WALLDREN, Allan Wade, 1934-
THE DEVELOPMENT OF AN INSTRUMENT TO ANALYZE
STUDENT QUESTIONS DURING PROBLEM SOLVING.
The Ohio State University, Ph.D., 1971
Education, general

University Microfilms, A XEROX Company, Ann Arbor, Michigan

© Copyright by
Allan Wade Walldren
1971
THE DEVELOPMENT OF AN INSTRUMENT TO ANALYZE
STUDENT QUESTIONS DURING PROBLEM SOLVING

Dissertation
Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By
Allan Wade Walldren, B.S., M.A.T.

The Ohio State University
1971

Approved by

[Signature]
Adviser
Faculty of Curriculum and Foundations
PLEASE NOTE:

Some Pages have indistinct print. Filmed as received.

UNIVERSITY MICROFILMS
ACKNOWLEDGEMENTS

As author of this study I wish to publicly acknowledge and thank several people for their assistance, guidance and support in this investigation. First and foremost is Louise, my wife, who shared equally in all of the rigors and frustrations of this investigation. She has been sounding board and colleague, critical editor and typist, and most of all an understanding and sensitive friend.

I also wish to acknowledge Dr. Charles M. Galloway, who not only supported the study and assisted in its design, but also encouraged me to bring my past experiences into a new focus and to report them in a style that was comfortable.

Dr. Paul R. Klohr not only supported the study but sensitively encouraged the investigator at the most appropriate times. Constructive criticism was always blended with a twinkle of humor.

Both the study and the investigator were significantly strengthened by Dr. John C. Belland who unravelled some of the data and finely focused the scope of the investigation.

Finally I'd like to acknowledge Dr. Roger T. Cunningham who inspired the study, and Mr. Donald R. Teffner who field tested the instrument in all of its phases of development.
VITA

Allan Wade Walldren  
Born, January 29, 1934; Chicago, Ill.

Education

The Ohio State University (1970-71) doctoral program in curriculum and instruction
Indiana University (Summer 1963) science
Michigan State University (1961-62) MAT in science education
Utah State University (Summer 1961) science and math
University of Illinois (1959-60) education
Lawrence College (1953-57) BS in history; minors: science and education
Marquette University (1952-53) general undergraduate program

Experience

1970-71 The Ohio State University. Teaching Associate for Introduction to Education (108) and General Methods for Secondary Teachers (435) courses.

1966-69 Science Research Associates, Chicago. Project Director for Teacher Education Division; Editor for Science Department; Manager of Marketing Division, Science Products.

1964-66 American Appraisal Co., Milwaukee. Training Director and Field Supervisor for Institutional Division; Representative for Industrial Division.

1962-64 Muskego-Norway School District, Wisconsin. Teacher of 9th, 10th, 11th grade biology and physical science.


Publications


# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>VITA</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td>I. THE PROBLEM</td>
<td>1</td>
</tr>
<tr>
<td>Need for the Study</td>
<td></td>
</tr>
<tr>
<td>Design of the Study</td>
<td></td>
</tr>
<tr>
<td>II. A REVIEW OF THE RELATED LITERATURE</td>
<td>11</td>
</tr>
<tr>
<td>Introduction</td>
<td></td>
</tr>
<tr>
<td>Studies of Classroom Observational Systems</td>
<td></td>
</tr>
<tr>
<td>Studies of the Question in Classroom Interaction</td>
<td></td>
</tr>
<tr>
<td>Efforts to Improve Classroom Questioning Practices</td>
<td></td>
</tr>
<tr>
<td>Summary</td>
<td></td>
</tr>
<tr>
<td>III. PROCEDURES USED IN DEVELOPING THE INSTRUMENT AND IN TRAINING ANALYSTS TO USE IT</td>
<td>42</td>
</tr>
<tr>
<td>Background</td>
<td></td>
</tr>
<tr>
<td>The Search for the Instrument</td>
<td></td>
</tr>
<tr>
<td>Suchman's Analysis of Inquiry</td>
<td></td>
</tr>
<tr>
<td>Carner's Levels of Questioning</td>
<td></td>
</tr>
<tr>
<td>Dodl's System</td>
<td></td>
</tr>
<tr>
<td>Bloom's Taxonomy</td>
<td></td>
</tr>
<tr>
<td>Sanders' Modification</td>
<td></td>
</tr>
<tr>
<td>&quot;QSOS&quot;</td>
<td></td>
</tr>
<tr>
<td>Other Systems</td>
<td></td>
</tr>
<tr>
<td>The Gallagher-Aschner System</td>
<td></td>
</tr>
<tr>
<td>The Development of the Instrument</td>
<td></td>
</tr>
<tr>
<td>The Training Program</td>
<td></td>
</tr>
<tr>
<td>The Source of Data</td>
<td></td>
</tr>
<tr>
<td>IV. ANALYSIS OF THE DATA</td>
<td>74</td>
</tr>
<tr>
<td>The Ice Cubes</td>
<td></td>
</tr>
<tr>
<td>Joe</td>
<td></td>
</tr>
<tr>
<td>Summary</td>
<td></td>
</tr>
</tbody>
</table>
V. SUMMARY, FINDINGS, INFERENCES, IMPLICATIONS AND RECOMMENDATIONS ............... 92

Summary
Findings
Inferences
Implications
Recommendations

Appendix

A. JOE: A CASE STUDY ................. 103

B. BEHAVIORAL SCIENCE PROBLEM SOLVING SESSION WITH SELECTED EDUCATION 435 STUDENTS: CONDUCTED BY ALLAN WALLDREN--"Joe: A Case Study" .... 104

C. PHYSICAL SCIENCE PROBLEM SOLVING SESSION WITH SELECTED EDUCATION 435 STUDENTS: CONDUCTED BY ALLAN WALLDREN--"The Ice Cubes" ....... 107

D. ANALYSTS TRAINED IN THE STUDY .......... 110

E. INQUIRY SESSION WITH J. RICHARD SUCHMAN AND 6TH GRADERS, "THE KNIFE," Film 15 ........ 111

F. DEMONSTRATION WORKSHOP: INQUIRY DEVELOPMENT PROGRAM, "DRINKING BOILING COFFEE," Film 16 114

G. DEMONSTRATION WORKSHOP: INQUIRY DEVELOPMENT PROGRAM, "THE PULSE GLASS" ............. 116

H. INQUIRY SESSION WITH 7TH GRADERS, NORTON MIDDLE SCHOOL, "Joe: A Case Study" .... 118

I. INQUIRY SESSION CONDUCTED BY DONALD TEFFNER WITH 7TH GRADERS, NORTON MIDDLE SCHOOL, "Joe: A Case Study" ............. 126

J. INQUIRY SESSION WITH 7TH GRADERS, NORTON MIDDLE SCHOOL, "Joe: A Case Study" .... 129

K. INQUIRY SESSION WITH 5TH GRADERS, "Drop the Dollar" ............. 133

L. INQUIRY SESSION WITH 9TH GRADERS, "The Rosenthal Effect" ............. 135

REFERENCES ............. 138
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Item-by-item Analysis of the &quot;Ice Cubes&quot; Typescript</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Composite Analysis of Coding by Category (The &quot;Ice Cubes&quot;)</td>
<td>76</td>
</tr>
<tr>
<td>2.</td>
<td>Ratios of Interanalyst Agreement: All Categories (The &quot;Ice Cubes&quot;)</td>
<td>77</td>
</tr>
<tr>
<td>3.</td>
<td>Ratios of Interanalyst Agreement: Verifying Category</td>
<td>78</td>
</tr>
<tr>
<td>4.</td>
<td>Ratios of Interanalyst Agreement: Seeking Category</td>
<td>79</td>
</tr>
<tr>
<td>5.</td>
<td>Ratios of Interanalyst Agreement: Inferring Category</td>
<td>79</td>
</tr>
<tr>
<td>6.</td>
<td>Ratios of Interanalyst Agreement: Concluding Category</td>
<td>80</td>
</tr>
<tr>
<td>7.</td>
<td>Ratios of Interanalyst Agreement: Foil Category</td>
<td>81</td>
</tr>
<tr>
<td>8.</td>
<td>Item-by-item Analysis of the &quot;Joe&quot; Typescript</td>
<td>81</td>
</tr>
<tr>
<td>9.</td>
<td>Composite Analysis of Coding by Category (&quot;Joe&quot;)</td>
<td>83</td>
</tr>
<tr>
<td>10.</td>
<td>Ratios of Interanalyst Agreement: All Categories (&quot;Joe&quot;)</td>
<td>84</td>
</tr>
<tr>
<td>11.</td>
<td>Ratios of Interanalyst Agreement: Verifying Category</td>
<td>85</td>
</tr>
<tr>
<td>12.</td>
<td>Ratios of Interanalyst Agreement: Seeking Category</td>
<td>86</td>
</tr>
<tr>
<td>13.</td>
<td>Ratios of Interanalyst Agreement: Inferring Category</td>
<td>86</td>
</tr>
<tr>
<td>14.</td>
<td>Ratios of Interanalyst Agreement: Concluding Category</td>
<td>87</td>
</tr>
<tr>
<td>15.</td>
<td>Ratios of Interanalyst Agreement: Foil Category</td>
<td>89</td>
</tr>
<tr>
<td>16.</td>
<td>Ratios of Interanalyst Agreement: Foil Category</td>
<td>89</td>
</tr>
</tbody>
</table>
CHAPTER I
THE PROBLEM

The major problem investigated in this study was to determine whether the questions students asked during problem solving would prove to be a fruitful source of data for analyzing patterns of inquiry. For this study a five category analytical system was inductively developed by the investigator who then trained four analysts in its use. The validity and reliability of the instrument was then tested by having the analysts and the investigator code typescripts of classroom problem solving sessions conducted with college students.

