no criteria and rejected for the other comparisons.

Summary of the Findings

The findings support many of the hypotheses and, thus, indicate differences between the methods of hypothesis generation intensive instruction utilized in this study. For each of the dependent variables, the findings involve the following.

a. Dependent Variable One--Differentiated reinforcement as an intensive instruction method is responsible for a higher quantity of written hypothesis after intensive instruction than the instruction method which involves no intensive instruction.

b. Dependent Variable Two--Participants who received intensive instruction which emphasizes either differentiated reinforcement, criteria or both generate a higher quality of written hypotheses, following intensive instruction, than participants who receive only undifferentiated reinforcement or no intensive instruction.

c. Dependent Variables Three and Four--No form of hypothesis generation intensive instruction improves the participants' ability to generate a greater quantity or diversity of written information search questions following intensive instruction.

d. Dependent Variable Five--Differentiated reinforcement alone as an intensive instruction condition is responsible for a greater quantity of written hypotheses than no intensive instruction following the group discussion.

e. Dependent Variable Six--Criteria as an intensive instruction method is responsible for a higher quality of written hypotheses, after the group discussion, than the instruction method of undifferentiated reinforcement or no intensive instruction.

f. Dependent Variable Seven--In the presence of the group discussion, diversity of oral information search questions is not significantly improved by hypothesis generation intensive instruction.

g. Participants who received intensive instruction which utilized criteria were not effected by prior exposure to written hypothesis generation when they generated a higher quality of written hypotheses--than participants who received only undifferentiated reinforcement or no intensive instruction--after the group discussion.

CHAPTER V

CONCLUSIONS, IMPLICATIONS, AND RECOMMENDATIONS

> If ever there was a cause, if ever there can be a cause, worthy to be upheld by all the toil or sacrifice that the human heart can endure, it is the cause of education.
>
> Horace Mann, 1796-1859

Summary

The formation of hypotheses is such an important objective of science education that curricula designed over the last fifteen years have emphasized specific learning activities which encourage students to isolate the variables inherent within a problem and, then, formulate reasonable hypotheses which may account for the principles of causality. In essence, students are urged to learn the content of a specific scientific discipline while developing and utilizing problem solving skills. Previous research has illustrated that problem solving skills can be enhanced by exposing students to specific strategies designed to encourage isolating and relating variables by asking questions or generating hypotheses about a discrepant event (Suchman, 1962b; Quinn, 1971; Wright, 1974). However, none of the previous work included comparisons of various hypothesis generation intensive instruction methods or assessments of the value of prior instruction during peer group discussion. Therefore, the purpose of this investigation was to:

a. compare the effects of various hypothesis generation intensive instruction procedures on the ability of ninth grade students to generate written hypotheses and information search questions about a discrepant event, and

b. determine if students, who have received hypothesis generation intensive instruction, in a setting free of peer interactions, exhibit a greater diversity of questions during a group discussion and greater written hypothesis generation behaviors after the discussion.

Procedures

The intact science classes of two ninth grade science teachers in a suburban junior high school comprised the participants for this study. Since 205 students were involved in the final analyses, forty-one students were present in each of four instruction groups and one control group. While the assignment to each group was random, there was stratification--according to the sex composition of each intact class--for the sex of the individuals. The procedures for the four instruction groups included:

a. listening to general instructions about hypothesis formation and watching an introductory discrepant event,

b. watching the intensive instruction discrepant event until six acceptable hypotheses were written, and

c. individual discussions during which the investigator evaluated each of the six hypotheses by one of the following predetermined standards:

   1. differentiated reinforcement and criteria group--each of the hypotheses had to meet a predetermined level of acceptability and, after each was stated, the student was both positively reinforced (e.g. good) and told the criteria for good hypothesis formation,

   2. undifferentiated reinforcement and criteria group--each of the hypotheses had to meet a predetermined level of acceptability and, after each was stated, the student was only told the criteria for good hypothesis formation,

3. differentiated reinforcement only group--each of the hypotheses had to meet a predetermined level of acceptability and, after each was stated, the student was only positively reinforced (e.g. good), or

4. undifferentiated reinforcement only group--had to generate six hypotheses and, after each was stated, the student received only acceptance (e.g. I can accept this hypothesis).

The control group only listened to general instructions about hypothesis formation and watched the introductory discrepant event. All the discrepant events were selected from the Inquiry Development Program Films (Suchman, 1962b).

Upon completion of the instruction sessions for all the experimental groups, the participants from both science classes during each period were shown another Inquiry Development Program Film (discrepant event) and were requested to write as many hypotheses as possible. Then, another film was shown and the participants were requested to write as many questions as possible which would provide information to help explain the discrepancy. Five days later, group discussions began using one experimental group during each class period until all the experimental groups had completed the discussion (five days). During the discussion, the students observed another Inquiry Development Program Film and, then had the opportunity, to voluntarily, ask questions to the investigator about the discrepant event. The questions were specific so they could be answered with yes or no. After twenty questions, the discussion was terminated and the students were requested to individually write hypotheses that might explain the discrepancy.

Thus, this study gathered data about seven dependent variables:

a. both the quantity and quality of written hypotheses following intensive instruction,

b. both the quantity and diversity of written information search questions following intensive instruction,

c. the diversity of information search questions during the group discussion, and

d. both the quantity and quality of written hypotheses after the group discussion.

The dependent variables for written results were analyzed utilizing planned comparisons which were further delineated by the Newman-Keuls or Dunnett analyses. Finally, the analysis of covariance was conducted to determine the effect of previous exposure to written hypothesis generation on hypothesis generating ability after the group discussion. The level of significance which was needed to support the hypothesis was 0.05.

Conclusions

The findings of this study lead to the formation of several conclusions. It should be emphasized, however, that each conclusion is restricted by the procedures employed (pp. 54-[illegible]) and the limitations of this study (p. 22).

1. The methods of hypothesis generation intensive instruction which employed differentiated reinforcement, whenever a hypothesis of a predetermined standard was generated, were more effective than no hypothesis generation intensive instruction (control group) in promoting a greater quantity of written hypotheses about a discrepant event following instruction. Further, while differentiated reinforcement intensive instruction was better than no instruction, there were no significant differences in the quantity of written hypotheses between the four intensive instruction groups. Therefore, differentiated reinforcement (either alone or with criteria) is only superior to no instruction and equal to other forms of intensive instruction for the quantity of written hypotheses following intensive instruction.

2. Those participants who received differentiated reinforcement intensive instruction (alone or in combination with criteria) or undifferentiated reinforcement and criteria intensive instruction produced a significantly higher quality of written hypotheses following instruction than participants who received either (a) only undifferentiated reinforcement or (b) no intensive instruction. In addition, there was no significant difference between either of the differentiated reinforcement intensive instruction groups and the undifferentiated reinforcement and criteria group. It seems reasonable to conclude that the presence of differentiated reinforcement and the addition of criteria does not enhance the quality of written hypothesis generation following intensive instruction. Although the presence of differentiated reinforcement, criteria or both cause a significant improvement in the ability to generate a higher quality of hypotheses following intensive instruction, differentiated reinforcement alone is sufficient to cause the same result.

3. None of the four methods of hypothesis generation intensive instruction improved the ability of the participants to generate a significantly higher quantity or diversity of written information search questions about a discrepant event following intensive instruction (when compared to no intensive instruction).

4. Differentiated reinforcement only hypothesis intensive instruction was found more effective than no intensive instruction (control group) in promoting a higher quantity of written hypotheses after the information search group discussion. There were no significant differences among the four intensive instruction groups. Also, no significant differences were detected between the other three intensive instruction groups and the

control group. Therefore, intensive instruction which gives only differentiated reinforcement to participants is significantly better than no intensive instruction but not significantly different to the other forms of intensive instruction after the group discussion.

5. Participants who received the criteria (either with differentiated or undifferentiated reinforcement), as the condition of hypothesis generation intensive instruction, produced a higher quality of written hypotheses after the group discussion than participants who either (a) received only undifferentiated reinforcement, or (b) received no intensive instruction. There was no difference in the quality of hypotheses when each criteria group was compared to the differentiated reinforcement only group. So giving criteria about good hypothesis formation was only more effective than no intensive instruction or only undifferentiated reinforcement for the quality of written hypotheses following the group discussion.

6. None of the four forms of hypothesis generation intensive instruction significantly effected the diversity of questions submitted during a group discussion when compared to no intensive instruction (control group).

7. The analysis of covariance indicated that those participants who received criteria as a form of intensive instruction were not effected by prior exposure to written hypothesis generation when this group of participants generated a higher quality of written hypotheses than participants who received no intensive instruction. However, the criteria only participants generated hypotheses of equal quality as the participants from the other intensive instruction groups.

The above conclusions are important because they indicate that a variety of hypothesis generation intensive instruction methods can be used to increase the quantity and/or quality of hypothesis generation. Basically, intensive instruction which emphasizes differentiated reinforcement is significantly better to no intensive instruction in promoting a higher quantity of hypotheses following intensive instruction. Likewise, for the quality of hypotheses following intensive instruction, either undifferentiated reinforcement and criteria or differentiated reinforcement and criteria, or differentiated reinforcement alone are the most effective forms of instruction when compared to no intensive instruction and only undifferentiated reinforcement intensive instruction. Since there is no difference between the differentiated reinforcement and criteria groups, reinforcement alone as an intensive instruction method is sufficient to produce a higher quality of written hypotheses. This conclusion substantiates the findings of Wright (1974) who used only positive reinforcement (e.g. good) as the intensive instruction condition in his work with ninth grade students. In summary, the methods of intensive instruction which emphasize differentiated reinforcement (either alone or with criteria) are most effective for written hypothesis generation following intensive instruction.

For written hypothesis generation after a group discussion, differentiated reinforcement alone was only significantly different (better) than no intensive instruction for the quantity of hypotheses. Further, a form of criteria (with differentiated or undifferentiated reinforcement) intensive instruction was the superior method for a higher quality of written hypotheses after the group discussion. It appears, therefore, that the effect of differentiated reinforcement on the quality of hypothesis generation is short-term. The previous statement is consistent with studies in concept attainment which found that students best attain and utilize a concept by instruction methods which allow the student to discover the concept with practice in application (Gagné and Brown, 1961; Wittrock, 1963).

In addition, it is interesting to note that this study--unlike previous work--utilized two forms of reinforcement. The differentiated form employed three different terms (good, very good, and excellent) which corresponded to each of three levels of acceptable quality hypothesis generation during intensive instruction. The undifferentiated form, on the other hand, merely utilized acceptance (in the absence of positive terms) for contributed hypotheses which either did or did not reach a predetermined level of acceptability depending upon the intensive instruction group. When each form of reinforcement is examined in terms of the quantity and quality of written hypotheses, worthwhile observations become apparent.

While no statistical differences occurred between the differentiated reinforcement only group and the undifferentiated reinforcement only group for the quantity of written hypotheses, significant differences did occur in favor of the differentiated reinforcement only group when compared to the control group (both after intensive instruction and the group discussion) These findings are consistent with the work of Byers (1965) who found positive reinforcement (in terms of selecting correct alternatives) increases the frequency of hypothesis formation. Further, for quality of written hypotheses, differentiated reinforcement was the important factor for significantly better hypotheses following intensive instruction. Although this finding did not continue for long-term hypothesis generation, it is crucial for short-term improvement in hypothesis generating behaviors. What the preceeding discussion indicates is that undifferentiated reinforcement may be useful as a precursor of differentiated reinforcement in the shaping of quantity and, then, quality hypothesis generation. Future study should investigate the specific effects of reinforcement (e.g. words used, frequency) on the entire process of hypothesis generation.

The significance of the conclusions is that each illustrates that hypothesis generation behaviors of ninth grade students can be improved by specific intensive instruction methods. The study indicates that students possess the cognitive ability to formulate acceptable hypotheses and intensive instruction further enhances these abilities.