Need for the Study

Interest in classroom questions has primarily focused on the teacher. Effective questioning tactics and skills have been considered most important in preparing teachers. Books of educational methods and "methods courses" usually have advocated that teachers acquire and practice effective questioning techniques. In this vein several systems and instruments have been developed to classify teacher questions. Carner's "Levels of Questioning,"1 Amidon and Hunter's

Verbal Interaction Category System (VICS)\(^2\) and Gallagher and Aschner's Convergent-Divergent Analysis\(^3\) are but three examples. The six levels of Bloom's *Taxonomy of Educational Objectives*\(^4\) and the Guilford model of the intellect\(^5\) have served as the bases for several of these systems. Sanders\(^6\) modified the *Taxonomy* to improve the question asking prowess of teachers. Blosser\(^7\) and Cunningham\(^8\) are two of many who have developed instructional sequences to train teachers in their questioning skills.

---


The concern for teacher questions, though, has not been matched in the area of student questions. Student questioning has been virtually neglected. Dodl in his classification system designated two of his twelve categories for student questions. This designation, it should be added, was twice the consideration given to student questions by most observational systems.

Perhaps student questioning has been neglected because the presumptive richness expected by researchers had not been found. Dodl found that students ask few questions in classrooms. His findings may have been accepted as the way things were, causing researchers to focus on the teacher talk of classroom interaction.

Another reason for neglect may have been that no instrument had been developed which discriminated among and classified student questions. Such an instrument would have been unnecessary as long as the numbers and types of student questions were few. The priorities of classroom interaction, however, are changing and more student involvement is being encouraged. With more student involvement there certainly will be more student questions.

---

Suchman with his inquiry development programs\(^1\) attempted to shift the emphasis from a teaching agenda to a learning agenda. Autonomous learning was the basic objective, and student questions were central to this approach. Science was Suchman's curricular vehicle, but soon other areas began to include student inquiry as a basic objective. The "new math," the "new social studies," virtually all science programs, transformational grammar and the linguistic approach to reading are examples. Process became as important as product, and student involvement became a necessary ingredient. Increased student participation and interaction generated more questions, especially student questions. Yet no lens existed with which to examine them, no system to sort the data.

An analysis of student questions has significance for curriculum, instruction, measurement and evaluation. Questions are verbal behaviors that reflect cognitive set, interest, disinterest, progress, frustration and the like. Viewed in context the investigator believed that student questions were observable behaviors that might reveal a portion of the thinking process in operation.

The investigator has had considerable experience with inquiry development programs—developing, demonstrating, editing, evaluating and producing some. It was his experience that problem solving was fundamental to all programs of this type. And the problem presented to students was one that generated interaction—that is, evoked hypotheses, theories, specific explanations and generalizations. With a problem solving format student questions were abundant.

For the most part schools and teachers have not readily accepted inquiry based programs and materials. The inquiry approach normally requires sweeping changes in teacher behaviors and attitudes. Consequently, most of the efforts have been directed toward guaranteeing a teaching success. Interest in the methods for teaching has precluded interest in the learner's method of problem solving. Only the product, the solution to the problem, has received much attention.

Suchman attempted some cognitive mapping. His analysis of a single student's verbal probing, however, has been purely descriptive and somewhat cavalier. No theoretical base, ideal standard or other criteria were established. Also Suchman's classification of student questions was severely restricted by the nature of inquiry specified in his programs.

---

Student inquiry is not limited to problem solving in the physical sciences. Inquiry into geometry or grammar, economics or English is also possible and quite desirable. In these fields the types of questions students ask are different, and the model of question classification used in physical science does not fully apply.

A review of the literature has revealed a singular lack of effort in analyzing student questions. Teacher questioning and observations of classroom interaction have dominated the research. The seventy-nine significant observational instruments reported in Simon and Boyer's *Mirrors for Behavior* focus mainly on teacher behaviors.

This emphasis carries over into the general practice of many teacher educators, which has been to examine the scene, isolate teaching behaviors, and then provide the instruction to teach teachers the skills needed to foster learning. Oversimplified, the general recommendation for improving questioning skill has been that the teacher should ask broader questions.

Research in educational psychology has dealt mainly with unraveling the learning process and examining the effects of various instructional inputs on the learning

---

outcomes. Planaria, pigeons and people have been subjected to different conditions and treatments and their responses measured and analyzed. The interpretations of these data patterns have focused mainly on the instructional end of the teaching-learning act. In the bulk of these studies a learner's questioning behavior has been but a small part of his overall response pattern and was quite often overshadowed by more observable behaviors.

Certainly the efforts to sort out and examine the questioning behaviors of teachers are important. Instruction that has broadened the teacher's questioning pattern has generally broadened the observable learning behaviors of students. The investigator believed that these benefits might be increased, however, if research could describe the inquiry patterns of students as they engaged in problem solving. That premise was the promise of this study.

**Design of the Study**

This study began with the investigator's review of the literature and research that related to classroom observational systems, the question as a component of classroom interaction, and the efforts to improve classroom questioning practices. This selected review is reported in detail in Chapter II of this study.

The investigator then examined the possibility of adapting or modifying an existing observational system or
analytical system to discriminate and classify questions students asked during problem solving. At the outset the optimum in performance was sought. The acceptance criteria were that the system would be all inclusive (for the data) and would have mutually exclusive categories. Later these stringent requirements were loosened and an instrument would have been accepted if it accommodated 90 per cent of the data and if the ambiguity in classifying questions was limited to a 10 per cent variation. The evaluation of each system is presented in Chapter III.

Since no existing instrument was found to be appropriate, the investigator inductively developed his own instrument by analyzing typescripts of problem solving sessions that had been conducted with elementary, secondary and college students. This system and the descriptive specification for each category are described in Chapter III. The procedures for selecting and training four analysts to use the classification system are also explained.

The validity and interanalyst reliability of the instrument were tested by coding typescripts of two inquiry sessions conducted by the investigator with education students at The Ohio State University. The reasons for selecting this population were: (1) Large numbers of students available from which to create random groups; (2) Various socio-economic backgrounds represented; (3) Various academic majors
represented; (4) a general mental maturity not necessarily possessed by younger students; and (5) convenience of the investigator.

The problems for inquiry in these sessions consisted of a discrepant event in physical science, "The Ice Cubes," and a presentation of a case of chronic late behavior.

The two different types of problems were used since the potential solution for each problem was of a different genre. Solutions for problems in physical science usually are more testable and verifiable than are explanations in the behavioral sciences. It was assumed that this variety of problems would yield data that were representative of the full spectrum of possible student inquiry in contrived classroom situations.

In Chapter IV the item-by-item coding record of each analyst for each session is presented and compared. The investigator determined that measures of the instrument's reliability could be obtained by calculating the ratios of agreement between analysts in their use of each category as well as for all classification categories. The ratios of interanalyst agreement, presented in tabular form, are discussed as major findings of this study.


14 Allan Walldren, "Joe: A Case Study," 1967. (Mimeographed.)
The final phase of this study was evaluative. The investigator interpreted the findings as they related to the problem statement. Inferences and implications were drawn and recommendations for further research were suggested. These are reported in Chapter V of this study.

**Definition of Questioning**

To analyze the questioning of students during problem solving it was necessary to consider all of the student verbalizations as questions. This was consistent with the spirit of inquiry. When students are engaged in problem solving simple sentences, short phrases and even single words can carry the verbal probing of inquiry. Consequently all of the student talk during these inquiry sessions was considered interrogative and each student entry in the dialog was coded as a question.
CHAPTER II

A REVIEW OF THE RELATED LITERATURE

Introduction

The thrust of this study was to determine whether student questions provided useful data for analyzing patterns of inquiry. Basically the investigator's purpose in reviewing the literature was to determine what had been reported about classroom questioning practices. Since the earliest of times questions and questioning have been the tools of educators; many regard the question as the most important pedagogical tool. It would be impossible, however, to survey the literature of questioning beginning with Socratic dialog and ending with the inquiry development programs of the 1970's. Consequently the investigator narrowed the scope of the review to explore three general topics that related directly to the purposes of this study.

This review was selective. The investigator attempted to integrate and report significant and representative studies. The first section relates to classroom observational systems. The second section concerns the question as a component of classroom interaction. The third section deals with efforts to improve classroom questioning practices. A short summary concludes the chapter.
Studies of Classroom Observational Systems

Educational research abounds with observational techniques and systems to analyze classroom interaction. The seventy-nine significant studies reported by Simon and Boyer\(^1\) focus mainly on observational systems that encode teaching behaviors. The subsequent analysis of a data record would then yield a basis from which to attempt behavioral modification.

Very few of the systems presented in *Mirrors for Behavior* designate more than two separate categories for student initiated verbal behaviors. Flanders' Interaction Analysis\(^2\) is one of the more fundamental systems and is representative of many. Of the ten categories in the Interaction Analysis, seven are specifically designated for teacher talk: (1) Accepts feeling; (2) Praises or encourages; (3) Accepts or uses idea of students; (4) Asks questions; (5) Lectures; (6) Gives directions; and (7) Criticizes or justifies authority.

Two categories represent the spectrum of classifiable student talk. The first of these, number 8, is used to code

---


a student's response to a teacher's solicitation. A number 8 would normally follow a teacher's number 4. The other category, number 9, is used when a student initiates interaction. A final category, number 10, is used to code both classroom silence and confusion as well as unclassifiable verbal behaviors. Thus, an analysis of student initiated participation in classroom interaction would focus on the frequency with which category 9 and possibly category 10 occurred in the record. In examining category 9, however, an analyst would be unable to sort out the questions students asked from the challenges, interruptions, exclamations, opining, and the like that students offered. Even greater confusion is presented by looking at category 10, which is used to start and end the coding, to classify the "unclassifiable" behaviors, and to collect the data of silence along with the units of pandemonium. The Flanders system, like other systems having teacher centrality, is an inadequate instrument for analyzing student questioning behaviors.