Implications

The implications suggested by the conclusions are important because each offers ideas of specific value for classroom instruction.

1. If it is desirable to promote an immediate higher quantity of hypotheses about a discrepant event, then students should be given verbal differentiated reinforcement as part of the intensive instruction. Differentiated reinforcement is logical since it encourages students to continue the generation of hypotheses. Giving students only criteria or only practice at writing hypotheses is not as effective since students are probably hesitant to generate hypotheses which do not meet acceptable criteria.

2. Since at least one form of intensive instruction--reinforcement, criteria, or both--is necessary to produce a higher quality of hypotheses following intensive instruction, the implication is that students can determine what constitutes an acceptable hypothesis either by being told the criteria or by gaining confidence through differentiated reinforcement to figure out the criteria. Therefore, teaching strategies should involve a predetermined level of acceptability and one of the above intensive instruction methods. It should be noted that the differentiated reinforcement intensive instruction methods improve both the quantity and quality of written hypothesis generation following intensive instruction.

3. Since both the quantity and diversity of written questions were not significantly effected by hypothesis generation intensive instruction, the instruction did not cause students to change their approach to isolating variables by questioning but merely by hypothesizing. These results are at variance with Wright (1974) who intensively instructed ninth grade students by positive reinforcement (e.g. good) and by asking if they had further hypotheses or details about the discrepant event. Further, when Wright assessed the impact of the intensive instruction he used oral rather than written measures. In addition, Salomon (1970), working with college students and written measures of information search questions, found a significant difference due to intensive instruction. The findings of Wright and Salomon, when considered with the findings of this study, indicate further research must be conducted to determine the effect of hypothesis generation intensive instruction on information search questioning behaviors.

4. Even after a group discussion, previous differentiated reinforcement,

as an instructional condition, yielded a higher quantity of hypotheses. This implies that differentiated reinforcement gives students confidence to generate hypotheses. It is important to note that, among the intensive instruction groups, there were no differences, but the reinforcement only group was significantly better than no instruction. This implies that, unless students receive reinforcement, the quantity of hypothesis generation is not increased.

5. The implication is strong that criteria as an intensive instruction condition is important to increase the quality of hypotheses after a group discussion since students could recall the appropriate criteria. The interesting fact is that the group which received both the differentiated reinforcement and criteria was significantly better for quality but not quantity after the group discussion. This tends to indicate that criteria instruction is effective due to a clear understanding of what is expected since it does not compete with the desire to generate a greater number of hypotheses (regardless of quality) for the purpose of receiving praise (differentiated reinforcement).

6. The conclusion that instruction does not enhance diversity of group information search questioning implies again that students did not modify their approach to isolating variables by questioning as a result of intensive instruction. Further, in the group setting individuals with extremely complex or evaluative questions may have been hesitant to volunteer them due to adverse peer pressure. Nevertheless, the differences in hypothesis generation behaviors indicate students did improve their ability to isolate and explain the relationship between variables. Unfortunately, this improvement did not manifest itself by either the individual written

or the group discussion information search activities. Since the present study only employed twenty questions--as the extent of the group discussion--future research should attempt to determine the effect of an increasing number of questions (e.g. twenty-five, forty, seventy-five) on the diversity of questions contributed during the group discussion.

7. Since the effect of previous exposure to hypothesis generation is not apparent with the participants who received criteria only for the quality of written hypotheses, the implication exists that when students know the criteria they can utilize additional information. It could be these participants were not confused in their approach to solving problems.

8. Further, the successful implementation of the instructional procedures in a secondary school setting implies that most teachers could devise similar methods to teach hypothesis formation through intensive instruction models.

In summary, the implications of the present study are worthwhile for educators to consider in planning activities to promote the goals of inquiry instruction. In the short-term, intensive instruction which emphasizes only differentiated reinforcement as a condition of instruction is superior to other forms. However, for the long-term, instruction which emphasizes criteria as the method of hypothesis generation intensive instruction is preferable.

Recommendations

As a result of this investigation, several recommendations for further study are important to mention.

1. This investigation should be repeated to see if the same results are achieved by junior high students.

2. This investigation should be repeated to determine if the same results are achieved by senior high students.

3. The investigation should be repeated with individual oral responses collected from each individual instead of written responses.

4. The investigation should be repeated using the same instructional procedures but the group discussion should be varied in the following ways:

a. allow the students to ask as many questions as they desire and not the twenty question limit imposed by this study,

b. combine the various intensive instruction groups during the discussion, and

c. allow the entire class (all experimental groups) to simultaneously participate.

5. Since development of formal reasoning skills depends on several factors as age and correlated to factors as grades (Sayre and Ball, 1975), it would be interesting to assess the correlation between the various intensive instruction methods which are best to utilize for specific student traits (e.g. grade point average).

6. The form of stimulus for the discrepant events should be varied. Suggested formats include: (a) live demonstration, (b) videotape recordings, and (c) discrepant events where students present the discrepancy either on film or in a live setting.

7. The correlation between diversity of information search questions and hypothesizing behaviors needs to be determined.

8. A scale should be developed to better equate information search questioning diversity with hypothesis generation.

9. The effect of increasing the number of acceptable hypotheses required during intensive instruction should be assessed.

10. The effects of specific reinforcement techniques should be examined to include the following:

a. the effect of specific words (e.g. good, very good, excellent) as differentiated reinforcement,

b. the effect of specific phrases as undifferentiated reinforcement (e.g. I accept this hypothesis, This hypothesis is acceptable), and

c. the effects of a mixture of reinforcement (e.g. vary the words or phrases after each hypothesis).

APPENDIX I

PARENTAL PERMISSION LETTER

138

PRINCE GEORGE'S COUNTY PUBLIC SCHOOLS

GEORGE W. ANDERSON, JR.
PRINCIPAL

KARL TASCHENBERGER
VICE PRINCIPAL

Frederick Sasscer Junior High School

UPPER MARLBORO, MD. 20870

627-4551

Dear Parent,

Because of my interest in developing teaching methods which will enhance science instruction, I am on educational leave from Our Public School System to pursue my doctorate in Science Education at the University of Maryland. The purpose of my research study is to determine ways which will help students generate better hypotheses.