Dodl developed a twelve-category system to record classroom interaction during social studies discussions in elementary schools. He modified and expanded Flanders' system in an attempt to discriminate the types of student talk. His first nine categories are teacher talk units.

---

Category 10 is used to code the solicited responses from students (like Flanders' number 8). Dodl's category 11 applies to student initiated talk and is subdivided into four separate subunits. These are subscripted as: (11a) Request for aid, clarification, direction, etc.; (11b) Thought questions; (11c) Role reversed questions; and (11d) Routine questions. The thought questions (11b) are further subdivided into 11b₁, seeking information or structure, and 11b₂, seeking corroboration, hypotheses. Like Interaction Analysis both silence and confusion are enveloped by the last category, in this case number 12.

The Dodl system was useful in recording the classroom interaction of fourth grade social studies discussions. The number of student questions tallied, however, was alarmingly small. Only 728 pupil questions were recorded out of a total of 43,531 noted behaviors. This means that student questions fill 1.67 per cent of the total spectrum of verbal behavior. In a question saturated environment Dodl's system would be too cumbersome to use in its entirety.

Medley and Mitzel⁴ describe several of the earlier systems used to record classroom interaction. They credit Horn⁵


⁵Ibid.
as being a pioneer in these efforts. Horn used a symbol system (dots, circles, and squares) to capture and record pupil participation. The basic criticism of this system is that it is static in representation.

By 1928 Puckett expanded Horn's system to include evaluations of pupil responses—that is, he qualified the participation with additional symbols that stood for inadequate, adequate, fair, or good responses. Significantly he added two symbols: the figure \( \triangleright \) was used when a pupil asked a question and the symbol \( / \) was used when a student spoke without being addressed.

Graphic systems such as these are partially useful if the situation is highly structured and teacher regulated. As more student participation is solicited and encouraged, however, the "illustrative" systems become very unwieldy and rather impractical.

In classroom observation graphic diagrams were soon replaced by climate maps. The climate, of course, referred to the socio-psychological condition of the classroom. Anderson developed an observational schedule to measure domination and socially integrative behavior. All twenty-three categories of this instrument were descriptive of

\[ \text{Ibid.} \]

\[ \text{Harold N. Anderson, "The Measurement of Domination and of Socially Integrative Behavior in Teachers' Contacts with Children," Child Development, 10 (1939), 73-89.} \]
teaching behaviors. Nearly ten years passed before the "Observation Blank" was revised by Anderson and Brewer to include significant numbers of categories for pupil verbal and nonverbal behaviors.\textsuperscript{8} But by then the die had been cast and virtually all the instruments developed for classroom observation and analysis focused on the adult behavior in the group.

Lewin, Lippitt, and White\textsuperscript{9} manipulated leadership styles with boys' clubs to see what effects were produced. The researchers classified the leadership types as authoritarian, democratic, and laissez-faire. Their findings were not explicitly analogous to classrooms, but there were definite implications for schools.

In 1949 Whithal named (or renamed) this educational phenomenon the Social-Emotional Climate.\textsuperscript{10} Operationally defined it was considered to influence: (1) the inner private world of each individual, (2) the esprit de corps

\textsuperscript{8}H. H. Anderson and Helen M. Brewer, Studies of Teachers' Classroom Personalities, Part I, Dominative and Socially Integrative Behavior of Kindergarten Teachers (Stanford: Stanford University Press, 1945).


of the group, (3) the sense of meaningfulness of group and individual goals and activities, (4) the objectivity with which a problem is attacked, and (5) the kind and extent of interpersonal interaction in a group.\(^1\) To measure cause-effect relationships in the Social-Emotional Climate Whithal developed an instrument having seven categories, each category being directly related to teacher behavior. This system, it should be added, was the precursor for the Interaction Analysis Systems developed by Flanders and others.

The term "Interaction Analysis Systems" is used here in its generic sense rather than as a specific reference to the Flanders Interaction Analysis System. The Flanders system is one of the nominal systems that were developed to record classroom discourse. It seems that after Whithal had given the phenomenon a name educational researchers and field theorists felt they could only explore the classroom with tools of their own design. Instruments were developed for varieties of purposes, but most were modifications of the Flanders original system. The Amidon-Hunter Verbal Interaction Category System (VICS)\(^2\) and the Hough Observational System for Instructional Analysis (OSIA)\(^3\) are two examples.

---

\(^{11}\) Ibid.


Medley and Mitzel\textsuperscript{14} coined the acronym OScAR (Observation Schedule and Record) for a system which was useful in sorting out three factors of classroom interaction: emotional climate, verbal emphasis, and social structure. OScAR has been modified and is still being revised. At this writing five major "editions" of OScAR have been produced.

Macdonald and Zaret\textsuperscript{15} took a different tack in looking at classroom behaviors. Their system placed teacher and student behaviors on a continuum which proceeded from learner-centeredness to teacher-centeredness and followed the behaviors episodically from opening behaviors to closing behaviors.

Morsh\textsuperscript{16} devised a system for supervisors to use to rate instructional effectiveness. The checklist contained categories relating to the instructor's words and deeds as well as the pupils' verbal and nonverbal behaviors. Question asking was one of the twelve overt behaviors Morsh designated to be recorded when observing students. His complete list for student behaviors follows:

\begin{itemize}
\item Medley and Mitzel, "Measuring Classroom Behavior by Systematic Observation," pp. 683-714.
\end{itemize}
Using this system, student question asking would be as important to note as would be student dozing, doodling, or smiling.

It is not the intent of this review to enumerate and briefly describe each of the many observational systems that sprung from the Social-Emotional Climate movement. The studies cited are significant and representative.

The work of Bloom\(^{17}\) had an equally strong influence upon the systematic research in education. The Taxonomy of Educational Objectives: Cognitive Domain was used not only as the basis for evaluating examination questions but also as a take-off for classroom analysis. Sanders\(^{18}\) modification of the Taxonomy was heralded as a significant cognitive measuring stick against which to study teacher questioning patterns. Davis and Tinsley,\(^{19}\) Farley,\(^{20}\) and


Davis, Morse, Rogers and Tinsley\(^2\) basically used the Sanders modification of the Taxonomy to rank the questions asked by teachers and student teachers in training. The intent in each of these studies was the same—to broaden the cognitive range of questions teachers ask.

Concern for recording cognitive behavior increased and classification systems and research designs broadened in scope. The Guilford\(^2\) model of intellectual performance had a great influence in the design and modification of category systems. Gallagher and Aschner\(^3\) developed a coding system directly from the Guilford model. It consisted of the following categories: (1) Cognitive-Memory operations, (2) Convergent thinking operations, (3) Divergent thinking operations, and (4) Evaluative thinking operations. A fifth category, Routine, was used to code procedural and miscellaneous verbal behaviors.


Hunter\textsuperscript{24} modified the Verbal Interaction Category System (VICS) to incorporate the four basic cognitive categories of the Gallagher-Aschner system. This added dimension provided greater feedback information concerning the effectiveness of instructional technique and intent.

The Topic Classification System was developed by Gallagher, Shaffer, et al.\textsuperscript{25} to measure and record classroom interaction within the cognitive realm. The intent was similar to that of Taba, Levine and Elzey\textsuperscript{26} who attempted to show the relationship between what the teacher was seeking and the cognitive responses of the children.

Smith and Meux\textsuperscript{27} in their cooperative research project developed a classification system to study the "logical operations" in classrooms. The logical operations included defining, explaining, stating, designating, reporting, substituting, evaluating, opining, classifying, comparing and

\textsuperscript{24}Elizabeth Hunter, "The Effects of Training in the Use of New Science Programs Upon the Classroom Verbal Behavior of First Grade Teachers as They Teach Science," Classroom Interaction Newsletter, IV (May, 1969), 5-11.


\textsuperscript{26}Hilda Taba, Samuel Levine and F. F. Elzey, Thinking in Elementary School Children, Cooperative Research Project No. 1574 (San Francisco: San Francisco State College, 1964), p. 207.

contrasting, conditional inferring, and directing and explaining. Smith and Meux used the entry of episodes as the basis for their classification scheme.

The observational systems and category schemes cited thus far represent the measuring tools of significant movements in education. It seems that the "toolbox" for each educational movement is quite large. In 1968 Komisar commented on the bewildering number of repetitive systems then in use.

We are rapidly approaching chaos in the production, by researchers, of "new" category systems. No one seems able or willing to tell us why certain categories are chosen or how one researcher's categories bear on those of another.²⁸

Biddle²⁹ suggested that because so many of the new systems lacked sound or systematic theoretical bases, each researcher was bound to develop his own system to meet his own needs.

Biddle's commentary may be a trifle harsh. Most would agree that the type of instrument developed will depend upon the intent of the investigator. The first instruments are usually checklists, tally sheets, perceptual organizers, in other words. They are nominal or descriptive and yield totaled measures. When the parameters of "what exists" have


²⁹Bruce J. Biddle, "Facets of Teacher Role Research," 1968. (Mimeographed.)
been charted and named, investigators and their systems can begin to rank behaviors in an ordinal fashion. Systems of this type are more judgmental and are usually intended to document the need for change. Finally systems are developed and modified to incorporate the "action-in-time" dimension. Sequence and chronology represent the dynamics of interaction and group process. An analysis of a process is quite inferential and subjective which may account for the great diversities in interpretation.

Studies of the Question in Classroom Interaction

In 1912 Stevens\(^3\) estimated that 80 per cent of classroom time was spent in question asking discourse. The rapid fire questioning practices of teachers was most alarming to her. She felt that a practice of asking 79 questions per class period would prevent assimilation and association of ideas as well as stifle pupil verbalization. When the number of teacher questions tallied 150 in a forty minute period, Stevens felt that a definite negative effect was produced.

Yamada\(^3\) concurred with Stevens, adding that children are unable to recall specific experiences and relate them

\(^3\)Romiett Stevens, *The Question as a Measure of Efficiency in Instruction* (New York: Teachers College, Columbia University, 1912).

\(^3\)Sochichi Yamada, "A Study of Question," *The Pedagogical Seminary*, 20 (June, 1913), 129-86.
in a meaningful way if the students are being continually bombarded with teacher questions. Yamada also added that over two-thirds of normal classtime was spent in the question-answer mode.

Corey\textsuperscript{32} observed classes in a laboratory high school for one year. He compiled a total of 39,000 questions during this time. Of these questions, he noted that fewer than 4000 had been asked by students. Corey further reported that approximately 38 per cent of the teachers' questions went unanswered by the students, meaning that the "no response" was generally followed by the teacher answering his own question.

Houston\textsuperscript{33} noted the same paucity of student questions in junior high school classrooms. He found that less than one student initiated question was asked per class period.

Corey\textsuperscript{34} conducted another study, recording for one week all talk in junior and senior high school English and science classes. In one science class students accounted for 17 per cent of the questions asked, but in the English classes only 1 per cent of the questions were student queries.

\textsuperscript{32}S. M. Corey, "The Teachers Out-Talk the Pupils," The School Review, XLVIII (December, 1940), 745-52.

\textsuperscript{33}V. M. Houston, "Improving the Quality of Classroom Questions and Questioning," Educational Administration and Supervision, 24 (1938), 17-28.

\textsuperscript{34}Merideth D. Gall, "The Use of Questions in Teaching," Review of Educational Research, 40 (December, 1970), 715-16.
In forty primary classrooms Floyd found that teachers also dominated the verbal interactions. Student initiated questions constituted only 3.75 per cent of the first grade, 5.14 per cent of the second grade, and 3.64 per cent of the third grade classrooms he analyzed. Floyd's finding showed that teachers asked 3½ to 6½ questions per minute with the average teacher asking 348 questions per day.

In Dodl's study of elementary classrooms, of the 43,531 behaviors recorded only 728 were pupil questions. Dodl categorized the pupil questions into five major categories. The order of frequency of occurrence was: (1) information seeking, (2) clarification, (3) hypothetical, (4) role reversal, and (5) routine-permission.

In 1968 Johns, using a modification of Flanders' Interaction Analysis, sought to determine the quality of questions asked by high school English teachers in Michigan. His hypothesis was that the more indirect teacher (according

---


to Flanders' scale) would ask more thought-provoking questions and so would the students in the class. His findings seem to indicate that more thought-provoking questions were asked by the indirect teachers, but he noted that very few student questions were asked in either environment, direct or indirect.

Adams\textsuperscript{38} assayed the questioning practices in junior and senior high school English and social studies classes. Compared with Stevens' data, Adams' findings yielded significantly fewer memory type questions. Perhaps even more importantly, though, he found that the question categories differed significantly in the separate subjects as well as with the grade levels and ability groups.

Kondo\textsuperscript{39} examined differences in question type and question frequency with elementary teachers using materials from the Science Curriculum Improvement Study (SCIS). In his reporting he inferred that the differences in questioning styles were caused more by individual characteristics of the teachers than by the types of lessons presented.


\textsuperscript{39} Allan K. Kondo, "The Questioning Behavior of Teachers in the Science Curriculum Improvement Study Teaching Material Objects" (paper presented at the National Association for Research in Science Teaching meeting, Pasadena, Calif., Feb. 7, 1969).
Kleinman\textsuperscript{40} reported similar findings with general science teachers. She found that the kinds of questions asked are fairly stable for each teacher. Kleinman also noted that the teachers who asked more critical thinking questions tended to ask fewer questions.

Classifying the structure of science questions according to their intent was the goal of a study by Moyer.\textsuperscript{41} He found that many teacher questions were unanswered and that answers to 40 per cent of all questions required little mental effort on the part of the students.

Hunkins\textsuperscript{42} suggested that elementary teachers do not adequately phrase their questions. A teacher's excessive verbal qualification of a spontaneously generated query often clouds the intent of the question.

The writings of Fristoe,\textsuperscript{43} Gagnon,\textsuperscript{44} and Batchelder,

\textsuperscript{40}Gladys S. Kleinman, General Science Teachers' Questions, Pupil and Teacher Behaviors, and Pupils' Understandings of Science (Ann Arbor: University Microfilms, 1964).


\textsuperscript{42}Francis P. Hunkins, "Using Questions to Foster Pupils' Thinking," Education, 87 (October, 1966), 83-87.


McGlasson and Shorling warn against the general practice of low level questioning. Perhaps the most succinct observation is Moyers' generalization that teachers basically are unaware of the classroom potential of effective questioning practices.

The classic study of classroom interaction was conducted by Bellack, Kliebard, Hyman and Smith. They analyzed the verbal interaction of 15 high school teachers and 345 students in social studies classes. One of their findings was:

The teacher is the chief solicitor; he speaks 86.9 per cent of all soliciting moves. For the teacher 46.6 per cent of his moves are solicitations, while for the pupil only 11.3 per cent are solicitations.

The investigators further explained:

When the teacher solicits he usually expects one pupil to respond; when the pupil solicits, he usually expects the teacher to respond. The intended interaction regardless of who solicits, is almost always between teacher and pupil, rather than between pupil and pupil.


47 Ibid., p. 129.

48 Ibid., p. 130.
Concerning student questioning Bellack states, "When the pupil launches on his own initiative, he generally does so by announcing his intention of asking the teacher a question."\textsuperscript{49} This restrictive pattern of student questioning behavior seems to be one of the "rules of the classroom game" that students have accepted and teachers have reinforced.

From the studies mentioned it can be seen that the paucity of student questions is an established finding. It is also obvious that much of the teacher's verbal behavior is of the question asking type. Therefore, some attention has been given to improving the questioning practices of teachers and student teachers.

\textbf{Efforts to Improve Classroom Questioning Practices}

There are a number of programs intended to change teacher questioning behaviors. Most of them were developed for the curricular areas of science, social studies, and reading and are primarily geared to elementary classroom practice.

Motivated by his observational data of classroom practice, Houston\textsuperscript{50} developed an inservice program to improve questioning practices of teachers. Group conferences,

\textsuperscript{49}\textit{Ibid.}, p. 245.

\textsuperscript{50}Houston, "Improving the Quality of Classroom Questions and Questioning," pp. 17-28.
stenographic reports, self-analyses, and supervisory evaluations were the crux of the training program. Working with eleven teachers Houston reported that the quality of questioning practices was significantly raised. For example, relevant questions increased from 41.6 per cent to 67.6 per cent; student participation rose from 40.4 to 56.1 per cent in those classes taught by the eleven teachers. Bothersome teaching habits also declined. The teacher's repetition of his own question fell from 4.8 to zero occurrences; repetition of student answers fell from 5.5 to .6 occurrences; and answering one's own questions dropped from 3.5 to .3 occurrences.

Schreiber\textsuperscript{51} researched the questioning practices of elementary teachers and developed an inservice instructional program that enabled the teachers to examine and analyze their own questioning techniques. The results of her study indicated that instruction in questioning did make a difference in classroom performance.

Following Schreiber's lead Cunningham\textsuperscript{52} attempted to improve the question phrasing skills of elementary teachers


\textsuperscript{52}Roger T. Cunningham, "A Descriptive Study Determining the Effects of a Method of Instruction Designed to Improve the Question-Phrasing Practices of Prospective Elementary Teachers" (unpublished Ph.D. dissertation, Indiana University, 1968).
in training at Indiana University. His study used the Gallagher-Aschner classification as a means for designating question type. His findings showed that student teachers could be trained to ask more of the "broader questions" within a short period of time. Awareness of cognitive intent and impact of the question upon the learner were prerequisites for success.

Blosser\textsuperscript{53} tested an instructional program to improve the questioning skills of prospective science teachers. Her performance criteria were: (1) to increase the number and quality of "open" questions asked by student teachers, (2) to pause long enough to allow pupils time to think before responding, and (3) to question in such a manner as to decrease the amount of teacher talk during a lesson. Her findings indicated that questioning appeared to be a skill that could be developed through instruction and practice.

Clegg, Farley, and Curran\textsuperscript{54} attempted to train student teachers to recognize the cognitive levels of questions


asked in classrooms. Of course the second purpose was to upgrade the questioning practices of the student teachers. Bloom's Taxonomy was used as the basis for classification in this study. The investigators found that student teachers asked a wide range of questions; however, 54 per cent of all the questions tallied fell into the two lowest levels of the Taxonomy (Knowledge and Comprehension).

Farley found that a group of student teachers who received training with the Sanders modification of Bloom's Taxonomy asked a degree of "higher order" questions than a group of student teachers trained in Flanders' Interaction Analysis.

Farley and Clegg in a paper presented at the annual AERA convention in 1969 reported that training in the use of the Taxonomy did upgrade the cognitive level of questions teachers ask, but the level of questions seldom rose above the third level, Interpretation.

The results of Cross's study show that intern teachers could formulate "higher order" questions. She attempted

---

55Farley, Increasing the Cognitive Level of Classroom Questions.


57Janet S. Cross, The Development of an Instructional Program to Enable English Teachers to Improve Discussion Skills (Ann Arbor: University Microfilms, 1968).
to train intern English teachers to broaden the range of questions that usually would be used in running discussion groups. After the training sessions had been completed, Cross found that the interns did indeed ask higher order questions. In fact the interns asked so many of the "better questions" with such frequency that the discussion sessions proved almost unproductive for the pupils.