Basically, this study will not interfere with normal class instruction. The procedures will be part of regular class activities for a few days. These will involve:

a. the random placing of students into one of five groups,
b. the presentation of a lesson to each group in a setting removed from the classroom,
c. the participation of all the groups in a class discussion when the students' written and verbal responses will be recorded, and
d. the analysis of the students' responses.

The instruments that will be utilized are the Hypothesis Quality Scale and the Diversity of Information Search Scale. Both of these will be available for your review at the school. Further, the instruments do not reflect the science ability or potential of your child but merely the success or failure of the study. Your child will remain completely anonymous during and after the study. TO INSURE THIS, NO CHILD'S NAME WILL BE INCLUDED ON DATA GATHERING SHEETS OR TAPES!

In accordance with existing Board of Education Policy, Mr. Anderson and I are informing and requesting permission for your child to participate. If you have any questions, please call me at 864-8951. We would appreciate the return of this form within a week.

Thank you for your cooperation.

Sincerely,

Chris A. Pouler
Ph.D. Candidate, University of Maryland

George W. Anderson, Jr.
Principal

☐ I do give permission for my child to participate in the above study.

____
Name of Student

☐ I do not give permission for my child to participate in the above study.

____
Signature of Parent

APPENDIX II

DISCREPANT EVENT FILM QUESTIONNAIRE

NAME: ____________________

DATE: ____________________

INSTRUCTIONS: Please check those items which would apply for each film if presented to a ninth grade audience in a Prince George's County school.

The Restaurant	The Knife	The Ice Cubes	The Sailboat and the Fan	Drinking Boiling Coffee	The Cannon	TECHNOLOGY:
						1. The sequence is easy to observe.
						2. The sequence is presented in an interesting manner.
						3. There are no major distractions which prevent students from observing the sequence.

The Restaurant	The Knife	The Ice Cubes	The Sailboat and the Fan	Drinking Boiling Coffee	The Cannon	CONTENT:
						1. The discrepancy depicted by the sequence is understandable.
						2. The discrepancy is probably a new experience for the student.
						3. It is possible to state several hypotheses which attempt to explain the discrepancy.
						4. It is possible to ask several questions which may clarify the discrepancy so hypotheses can be generated.

APPENDIX III

HYPOTHESIS AND QUESTION EVALUATION INFORMATION

INSTRUCTIONS: After viewing "The Sailboat and the Fan," please rate each of these student hypotheses by using the number that "best" fits each.

NUMBER:

_____ 1. The boat didn't move because maybe the air wasn't in a certain angle.

_____ 2. The boat didn't move because there was not enough air to push the boat and the fan both.

_____ 3. The boat didn't move because maybe there were two speeds on the fan and the first time he stuck the speed on high and the second it was on low.

_____ 4. The tank of water may have been tilted the first time and not the second.

_____ 5. The man may have docked the boat to something so it would not move the second time.

_____ 6. Maybe the second time the fan only circulated the air around.

_____ 7. It could have been magic.

_____ 8. It could have been because of the thickness of the material of the sails.

_____ 9. The water could have been moving the first time and stayed still the second.

_____ 10. The wind from the fan was blowing at the sail and not getting caught into the sail to make it go.

_____ 11. The ship moved the first time since the fan was moving along with the boat.

_____ 12. The boat was not in the same position.

_____ 13. A magnet could have been on the bottom.

_____ 14. When the fan is lifted the boat is light and can move freely.

_____ 15. The first time the wind was not hitting the pole on the sail and the second time it was.

Hypothesis Quality Scale

The following was devised by Quinn (1972) and utilized by Quinn (1972) and Wright (1974) in hypothesis generation studies. Hypotheses are rated by the experimenter using the number that corresponds to the description that best "fits" the generated hypothesis.

Points	Classification
0	No explanation: for example a nonsense statement, a question, an observation, a single inference about a single concrete object
1	Nonscientific explanation: for example, ". . . because it's magic" or "because the man pushed the button."
2	Partial scientific explanation: for example, incomplete reference to variables, a negative explanation, analogy
3	Scientific explanation relating at least two variables in general or nonspecific terms
4	Precise scientific explanation, a qualification and/or quantification of the variables
5	Explicit statement of a test of a hypothesis (An inference is made here that the child who states a test can also hypothesize adequately and precisely.)

INSTRUCTIONS: After viewing "The Ice Cubes," please rate each of these student hypotheses by using the category that "best" fits each.

CATEGORY:

______ 1. Were the two cubes ice?

______ 2. Did each glass have the same amount of liquid in it?

______ 3. Did he put something in one of the cubes that would make it float?

______ 4. Did he use the same temperature liquid in each glass?

______ 5. Was the table level?

______ 6. Did the man put a solution in the glass when he pushed the ice cube down?

______ 7. Does the heat of the man's hand have any effect on the cubes?

______ 8. Were the liquids greater than room temperature when the cubes were added?

______ 9. Was there cold air blowing over one of the glasses?

______ 10. Were the cubes the same weight?

______ 11. Did the ice cubes have the same density?

______ 12. Are both cubes fully frozen?

______ 13. Could one glass have salt water in it?

______ 14. Was the spoon made of plastic?

______ 15. Did the man have something on one of his hands rubbing off on one of the cubes?

DIVERSITY OF INFORMATION SEARCH SCALE

The following scale was devised by Suchman to determine the diversity of questions during an information search activity. Wright (1974) found it applicable for the analysis of questions generated by ninth graders. Basically, a question is classified into one of the sixteen categories. After all the questions have been classified, mathematical manipulations make it possible to determine the diversity.