Wilson^58 noted that during science lessons elementary teachers trained in inquiry programs asked a greater number of questions that required more analytical thinking than did teachers not similarly trained. Porterfield^59 found similar results with elementary reading teachers who asked significantly different proportions of divergent questions. At Michigan State University Moon^60 and Bruce^61 achieved equally significant improvements in teacher questioning strategies after teachers had been rigorously trained in the SCIS program.

---

58 John H. Wilson, "Difference Between the Inquiry Discovery and the Traditional Approaches to Teaching Science in Elementary Schools" (unpublished Ph.D. dissertation, University of Oklahoma, 1967).


60 Thomas C. Moon, A Study of Verbal Behavior Patterns in Primary Grade Classrooms during Science Activities (Ann Arbor: University Microfilms, 1969).

61 Larry R. Bruce, A Determination of the Relationships Among SCIS Teachers' Personality Traits, Attitude toward Teacher-Pupil Relationship, Understanding of Science Process Skills and Question Types (Ann Arbor: University Microfilms, 1969).
Also to be included in programatic attempts to improve questioning practices would be Sanders' book, *Classroom Questions: What Kinds?* This work evolved from Sanders' work with the Wisconsin Improvement Program and four summer workshops held in Manitowoc for the purpose of devising guidelines to help teachers improve their questioning practices. While this book uses social studies as its vehicle of illustration, its intent is to apply questioning practices to all classroom disciplines.

The *Minicourse* developed by Borg, Kelley, Langer, and Gall represents still another attempt to improve questioning strategies. The following description is by Gall.

...it is a self-contained, inservice training package requiring about 15 hours to complete. The minicourse relies on techniques to effect behavioral change. In a field test with 48 elementary school teachers, the minicourse produced many highly significant changes in teachers' question behavior, as determined by comparisons of pre- and post-course videotapes of 20 minute classroom discussions: increase in frequency of redirection questions (questions designed to have a number of students respond to one student's original question) from 26.7 to 40.9; increase in percentage of thought questions from 37.3 per cent to 52.0 per cent; and increase in frequency of probing questions (questions which require students to improve or elaborate on their original response) from 8.3 to 13.9.

---


64 Gall, "The Use of Questions in Teaching," p. 717.
Oliver and Shaver\textsuperscript{65} developed a program to train teachers to use questioning methods appropriate to discussing controversial issues in social studies. The two basic styles were Recitation teaching and Socratic teaching. To Oliver and Shaver a teacher's style difference was measured only by different learning outcomes.

Taba\textsuperscript{66} developed a system of teacher training built around questioning strategies. Her program focused on social studies and the inquiry strategies teachers were to employ would help students to form concepts, explain cause and effect relationships, and explore implications.

Sloan and Pate\textsuperscript{67} spoke for more than the School Mathematics Study Group when they called for strong inservice training programs to help teachers acquire the questioning skills needed to teach the "new Math" (SMSG) curriculum. Generally a prerequisite for each new curriculum package was that teachers must modify their behaviors to effectively teach the program. Questioning tactics, of course, were high on the list for change.


Perhaps the program requiring the greatest behavioral change in teachers was the Inquiry Development Program conceived by Suchman\textsuperscript{68} at the University of Illinois. The original program, often called Illinois Inquiry Training, was intended to develop the inquiry skills of gifted children. Suchman's materials consisted of problems--discrepant events--in the physical sciences. In presenting a problem to a class the teacher became an answerer of student questions rather than a questioner of students. The subsequent studies and cooperative research sponsored by the United States Office of Education\textsuperscript{69} indicated that this dramatic role reversal did indeed engender active student participation and productive inquiry, and soon inquiry development programs were being expanded and utilized with the not-so-gifted populations.

The inquiry development programs that followed were tangible programs that incorporated many of the questioning practices that had been suggested for many years by prominent educators and classroom chroniclers.

In reviewing the literature of questioning practices the investigator has found recurring attempts to get teach-


\textsuperscript{69}Ibid., and \textit{Illinois Studies of Inquiry Training}, N.D. (Mimeographed and dittoed materials.)
ers to think about and improve their classroom questioning. Most of these articles appear in the more popular educational journals and are steeped in common sense rather than research findings. A few examples will illustrate the admonishing style.

Wellington and Wellington are critical of teachers who take sole ownership of the right to question in classrooms.

In no dictionary does it state that a question is defined by the term "teaching," or by the idea of asking someone to play a guessing game for the answer the teacher has thought of.  

Klebaner does not disagree with the Wellingtons, but the title of her article, "Questions That Teach," seems to reaffirm the importance of the question for the teacher.

Waetjen contends that the teacher's questions play a significant role in the thinking level of students. Cross and Nagle state:

The behavior of the teacher is, perhaps, the most critical determinant of how involved students can and will become during classroom discussion. And, ultimately it is a teacher's questioning skills which most influence both degree and level of his students' involvement.

---

70Jean Wellington and C. Burleigh Wellington, "What's a Question?" Clearinghouse, 36 (April, 1962), 472.


73Janet S. Cross and John M. Nagle, "Teachers Talk Too Much!" English Journal, 58 (December, 1969), 1363.
Aschner,^74 Carner,^75 and Pate and Bremer^76 all suggest that skillful teacher questioning yields productive student thinking. These writers and others are, of course, promoting the asking of higher order questions to engender higher order thinking.

To describe the characteristics of a "good" question would be purely an academic exercise and of limited value. It is rather the organization of kinds of questions that leads to fruitful inquiry. There is not a Socratic question but there is a Socratic dialog.

Summary

Our perceptions and understandings of the general classroom interaction have been biased by the instruments and techniques that have been used for observing. In general the focal lengths of those lenses were clearly ground to view the teacher. Field theorists usually had hunches about which behaviors were the significant and effecting behaviors. And these hunches were translated into separate categories of measurement scales. Although a researcher would emphasize


the nominal nature of his system, practitioners often would apply graduated pedagogical values to the summed data.

The criticism of this practice is not meant to be a chastisement, merely an observation. To be educationally useful an instrument or observational technique must help an investigator find cause and effect relationships. Usually reliability in measuring is high, but validity through interpretation may not be.

The instrumentation that evolved from the Social-Emotional Climate movement assumed that a teacher's classroom verbal behavior was more significant than other behaviors, conditions and events. The nonverbal dimensions of classrooms were virtually ignored as were the ephemeral forces of group dynamics.

The multidimensionality of classroom coding devices increased fantastically when cognitive activities were considered and recorded along with affective behaviors. The subsequent interpretation of a record of events then usually included many inferences about both cognitive level and intent of behaviors. Cause-effect relationships were postulated and, for the most part, accepted as fact.

What resulted, among other things, was a renewed attempt to train teachers to modify their verbal behaviors. Positive learning results were usually documented when a "new and improved" teaching style was compared to another style with
reference to pupil scores on some achievement test. Of course, the teaching styles were described by a classification system, usually a modification of the system originally used to record and report what went on in classrooms.

That course of events seems logical and most worthwhile. Yet when the sequence of events is viewed from a longer perspective there appear to be certain weaknesses. Most of the observational instruments were developed inductively—that is, evolved from classifying the "what is" of classroom education. When the reports of "what is" were evaluated, many educators felt that "what is" should become "what was." The task then became to determine "what should be," and to develop programs that would produce that state. In these efforts a classification system often was modified to be used as a self-evaluation inventory, a supervisor checklist, or as a basis to formulate guidelines which would outline the route to improved teaching practices.

Teachers were instructed, and feedback (measurement) for evaluation was collected by using the observational system (or a modification of it). The investigator questions this linear process. A physicist attempting to measure momentum without using a timer seems analogous. In any interaction the total number of variables clearly is not known. To deduce or conclude from inferences drawn from data that have been recorded by observers who used a system
that was designed to code overt behaviors as expressions of covert intentions within a contrived setting in which it is assumed that one person, the teacher, basically regulates the social and learning dynamics clearly is a limited generalization about "what can be" in education.

Most studies of classroom observation reported that student initiated questions were not considered indices of significant behavior. No doubt students had and still have questions. The investigator believes, however, that students have learned to answer questions in the classroom, not to ask them. Aside from procedural queries most instructional agendas have not allowed for or encouraged student inquiry. As long as the educational emphasis remained with teaching, there would be no need to have a system to catalog student questions.

Active student involvement and inquiry, however, have become objectives of several new educational programs. The teaching methodology necessary to effectively implement these inquiry based programs incorporates, invites and encourages student questions. Yet no system has been developed that helps to analyze the questions students ask while problem solving. The investigator believes that the development of such an instrument is prerequisite if we are to assess the educational significance of these new programs.
CHAPTER III

PROCEDURES USED IN DEVELOPING THE INSTRUMENT
AND IN TRAINING ANALYSTS TO USE IT

The broad concern of this study was to determine whether student initiated questions, asked during problem solving, were useful data for analyzing patterns of inquiry. The first specific task of this investigation was the development of an instrument that could be used for an analysis of inquiry when student questioning becomes the focus. The second task involved training analysts to use the instrument with reliability.

The basic methodology for this study evolved from the investigator's past experiences with inquiry development programs and his more recent exposure to educational research techniques. Fundamentally this study is descriptive, due to the heuristic nature of the investigation. The investigator decided that the simple statistical treatment of data was appropriate for the purposes of this study.

This chapter consists of two parts. The first deals with the formulation of the instrument, its applicability and general utility. The second part deals with the program employed to train analysts and describes the methods used to evaluate the reliability of the instrument and the training.
Background

In 1966 the investigator was employed by an educational publisher\(^1\) to head a national campaign to promote inquiry development programs and to secure their adoption. This meant "selling" school districts, supervisory personnel and teachers on the inquiry idea for their classes. At that time the inquiry development programs were not the most sought after educational innovation. Traditional teaching techniques had to be greatly modified in order to effectively implement an inquiry based curriculum. This requirement alone prevented mass adoption.

Other factors, too, weighed heavily against inquiry. For example, the subject coverage of an inquiry program often did not match the prescribed state curricular specifications which meant that adopting districts could not get state funds to purchase the materials. Also the format of the materials was different. Instead of having a single textbook with accompanying student workbook and teacher guide, an inquiry development program might (and usually did) consist of a collection of multimedia materials—films, records, books, kits and so on.

Much later some of these program deficiencies were regarded as program assets, but one ringing criticism of

inquiry programs remained, student evaluation. The investigator clearly remembers that that problem--not having a definite evaluation system, nay a grading system--was the reason most frequently cited by teachers, supervisors and administrators in rejecting any inquiry program.

To overcome this "shortcoming" the investigator encouraged and worked with the authors and developers to formulate evaluation instruments and procedures. These efforts resulted in a publication three years later.2

These experiences are mentioned here to serve as a preface to the study and to indicate some of the investigator's earlier concern with the evaluation of student inquiry during problem solving. Perhaps the basic problem of evaluating inquiry then as it is now is that a process must be measured as a product.

The Search for the Instrument

The investigator began this study by reviewing the literature and research that dealt with question classification systems. Virtually every system reported was intended to be used to code teacher questions, and only a very few systems had categories that could be used for classifying

student questions. The systems developed by Suchman were the only systems explicitly designed to code student questions exclusively. And these systems and instruments were expressly designed for evaluating student progress in the commercially published inquiry development programs developed by Suchman and by Suchman and McCombs.  

Suchman's Analysis of Inquiry

The investigator decided that since all of these systems and instruments evolved from Suchman's first analysis of inquiry, "The Question Record Sheet," this prototype might be a useful starting point. "The Question Record Sheet" is a device that was designed to help the teacher interpret the kinds of questions students asked during problem solving. The designations for the types of questions asked are two-dimensional—that is, a student's question is classified according to its focus (events, objects, conditions, or properties) and also classified according to its cognitive level (verification, experimentation, necessity, or synthesis). The coding symbols are shown below:


4J. Richard Suchman, Developing Inquiry, p. 60.

5Ibid., p. 57.
In the manual for his first inquiry development program, Suchman defined each category as follows:6

**Verification (V)** This class includes all questions that seek to identify or verify some aspect of the given event.

**Experimentation (E)** In this class are all questions that attempt to ascertain the consequences of some change in the given experiment. Experimental questions are always hypothetical.

**Necessity (N)** All questions that seek to determine whether a particular aspect of an event was necessary for the outcome are in this class.

**Synthesis (S)** This class includes all questions that seek to determine whether a particular idea about causation is valid. Actually, such questions are theories for which the child is seeking approval through authority, rather than attempting to verify them himself by gathering data.

**Events (e)** are any happenings, apart from any analysis of them.

**Objects (o)** are slightly more abstract than events. An object represents a separate part of a whole happening. Objects are timeless.

**Conditions (c)** are the states of objects.

**Properties (p)** are characteristics of objects that do not change with time.

---

6Ibid., pp. 56-57.
The investigator found that Suchman's scheme was useful in classifying student questions generated during inquiry sessions conducted by Suchman with sixth graders who were verbally exploring a discrepant event in physical science.\(^7\) The utility of the system waned greatly, however, when the investigator tried to code questions asked by seventh graders who were engaged in solving a problem in social science.\(^8\) The cognitive levels of seventh graders' questions had to be inferred and the focusing classification (objects, events, conditions, and properties) proved to be cumbersome and inappropriate designations.

When the investigator modified the Suchman system by dropping the focusing classification the system was not all-inclusive and category designations for many questions remained largely inferential. The investigator found his own designations changing each time he reviewed the typescript of the same inquiry session. Consequently the Suchman system was rejected since it was not completely useful for classifying the questions students asked in social studies problem solving.

\(^7\)Audio tape of Suchman and sixth grade students, Alexandria, Virginia, 1965.

\(^8\)Audio tape of Allan Walldren and selected seventh grade students of Norton Middle School, Columbus, Ohio, 1971.
Carner's Levels of Questioning

The levels of questioning suggested by Carner\(^9\) were examined next. Although Carner's purpose in formulating the three different levels of questioning was to improve teaching practices, the investigator felt the system should be tried with student questions.

The cognitive levels of Carner's system are:\(^{10}\)

**Level I Concrete**

Questions dealing with comprehension of detail—rote memory.

**Level II Abstract**

Questions dealing with:
1. perceiving relationships
2. sensing continuity and sequences
3. making inferences
4. drawing sound conclusions
5. evaluating

**Level III Creative**

Basically questions that begin "what would happen if...?"

When Carner's levels were held against the questions students asked during two separate inquiry sessions almost all the queries were classified as Level II (Abstract). The investigator then tried to further discriminate these questions according to the separate dimensions of the level.

---


\(^{10}\)Ibid., pp. 548-50.
This task proved to be highly interpretive since an "inference question" closely resembled a "conclusion question." And either of these could have been classified as a Level III (Creative) question.

Consequently, the investigator rejected Carner's Levels of question classification. It was not a reliable instrument to classify student questions. The system lacked adequate power for discrimination and some of the categories were rather ambiguous.

Dodl's System

Dodl\(^ \text{11}\) had developed a classroom observational system to measure pupil questioning behavior in the context of classroom interaction. Like many other classroom observational systems it had been derived from the Flanders Interaction Analysis and then extensively expanded. The categories relating to student talk were examined by this investigator to see whether that portion of Dodl's scheme could serve as the basis for analyzing student questions. Dodl had designated Categories 10 and 11 for student talk. He described each category in the following manner.\(^ \text{12}\)


\(^{12}\)Ibid., Appendix A.
10. **Solicited response:** Pupil complies verbally as expected by the teacher.

11. **Pupil initiated talk:** Pupil talk which goes beyond expected response of 10 or involves immediate decision involving the weighting of evidence, fact or opinion. If the pupil talk is a question it should be subscripted as follows:

   a. **Request for aid, clarification, direction:** These questions typically imply a failure to hear or comprehend.
   
   b. **Thought question:** Question from which it may be inferred the pupil is pursuing his own course in seeking to learn. These questions should be subscripted:
      1. Seeking information or structure
      2. Seeking corroboration or information which determines the next step for the pupil. May be termed hypothetical question.
   
   c. **Role reversed question:** Student assumes role of teacher or discussion leader.
   
   d. **Routine question:** Courtesy questions, permission.

The categories conceived by Dodl were useful designations for student talk in teacher dominated classrooms. The investigator found that most student questions that were asked during inquiry sessions (student domination) were covered by Dodl's Category 11b, Thought questions. The investigator found that most of the questions that were asked by sophisticated and experienced inquirers were classified by Category 11b, Thought questions, seeking information or structure.

A finer discrimination of questions was deemed necessary, but the investigator found that to further analyze a question with Dodl's criteria became largely an inferential
and interpretive task. Also two categories, Role reversal and Routine, were not used at all. The investigator rejected Dodl's system because it did not provide the power of discrimination necessary for question analysis.

Bloom's Taxonomy

Bloom's Taxonomy of Educational Objectives\(^{13}\) has been extensively used as a set of guidelines for measuring cognitive activity. Educators and researchers have used the six levels of cognition as measures of verbal activity and as a hierarchy for questions in test construction. Stodola and Stordahl\(^{14}\) describe the levels quite simply.

1. Knowledge (remembering facts, terms and principles in the form in which they were learned)
2. Comprehension (understanding material studied without necessarily relating it to other material)
3. Application (using generalizations or other abstractions appropriately in concrete situations)
4. Analysis (analyzing elements in a communication, the relation between and structure of these elements)
5. Synthesis (placing elements together to form a new pattern)
6. Evaluation (judging the value of material for a specific purpose)


3. Interpretation: The student discovers relationships among facts, generalizations, definitions, values and skills.

4. Application: The student solves a lifelike problem that requires the identification of the issue and the selection and use of appropriate generalizations and skills.

5. Analysis: The student solves a problem in the light of conscious knowledge of the parts and forms of thinking.

6. Synthesis: The student solves a problem that requires original, creative thinking.

7. Evaluation: The student makes a judgment of good or bad, right or wrong, according to standards he designates.

Obviously the rationale for each question class was derived from the intended pupil response. This modification of Bloom's Taxonomy was a tool that had been designed to upgrade teacher questioning practices. The investigator tried to convert the explicit teacher expectancies for each category into descriptions of possible student expectancies. For example, Category 3 could be restated in this way:

Interpretation: The student asks a question about relationships among facts, generalizations, definitions, values and skills.

While this minor modification seemed appropriate for Categories 1, 2 and 3, the sense of it all broke down for Categories 4, 5 and 6. Other attempts to revamp the criteria yielded either inappropriate classes or destroyed the sequence and logic from which the system evolved.
The order of levels in the Taxonomy represents a systematic progression of cognitive activity starting with the lowest order of thinking and ending with the highest. When this system was used as a basis for classifying student questions the investigator found that Categories 3 and 4 captured most of the questions. Categories 1 and 2 were not used. The system had no category in which to put questions that sought information. Also the investigator found it a confusing task to classify questions as Category 5 or 6. Since most of the student questions were not easily categorized, the Taxonomy was rejected as being inappropriate for analyzing student questions.

Sanders' Modification

The investigator believed that Sanders' modification of the Taxonomy would be similarly inadequate for analyzing student questions. Sanders had modified and expanded Bloom's system to create the "Taxonomy of Questions" consisting of the following categories:\(^{15}\)

1. **Memory:** The student recalls or recognizes information.

2. **Translation:** The student changes information into a different symbolic form or language.