	Events	Objects	Conditions	Properties
Verification	V_e	V_o	V_c	V_p
Experimentation	E_e	E_o	E_c	E_p
Necessity	N_e	N_o	N_c	N_p
Synthesis	S_e	S_o	S_c	S_p

A question is classified into one of sixteen categories which are defined as:

a. Events--refer to the occurance of events (e.g. Did he wipe the blade?)

b. Objects--refer to the nature of objects (e.g. Was the liquid water?)

c. Conditions--refer to the states of an object, in this context conditions can vary and are defined by numbers (e.g. Was the temperature of the water 85° F?)

d. Properties--refer to properties, in this context properties do not vary and refer to constant characteristics (e.g. Does an ordinary knife bend when heated?)

e. Verification--if the question seeks to identify or verify some aspect of the entire filmloop sequence,

f. Experimentation--if the question seeks to ascertain the consequences of some hypothetical change in the experiment presented in the film,

g. Necessity--if the question seeks to determine whether a particular aspect of a phenomenon in the film was necessary for the outcome (cause and effect), and

h. Synthesis--if the question seeks to determine if a particular idea or theory of causation is valid and explains totally or some aspect of the experiment.

APPENDIX IV

INDIVIDUAL VALUES FOR THE SIX WRITTEN DEPENDENT VARIABLES

TABLE 29.--Written hypothesis and question values for each individual participating in the study

Group	Student	HQU[1] FII	HQA[2] FII	ISQ[3] FII	ISD[4] FII	HQU[5] AGD	HQA[6] AGD
A	1	00	0.00	02	1.0000	02	2.50
A	2	02	2.00	02	1.0000	03	2.00
A	3	03	2.33	05	0.9603	04	2.75
A	4	03	3.00	03	0.9183	02	3.00
A	5	03	2.33	05	0.9709	03	2.33
A	6	03	3.00	06	0.0000	01	1.00
A	7	03	2.33	02	1.0000	03	2.80
A	8	06	3.17	03	1.0000	03	3.00
A	9	06	3.33	04	0.8113	02	3.00
A	10	03	2.60	02	1.0000	01	3.00
A	11	05	2.00	03	1.0000	04	2.75
A	12	11	2.40	05	0.9709	04	1.50
A	13	04	2.30	02	0.0000	01	2.67
A	14	02	2.00	02	0.0000	04	2.75
A	15	02	1.80	05	0.9603	05	2.20
A	16	05	3.67	08	0.9464	01	4.00
A	17	03	3.26	03	1.0000	03	3.33
A	18	04	3.00	10	0.9709	02	3.00
A	19	07	3.00	03	0.0000	00	0.00
A	20	05	3.00	07	0.8631	01	4.00
A	21	02	2.83	09	0.8538	02	4.00
A	22	06	3.50	05	0.9603	01	3.00
A	23	02	3.80	07	0.9852	01	4.00
A	24	05	3.00	06	0.9183	07	2.86
A	25	06	2.00	06	0.9206	03	2.33
A	26	04	2.00	04	1.0000	00	0.00
A	27	02	2.50	07	0.9212	01	3.00
A	28	04	2.14	06	0.7897	01	3.00
A	29	07	3.25	08	0.9528	01	4.00
A	30	08	3.00	04	0.9464	01	3.00
A	31	06	2.33	04	0.9464	01	2.00
A	32	03	3.25	07	0.8195	05	3.00
A	33	04	4.00	04	1.0000	05	3.60
A	34	01	2.20	05	0.9610	01	3.00
A	35	05	2.60	05	0.8650	01	4.00
A	36	05	3.00	03	1.0000	01	4.00
A	37	02	3.00	04	0.8113	02	3.00
A	38	04	3.60	05	0.8650	07	2.86
A	39	05	2.50	03	0.0000	02	2.50
A	40	04	3.33	09	0.7725	01	4.00
A	41	01	1.00	07	0.0000	01	2.00

TABLE 29.--Continued

Group	Student	HQU[1] FII	HQA[2] FII	ISQ[3] FII	ISD[4] FII	HQU[5] AGD	HQA[6] AGD
B	1	02	2.00	06	0.9206	03	2.50
B	2	04	3.00	05	0.9603	03	2.33
B	3	01	2.00	04	0.0000	01	3.00
B	4	05	1.80	09	0.8289	10	2.00
B	5	05	2.40	09	0.8289	01	3.00
B	6	01	1.00	00	0.0000	01	3.00
B	7	03	2.67	04	0.8113	01	3.00
B	8	03	2.00	03	0.9133	03	3.33
B	9	05	2.80	05	0.7219	10	2.10
B	10	04	2.50	08	0.9851	11	2.00
B	11	03	2.67	03	0.9183	04	2.75
B	12	01	2.00	02	0.0000	02	2.50
B	13	03	3.33	01	0.0000	02	3.00
B	14	02	3.50	04	0.9464	02	3.50
B	15	04	2.75	06	0.9206	03	2.33
B	16	05	2.80	10	0.8174	01	4.00
B	17	02	3.00	06	0.9206	00	0.00
B	18	04	2.75	06	0.9206	01	4.00
B	19	04	3.75	03	0.9183	07	3.00
B	20	04	2.75	02	1.0000	02	2.50
B	21	01	3.00	06	0.7897	02	2.50
B	22	02	2.50	03	1.0000	01	3.00
B	23	04	2.25	05	0.9603	02	2.00
B	24	03	3.00	09	0.8289	02	3.00
B	25	05	2.40	04	0.9464	02	2.50
B	26	04	3.25	07	0.9212	02	3.00
B	27	09	2.67	10	0.9603	06	2.33
B	28	05	3.60	06	0.9591	04	3.00
B	29	05	3.40	07	0.9141	04	2.50
B	30	05	2.80	04	1.0000	04	2.25
B	31	06	2.67	06	0.9591	02	3.50
B	32	03	2.67	04	0.9464	01	4.00
B	33	02	3.50	03	0.9183	02	3.50
B	34	02	3.50	09	0.8764	01	4.00
B	35	02	4.00	06	0.9206	01	4.00
B	36	01	4.00	03	0.9183	01	4.00
B	37	02	3.50	05	0.7219	03	3.00
B	38	04	3.50	04	0.8113	02	3.00
B	39	02	3.50	04	0.0000	02	1.50
B	40	05	3.40	08	0.8869	02	3.00
B	41	02	2.50	04	0.0000	01	3.00

TABLE 29.--Continued

Group	Student	HQU[1] FII	HQA[2] FII	ISQ[3] FII	ISD[4] FII	HQU[5] AGD	HQA[6] AGD
C	1	05	2.00	06	1.0000	04	2.20
C	2	02	2.00	02	0.0000	03	2.40
C	3	05	2.20	00	0.0000	05	2.20
C	4	07	2.80	10	0.8174	07	2.70
C	5	05	2.60	07	0.9754	04	1.75
C	6	06	3.00	10	0.9373	03	3.00
C	7	04	3.00	06	1.0000	01	4.00
C	8	03	2.00	04	0.9709	01	4.00
C	9	03	2.67	02	0.0000	01	3.00
C	10	02	2.40	04	0.8764	01	2.00
C	11	01	2.00	02	0.0000	03	2.33
C	12	02	2.50	04	0.0000	04	2.00
C	13	03	1.67	04	0.8113	02	2.00
C	14	03	3.33	06	0.9206	03	2.33
C	15	05	2.80	06	0.9206	05	2.40
C	16	03	3.00	03	0.9183	03	2.67
C	17	04	2.25	03	0.9183	00	0.00
C	18	02	3.00	04	0.8113	04	2.75
C	19	03	3.67	02	0.0000	02	3.00
C	20	03	3.67	03	0.9183	03	3.33
C	21	01	3.00	03	0.9183	01	4.00
C	22	02	2.00	05	0.9603	01	3.00
C	23	05	3.40	07	0.8699	05	2.80
C	24	09	2.11	13	0.7560	07	2.14
C	25	04	3.00	08	0.9528	04	2.50
C	26	04	2.25	05	0.9603	03	2.67
C	27	05	2.60	07	0.9212	04	2.00
C	28	04	2.50	05	0.9610	05	2.60
C	29	03	3.33	07	0.9610	03	2.70
C	30	06	2.80	06	0.7897	06	2.50
C	31	06	3.00	06	0.9206	02	3.00
C	32	07	3.14	10	0.8567	03	3.33
C	33	06	3.50	07	0.8699	01	2.00
C	34	02	3.50	04	0.9464	00	0.00
C	35	04	3.00	05	0.9709	05	2.60
C	36	03	3.33	02	0.0000	04	3.33
C	37	05	2.00	08	0.8869	06	3.20
C	38	05	3.00	05	0.9603	06	3.33
C	39	07	2.29	05	0.9709	05	3.20
C	40	04	3.00	00	0.0000	01	2.00
C	41	01	4.00	08	0.9851	04	3.25

TABLE 29.--Continued

Group	Student	HQU[1] FII	HQA[2] FII	ISQ[3] FII	ISD[4] FII	HQU[5] AGD	HQA[6] AGD
D	1	04	2.25	03	0.0000	02	2.50
D	2	01	2.00	04	0.9464	06	2.17
D	3	05	2.40	05	0.9603	03	2.33
D	4	04	2.50	03	0.9183	03	2.00
D	5	05	1.20	07	0.9212	03	2.00
D	6	04	2.25	06	0.9206	03	2.67
D	7	01	2.00	04	0.0000	02	1.50
D	8	03	2.67	01	0.0000	00	0.00
D	9	05	2.40	03	0.9183	03	2.00
D	10	02	2.50	02	1.0000	03	2.67
D	11	05	1.60	06	1.0000	06	2.67
D	12	05	1.60	07	0.8699	04	2.25
D	13	01	1.00	00	0.0000	01	2.00
D	14	04	2.25	06	0.9206	00	0.00
D	15	04	2.75	09	1.0000	03	3.00
D	16	04	2.50	06	0.9183	03	2.67
D	17	03	1.33	08	0.8869	03	2.00
D	18	06	2.67	08	0.9464	04	2.25
D	19	04	1.75	10	0.9477	01	4.00
D	20	02	3.00	03	0.9183	01	4.00
D	21	02	2.50	00	0.0000	01	1.00
D	22	05	3.00	06	1.0000	05	2.00
D	23	04	1.25	06	0.9183	02	2.00
D	24	04	2.25	04	1.0000	03	2.00
D	25	03	2.67	06	0.9206	01	3.00
D	26	03	3.00	04	0.8113	03	2.00
D	27	02	1.50	03	0.0000	02	2.00
D	28	04	2.25	11	0.7828	01	2.00
D	29	04	3.00	05	0.8650	03	3.33
D	30	03	3.00	03	0.9183	01	3.00
D	31	05	2.00	10	0.9373	02	1.50
D	32	05	2.00	10	0.9603	02	2.50
D	33	03	2.67	05	0.8650	04	2.25
D	34	06	1.00	06	0.7897	01	4.00
D	35	03	1.67	07	0.9751	01	3.00
D	36	03	1.67	05	0.7219	03	3.00
D	37	02	2.00	01	0.0000	00	0.00
D	38	04	3.00	06	0.7897	01	2.33
D	39	03	2.33	06	1.0000	03	2.33
D	40	01	2.00	02	0.0000	00	0.00
D	41	00	0.00	02	0.0000	03	3.00