Like its precursor, Bloom's Taxonomy, the Sanders' "Taxonomy of Questions" was found to be inappropriate for analyzing student questions. Too many questions could not be classified.

"QSOS"

The investigator next turned to the Questioning Strategies Observational System (QSOS) developed by Morse. This system contains twenty-four separate categories and was developed to code the classroom questioning of teachers. The present version of the QSOS does not contain categories for student initiated questions. In a monograph it is mentioned that categories for pupil initiated questions as well as textbook and preframed questions will be added to the basic QSOS format at a later date.

The present QSOS is basically a two-dimensional system. Teacher questions are coded according to their designation (Mass, No one, Group, or Individual Designated) and to their level of cognitive intent (Knowledge, Comprehension, Application, Analysis, Synthesis, Evaluation, Affectivity, or


Procedure). Using this system cognitive activity as well as classroom management can be monitored. However, until the QSOS incorporates categories to record student initiated questions it will not be a useful instrument for analysis of student inquiry. Therefore, it was rejected.

Other Systems

The investigator rejected virtually every other classroom observational system by merely examining the categories and noting the purposes of the developer. For the most part systems either were designed to capture the broad spectrum of teacher behaviors (such as OScAR\textsuperscript{18}, FIAC\textsuperscript{19}, and VICS\textsuperscript{20}) or were too specific in dimension to even consider and discriminate student initiated questions (such as the systems of Clements\textsuperscript{21}, Oliver and Shaver\textsuperscript{22}, and Smith and Meux\textsuperscript{23}).

\textsuperscript{18}Observation Schedule and Record (OScAR) developed by Donald M. Medley, et al. to code teacher verbal behavior in classrooms.

\textsuperscript{19}Flanders Interaction Analysis Categories (FIAC) developed by Ned Flanders and Edmund Amidon to code teacher verbal behavior in classrooms.

\textsuperscript{20}Verbal Interaction Category System (VICS) developed by Edmund Amidon and Elizabeth Hunter both to code teacher verbal behaviors and to classify cognitive levels.

\textsuperscript{21}Robert D. Clements' classification of Question Types, Patterns and Sequences Used by Art Teachers in the Classroom.

\textsuperscript{22}Donald W. Oliver and James P. Shaver's system to analyze Socratic and Recitation techniques employed by teachers in social studies classes.

\textsuperscript{23}B. Othanel Smith, Milton Meux, et al. systems to analyze the logic of teaching and the strategies of teaching.
Considering the purpose of this study the investigator concluded that systems of cognitive measure would be most useful. Most of these systems, however, focused on teacher questions and were not useful for analyzing student initiated questions. The developers of cognitive measuring systems seemed to have a single purpose—to improve or upgrade teaching practices. In a preliminary examination, only one system, the Gallagher-Aschner system, seemed to hold some promise for this investigation of student initiated questions.

The Gallagher-Aschner System

For their studies of classroom behaviors Gallagher and Aschner sought an instrument that would be useful in coding cognitive behaviors.24 Using the operations dimension of the Guilford model of the intellect, Gallagher and Aschner identified four basic categories that could be used descriptively to classify cognitive behaviors. The four categories and their definitions as they relate to teacher questions are taken from Cunningham.25


1. **Cognitive-Memory** (questions that probe student's recall, memory or recognition)

2. **Convergent** (questions that demand an integration of facts, usually leading to a conclusion)

3. **Divergent** (questions that usually ask for creative and imaginative responses to problems)

4. **Evaluative** (questions that require the student to judge or take a position)

Categories 1 and 2 can be regarded as classifiers of "narrow" questions and Categories 3 and 4 as classifiers of "broad" questions.

Using this four-category scheme Gallagher and Aschner noted that the cognitive level of a teacher's question yielded a student response at the same cognitive level. This was encouraging to educators who were seeking ways to raise the cognitive activity in classrooms.

The investigator tested the efficacy of the Gallagher-Aschner system by analyzing a typescript of questions seventh grade students had asked while problem solving.\(^{26}\) The cognitive levels for many student questions could not be determined. For the most part students were converging, trying to find explanations for an event. Most students asked questions that basically sought information not at hand. Several students asked for approval of their

---

\(^{26}\)Allan Walldren and selected seventh grade students from Norton Middle School, Columbus, Ohio, 1970.
theories and guesses. Using the Gallagher-Aschner criteria it was impossible to determine the purpose or cognitive level of many questions. Also the first category, Cognitive-Memory, was virtually unused by the investigator.

Since the system would not accommodate numbers of student questions and because the investigator could not consistently catalog the cognitive levels of several questions, the Gallagher-Aschner system was rejected.

At this point in the study the investigator was discouraged. None of the existing classification systems was particularly useful to analyze the questions students asked during problem solving. The investigator discovered that the instrument had to be developed inductively.

The Development of the Instrument

The preliminary search for generic categories of question types began with the investigator listening to tapes and examining typescripts of inquiry sessions conducted by Suchman and the investigator. In most of these sessions the students were generally slow to ask questions and the questions seemed to be either very narrow--asking for specific details--or very broad--asking for approval of a hypothesis or generalization.

At this time and for quite a different purpose the investigator was conducting some action-research with both of his undergraduate education classes at The Ohio State University. Both classes had been divided randomly into groups of equal size (18 members in each). The instructor's intention was to compare the efficiency of learning of the separate and relatively equal groups. One group of each class was "taught" to use a two-camera video tape system having a special effects generator. The other group in each class was not taught but encouraged to learn how to operate the same system on their own. Specific questions were answered but no other instruction was given.

This action-research with video taping was conducted for forty-five minutes each day for two consecutive days. Only four students (two in each class) had any prior experience with video tape equipment and these students by a stroke of good fortune were placed in separate groups through the random assignment.

The instructor audio taped all sessions. He attempted to answer all questions that were asked by the "taught" groups, but his basic purpose was to instruct these students to use the cameras and special effects generator to make a video tape. The "inquiry" groups had the same

---

28Education 435, Theory and Practice in Secondary Education (a general methods course).
objective, to make a tape, but had no instruction aside from the specific answers to the specific questions that were asked. The instructor only once interceded with one "inquiry" group. Worried about the precarious positioning of a camera, he said in a rather loud voice, "That's top heavy, you know. It could fall!"

The tapes of the first day of the action-research were transcribed and the interaction examined. As expected in the "taught" sections there were long monologs and few questions. And a good proportion of these few questions were not directly related to the function and operation of the equipment.

Examples are:

"Did Sony make this stuff first?"

"Who else makes this?"

"Have you seen the new Panasonics?"

"What's all this cost?"

With the "inquiry" groups not many questions were asked during the first ten minutes or so. Then the questions were mostly questions of application and implication.

"Will this clear it up?"

"Does this go here?"

"Can I make it darker?"

"Can we splice this tape like recording tape and make a movie?"
Midway through the second day of this action-research the "taught" groups had been instructed on all components of the system. In contrast the "inquiry" groups had taught themselves how to operate all the components in roughly forty minutes of the first day. The "inquiry" groups also had set tasks for themselves and had rather successfully tested the limits of the equipment.

To conclude this action-research each group was given a portable video tape unit made by a different manufacturer and told to make a target tape. This involved operating similar equipment, but not an identical unit. Each group had to unravel cables, connect them properly, thread the tape and use a camera to record an establishing shot, to zoom on a fixed target and to execute a smooth pan from one target to another. The quality of each target tape was independently evaluated by two judges. A record of the time each group took to make its target tape was recorded by the instructor.

Although this description of an action-research effort may seem inappropriate to the procedures of this study, this experiment provided the turning point for the investigation. Each "inquiry" group made its tape in less than

29 James D. Gammel, graduate student, Faculty of Educational Development, and Victor E. Edmonds, graduate student, Faculty of Curriculum and Foundations, both of The Ohio State University.
ten minutes (a little over nine minutes for each group). The "taught" groups completed their tapes in 22 minutes and 32 minutes respectively. The judges could not distinguish measurable differences (focus, contrast and smooth camera movements) in quality among the tapes.

Since the results of this action-research favored the "inquiry" group self-instruction, the investigator reexamined the typescripts of those sessions. The questions asked by the students in those sessions were felt to be more representative of attempts to solve a real problem. The investigator remembers that many questions never were verbalized. The immediate results of flipping a switch, turning a lens or doing some other manipulations probably answered a question before it could be verbalized. It was assumed, however, that the verbal record did span a significant spectrum of inquiry.

Most questions asked by the self-taught groups were found to be convergent in nature and sought new information. The investigator was tempted to hypothesize and classify the motives underlying each question. The investigator remembered though that the tendency and practice of inferring too much with too little information was a major factor used by the investigator in rejecting other analytical instruments and systems.

Since many questions did seek new information, the investigator felt that a category termed "seeking" was
appropriate to classify certain kinds of questions. Another type of question seemed to be of a verifying nature. These questions seemed to test some condition, quality or state. Consequently, "verifying" was tentatively accepted as a class. Both "verifying" questions and "seeking" questions appeared to have rather clear, descriptive characteristics.

Other questions, however, seemed more complex and seemed motivated by higher order cognitive needs. It was difficult to classify these questions without inferring the cognitive intent. The investigator felt that if a question consisted of an expressed cause-effect sequence— that is, a two-part question— it could be designated as an "inferring" question. In a sense an inferring question sought new information (seeking), but the more complex nature of the question was deemed a higher cognitive order. Inferring questions tended to put together a condition with an explanation. This type of question was similar to Suchman's "Experimentation" and "Necessity" categories as well as to Bloom's "Application" and "Synthesis." The investigator noted that inferring questions were both convergent and divergent but did not add that discrimination since that refinement depended on the analyst's interpretation.