TABLE 29.--Continued

Group	Student	HQU[1] FII	HQA[2] FII	ISQ[3] FII	ISD[4] FII	HQU[5] AGD	HQA[6] AGD
E	1	01	1.00	02	1.0000	02	2.00
E	2	02	1.50	03	1.0000	04	1.50
E	3	00	0.00	06	0.0000	06	0.17
E	4	05	2.20	05	0.7219	00	0.00
E	5	02	2.50	02	1.0000	01	2.00
E	6	02	1.50	02	0.0000	01	3.00
E	7	01	2.00	06	1.0000	01	2.00
E	8	03	2.67	03	0.9183	02	2.50
E	9	03	2.25	05	0.9709	02	1.50
E	10	03	1.00	04	0.8113	03	2.00
E	11	05	2.20	02	0.0000	01	3.00
E	12	00	0.00	02	0.0000	01	3.00
E	13	00	0.00	02	0.0000	01	2.00
E	14	03	2.33	05	0.9603	01	3.00
E	15	01	0.00	00	0.0000	01	2.00
E	16	04	3.00	03	0.0000	04	3.00
E	17	01	2.00	03	0.9183	01	0.00
E	18	05	2.25	07	0.8631	03	2.33
E	19	01	1.00	04	0.0000	02	2.00
E	20	01	3.00	02	0.0000	03	2.00
E	21	01	2.00	06	0.9206	02	3.00
E	22	03	2.33	05	0.9610	01	3.00
E	23	03	2.33	06	0.9206	03	2.00
E	24	09	1.33	04	0.9464	01	2.00
E	25	06	1.83	07	0.9852	01	3.00
E	26	03	2.50	09	0.9657	03	1.67
E	27	02	1.00	04	0.8113	02	2.00
E	28	01	4.00	03	0.9183	03	2.33
E	29	01	3.00	04	0.0000	03	2.00
E	30	04	0.75	02	0.0000	02	2.50
E	31	03	2.33	06	0.7897	01	3.00
E	32	02	2.50	02	0.0000	00	0.00
E	33	04	1.50	05	0.7219	01	3.00
E	34	02	3.50	05	0.9709	01	2.00
E	35	03	3.00	05	0.9709	04	2.25
E	36	01	1.00	06	1.0000	01	0.00
E	37	02	2.50	03	1.0000	02	3.00
E	38	01	0.00	02	0.0000	03	2.33
E	39	03	0.67	00	0.0000	02	2.50
E	40	02	3.00	05	0.9603	01	3.00
E	41	02	2.50	03	0.9183	01	2.00

[1] Hypothesis quantity following intensive instruction
[2] Hypothesis quality following intensive instruction
[3] Question quantity following intensive instruction
[4] Question diversity following intensive instruction
[5] Hypothesis quantity after group discussion
[6] Hypothesis quality after group discussion

BIBLIOGRAPHY

Ankney, Paul and Sayre, Steve A. "Starting Points for Creativity," The Science Teacher, 42 (December, 1975), 23-5.

Atkin, J. Myron. "A Study of Formulating and Suggesting Tests for Hypotheses in Elementary School Science Learning Experiences." Science Education, 42 (December, 1958), 414-22.

Babikian, Yeghia. "An Empirical Investigation to Determine the Relative Effectiveness of Discovery, Laboratory and Expository Methods Teaching Concepts." Journal of Research in Science Teaching, 8 (Number 3, 1971), 201-09.

Barker, Robert Halcomb. "Development and Evaluation of a Method of Structuring Inquiry Films to Elicit Hypotheses From High School Chemistry Students." Unpublished Doctoral Dissertation, The University of Texas at Austin, 1969.

Bills, Frank Lynn. "Developing Creativity Through Inquiry." Science Education, 55 (July-September, 1971), 417-21.

Brown, Dorothy McKenna. "The Relationship of Independent Study, Object Visualization, and Anxiety to Hypothesis Formation by College Freshman in the Biological Sciences." Unpublished Doctoral Dissertation, The University of Pennsylvania, 1973.

Brown, Talbert W. "The Influence of the Science Curriculum Improvement Study on Affective Process Development and Creative Thinking." The University of Oklahoma, 1973.

Bruner, Jerome S. "The Process of Education Revisited." Phi Delta Kappan, 53 (September, 1971), 18-21.

Burkman, Ernest. "Unhooking High School Science: The ISIS Project." The Science Teacher, 41 (October, 1974), 30-2

Byers, Joe L. "Hypothesis Behavior in Concept Attainment." Journal of Educational Psychology, 56 (December, 1965), 337-42.

Chamberlain, T. C. "The Method of Multiple Working Hypotheses." Science, 148 (May, 1965), 754-59.

Clyde, Dean; Cramer, Elliot M.; and Sherin, Richard G. Multivariate Statistical Programs. Biometric Laboratory: University of Miami, 1966.

Coleman, Charles H. "Demonstration of the Use of Alternative Hypotheses." School Science and Mathematics, 66 (January, 1966), 81-5.

Dayton, C. Mitchell. The Design of Educational Experiments. New York: McGraw-Hill, 1970.

Dayton, D. Mitchell, and Stunkard, Clayton L. Statistics for Problem Solving. New York: McGraw-Hill, 1971.

DeTure, Linda R., and Koran, John J., Jr. "Using Filmed Models to Prompt Childrens' Laboratory Activities." Journal of Research in Science Teaching, 12 (October, 1975), 449-52.

Elkind, David. "Piaget and Science Education." Science and Children, 9 (November, 1972), 9-13.

Frederiksen, Norman and Evans, Franklin R. "Effects of Models on Creative Performance and Ability to Formulate Hypotheses." Journal of Educational Psychology, 66 (February, 1974), 67-82.

Friedman, Charles P. "A Project to Individualize PSSC." The Science Teacher, 43 (April, 1976), 15-8.

Gagnē, Robert M. "The Learning Requirements for Enquiry." Journal of Research in Science Teaching, 1 (1963), 144-53.

Gagnē, Robert M., and Brown, Larry T. "Some Factors in the Programming of Conceptual Learning." Journal of Experimental Psychology, 62 (October, 1961), 313-21.

Gibbs, Ronald K. "An Analysis of the Effectiveness of Biological Science Curriculum Study Single Topic Films in Teaching Hypothesis Construction to High School Biology Students." Unpublished Doctoral Dissertation, Indiana University, 1967.

Glasser, William. Schools Without Failure. New York: Harper and Row, 1969.

Guilford, J. P. "Factors That Aid and Hinder Creativity." Creativity: Its Educational Implications. Compiled by John Curtis Gowan, George D. Demos and E. Paul Torrance. New York: John Wiley and Sons, Inc., 1967.

Guilford, J. P. "Potential for Creativity." Gifted Child Quarterly, 6 (1962), 87-90.

Graybill, Lettia. "Sex Differences in Problem-Solving Abilities." Journal of Research in Science Teaching, 12 (October, 1975), 341-46.

Hays, William L. Statistics for the Social Sciences. New York: Holt, Rinehart and Winston, 1973.

Jerrigan, H. Dean. "A Research Approach to Inquiry in High School Biology." The American Biology Teacher, 34 (November, 1972), 466-70.

Joyce, Bruce and Weil, Marsha. Models of Teaching. Englewood Cliffs, N. J.: Prentice Hall, 1972.

Kersh, Bert Y. "The Adequacy of 'Meaning' as an Explanation for the Superiority of Learning by Independent Discovery." Journal of Educational Psychology, 49 (November, 1958), 282-92.

_______. "The Motivating Effect of Learning by Directed Discovery." Journal of Educational Psychology, 53 (Number 2, 1962), 65-71.

Kilburn, Robert E. "Hypothesis Testing: A Student Activity Approach." The Science Teacher, 30 (November, 1963), 51-2.

Klein, Stephen P.; Frederiksen, Norman; and Evans, Franklin R. "Anxiety and Learning to Formulate Hypotheses." Journal of Educational Psychology, 60 (December, 1969), 465-75.

Lawson, Anton E., and Renner, John W. "Relationships of Science Subject Matter and Developmental Levels of Learners." Journal of Research in Science Teaching, 12 (October, 1975), 347-58.

Lloyd, Monte; Zar, Jerrold H.; and Karr, James R. "On Calculation of Information--Theoretical Measures of Diversity," The American Midland Naturalist, 79 (April, 1968), 257-72.

Mallon, Elizabeth J. "Cognitive Development and Processes: Review of Philosophy of Jean Piaget." The American Biology Teacher, 38 (January, 1976), 28-33.

McGinty, Robert LeRoy. "The Effects of Four Method of Instruction Upon the Ability of Second and Third Grade Students to Derive Valid Logical Conclusions From Verbally Expressed Hypotheses." Unpublished Doctoral Dissertation, Michigan State University, 1972.

Micciche, Frank and Keany, Michael. "Hypothesis Machine." The Science Teacher, 36 (April, 1969), 53-6.

Pugliese, Anthony C. "The Meaning of Inquiry, Discovery, and Investigative Approaches to Science Teaching." The Science Teacher, 40 (January, 1973), 24-7.

Pulaski, Mary Ann. Understanding Piaget. New York: Harper and Row, 1971.

Quinn, Mary Ellen. "Evaluation of a Method for Teaching Hypothesis Formation to Sixth Graders." Unpublished Doctoral Dissertation, The University of Pennsylvania, 1971.

______. "Hypothesis Formation Can Be Taught." The Science Teacher, 39 (September, 1972), 30-1.

Rapp, Marvin A. "The Brainstorming Attitude." Creativity: Its Educational Implications. Compiled by John Curtis Gowan, Goerge D. Demos and E. Paul Torrance. New York: John Wiley and Sons, 1967.

Renner, John W., and Stafford, Don G. Teaching Science in the Secondary School. New York: Harper and Row, 1972.

Salomon, Gavriel. "Cue Attendance and Hypothesis Generation as Two Procedures of Training for Producing Uncertainty in Teachers," U. S. Office of Education Technical Report 14: Stanford University. ED045622, 1970.

Sayre, Steve and Ball, Daniel W. "Piagetian Cognitive Development and Achievement in Science." Journal of Research in Science Teaching, 12 (April, 1975), 165-75.

Sechrest, Lee and Wallace, John. "Assimilation and Utilization of Information in Concept Attainment Under Varying Conditions of Information Presented." Journal of Educational Psychology, 53 (August, 1962), 157-64.

Sprafka, Sarah Anne. "The Effect of Hypothesis Generation and Verbalization on Certain Aspects of Medical Problem Solving." Unpublished Doctoral Dissertation, Michigan State University, 1973.

Suchman, J. Richard. "Creative Thinking and Conceptual Growth." Gifted Child Quarterly, 6 (Autumn, 1962a), 95-9.

______. Developing Inquiry: Inquiry Development Program in Physical Science. Chicago: Science Research Associates, Inc., 1966.

______. "The Elementary School Training Program in Scientific Inquiry," U. S. Office of Education Title VII Project Number 216: University of Illinois, 1962b.

______. "Inquiry Training Building Skills for Autonomous Discovery." Merrill-Palmer Quarterly, 7 (1961), 147-69.

______. "Learning Through Inquiry." Childhood Education, (February, 1965), 289-91.

______. "Rebuilding the Science Program Inquiry Training in the Elementary School." The Science Teacher, 27 (November, 1960), 42-7.

Trowbridge, Leslie W. "Trends and Innovations in Junior High School Science Teaching in the United States." The Science Teacher, 41 (April, 1974), 12-5.

Tyler, Ralph W. "Measuring the Ability to Infer." Educational Research Bulletin 9 (November 19, 1930), 475-80.

Wittrock, M. C. "Verbal Stimuli in Concept Formation: Learning by Discovery." Journal of Educational Psychology, 54 (August, 1963), 183-90.

Woodburn, John H. "Hypotheses: Crossroads to Success Versus Failure in Science and Science Teaching." School Science and Mathematics, 69 (April, 1969), 333-37.

Wright, Emmett L. "The Effect of Intensive Instruction in Cue Attendance and Hypothesis Generation Upon Open Exploration Behavior of Low and High Ability Ninth Grade Students." Unpublished Doctoral Dissertation, The Pennsylvania State University, 1974.

______. "Fifteen Months Later: The Effects of Intensive Instruction Upon the performance of Basic Inquiry Skills." The University of Maryland, 1976. (Mimeographed.)