A final descriptive category seemed necessary to code a few questions and many statements made by students. For lack of a better name the term "concluding" was selected. Not many questions were "conclusive" in nature, but many
The investigator believed that to construct a system that would be useful in analyzing student questions, a final category was necessary and logical. Verbal interaction usually concludes when a problem is solved or a bell rings to interrupt an interaction. When a problem is solved, the "concluding" verbal remark signals the end of a student's verbal involvement. When a problem session is interrupted "concluding" statements and questions often occur early in subsequent sessions. These "concluding" comments would then be viewed as the verbal capstone ending student inquiry into one problem.

The investigator found that these four descriptive categories were quite useful in classifying the questions students asked during problem solving. Almost every question was cataloged by the system. Those few questions that could not be labeled were classified by the investigator as "foil" questions, meaning that the analyst could not easily determine the cognitive intention of the probe. For the most part, questions of this designation were verbalizations that were followed by laughter or silence. To suggest that these questions emanated from the class "comedian" would be an inference inconsistent with the investigator's purpose in this study. Therefore, the "foil" category was used to capture any unclassifiable question. The category was seldom used.
The five categories developed by the investigator are defined as follows:

**Verifying (V)**

A question explicitly phrased to attest, reaffirm or verify some (usually one) fact, event, condition, characteristic, etc. that has been previously established.

**Seeking (S)**

A probing question explicitly phrased to procure one bit of information that has not yet been established.

**Inferring (I)**

A question of implication that specifically ties together two or more facts, events, conditions, characteristics, etc. as a hypothesis to explain any part of the problem.

**Concluding (C)**

An explicit question or statement of closure phrased as an explanation or an obvious summary of the student's verbal involvement with the problem.

**Foil (F)**

A question that does not readily fit into any of the other categories. The intent of a foil question can only be inferred. The student would know; the analyst would guess. Often a foil may interrupt interaction.

Using this five category system, the investigator was able to code all student questions that had been recorded during inquiry sessions which were conducted by the investigator. The system also was found useful in coding the questions that were recorded in several sessions conducted by Suchman. The investigator consistently coded each
question the same way. In comparing the first and second coding efforts of six typescripts, the investigator found that he coded only eleven questions differently. These findings are reported in Chapter IV.

At this point in the study the validity of the instrument had only been assumed. In conversations with colleagues it was suggested that problem solving sessions should be conducted by persons other than Suchman and the investigator and that these sessions should be taped. The investigator decided that if the questions generated in these sessions could be consistently coded, the validity and reliability of the instrument would be enhanced.

Two inquiry sessions were conducted by a former student of the investigator who was familiar with the teaching techniques normally effective in conducting inquiry sessions.\(^3\) These sessions were taped by him; the investigator was not present. An additional tape containing portions of two inquiry sessions conducted by teachers unknown to the investigator was borrowed.\(^4\) This tape had been recorded at an autumn inquiry workshop held in Illinois in 1968.

\(^3\) Donald R. Teffner, student teacher in social studies, Norton Middle School, Columbus, Ohio.

\(^4\) The tape is the property of Paul Farmer, Reading Coordinator, Evansville School Corporation, Evansville, Indiana.
These tapes were transcribed and then coded by the investigator. He was able to classify all questions by using the categories of the instrument, and the consistency of coding on a second trial remained high. Although these simple tests did not establish the validity and reliability of the instrument, the results were interpreted as a favorable indication that the instrument did measure what it was designed to measure.

The results of another application of the instrument were interpreted as favorable indicators pointing toward validity and reliability. It had been suggested that the category system might be useful to record the classroom behavior of a student or a teacher. After some deliberation the investigator expanded the system to include the factors of verbal attitude and verbal involvement. This meant adding three subcategories which seemed to accurately describe the spectrum of possible verbal participation or lack of it. These categories are defined as follows:

**Challenge** (-)

A question or statement phrased to attack the validity of another's finding, logic, inference, etc.

**Agreement** (+)

A question or statement phrased to accept all or part of another's finding, logic, inference, etc.

**Neutrality** (0)

Remaining verbally uncommitted to challenge, to agree, or even to participate in the discourse.
The investigator then used the five categories of the question classification instrument with the three attitudinal designations. He found that these categories constituted a comprehensive system which was quite easy to use in recording the verbal behaviors of one person in a classroom. The instrument, however, had not been designed as an observational system. Finding that the instrument was useful for observing and recording on-the-spot classroom behaviors was interpreted nevertheless as an indication of nominal validity.

The Training Program

To the developer the instrument seemed relatively appropriate and rather simple to use. Encouraged by Galloway\(^32\) and others, the investigator sought to demonstrate a broader, less biased applicability. At Galloway's suggestion a training program was developed and outlined and four potential analysts recruited.\(^33\) Three training sessions of about an hour each were conducted. A brief description of each session follows.

\(^32\)Charles M. Galloway, Professor, Faculty of Curriculum and Foundations, College of Education, The Ohio State University, and academic advisor to the investigator.

\(^33\)The four persons to be trained were graduate students during spring quarter of 1971 in the College of Education, The Ohio State University: Victor E. Edmonds (Curriculum and Foundations), James D. Gammel (Educational Development), Jane Gammel (Early and Middle Childhood Education), and Judith Morris (Curriculum and Foundations).
First Session

Observational systems were discussed generally. Their inadequacies for coding student questions were emphasized. A five-minute tape containing a portion of an inquiry session was played to illustrate these problems. The investigator related his reasons for developing his system and described the rationale for it. Each analyst was given a copy of the instrument and the five categories were discussed. The five-minute tape was replayed.

Second Session

The category system was reviewed and discussed again. The analysts were challenged to think of questions that could not be coded by the question classes. Discussion followed. A typescript of an inquiry session was distributed and coded by the members of the group. The codings were discussed.

Third Session

In this final session a two-page typescript was distributed and coded independently. The separate codings were then compared and discussed. A second typescript was similarly treated. The film, "The Ice Cubes,"\(^{34}\) was shown

and the story, "Joe: A Case Study," was distributed. These were discussed since each would be used as a focus for inquiry sessions the analysts would code.

**The Source of Data**

To test the reliability of the instrument and to compare the coding for each question by each analyst the investigator decided to conduct and record inquiry sessions with his undergraduate education classes. His reasons for choosing this sample population with which to conduct final inquiry sessions were: (1) He felt the students possessed a general mental maturity not necessarily found in younger children. From this he inferred that the students would ask questions that would span a relatively wide cognitive range. (2) The students in these classes were majoring in a wide range of academic disciplines. (3) The students had been assigned to these two classes in a random manner. They had not been grouped in any way which lead the investigator to assume the membership of his classes represented the larger university community. (4) It

---


37Art, English, foreign languages, math, music, physical education, speech and hearing therapy, science and social studies.
was convenient for the investigator to use these two groups. He met with these two classes for two hours per day, four days per week.

Prior to introducing inquiry development programs, the classes had been exposed to some selected materials from *The Humanities Curriculum Project*. This program was demonstrated by conducting a short discussion session and then having the students examine the items of evidence in one of the units. It should be noted that the rationale for this program and the techniques used to teach it are not unlike those of inquiry-based programs.

Before conducting and taping the final inquiry sessions (for this study) the instructor discussed inquiry development programs in general. He attempted to illustrate basic inquiry by demonstrating the *Inquiry Box*, a device that could be used without prerequisite subject matter knowledge. Following this introduction the final inquiry sessions were conducted.

One final consideration should be noted. Based on his previous experiences with inquiry, the investigator believed that the styles of student problem solving—that is, the

---


number and kinds of questions asked—might be influenced by either the nature of the problem or the nature of the problem solution. Most concede that an explanation for a discrepant event in physical science is more testable and verifiable than is an explanation for a behavior in the social sciences. Consequently the investigator chose a problem from the physical sciences ("The Ice Cubes") and one from the social or behavioral sciences ("Joe: A Case Study").

The Inquiry Sessions

For the first problem solving session the film "The Ice Cubes" was shown. In that film an ice cube was dropped into two glasses containing a colorless liquid. The ice cube in the first glass sank momentarily and then floated to the surface of the liquid. In the second glass the ice cube sank to the bottom of the glass and remained there. After a few minutes an experimenter used a spoon to push the floating cube to the bottom of its glass. This cube immediately floated to the surface again. He then pushed the "sunken" cube to the very bottom of its glass. When released, this cube then rose about halfway up in its glass. The same procedure was followed a few minutes later. The floating cube behaved normally and floated at the surface. The "halfway sunken" cube, however, behaved strangely. After it had been pushed to the bottom of the glass the cube
floated nearly to the surface of the liquid.

As expected many questions were asked in the inquiry session that followed the showing of this film. This session was taped and transcribed.

For the second inquiry session a behavior problem served as the focus for inquiry. The one-page story, "Joe: A Case Study," was given to the class and read aloud. In this story two specific instances of late behavior were described specifically and other instances were implied. (See Appendix A.) The investigator challenged the students to explain this chronic late behavior. Of course, many questions were asked to secure information not presented. This session was also taped and transcribed.

The typescripts were then duplicated and given to the analysts. The analysts coded the "Ice Cubes" script first. The record of this coding was tabulated in rough form and discussed briefly. Two days later the analysts independently coded the final typescript, "Joe."

The record of these analytic coding efforts as well as an analysis of the data are found in Chapter IV.
CHAPTER IV
ANALYSIS OF THE DATA

Introduction

This study dealt with student questions. A major consideration was to determine whether student questions were a useful source of data for analyzing student inquiry during problem solving. The investigator developed an analytic instrument and instructed others in its use. Each of the trained analysts and the investigator then independently coded typescripts of two problem solving sessions conducted by the investigator with two undergraduate classes at The Ohio State University. The film, "The Ice Cubes," which portrayed a discrepant event in physical science, was used as the focus for the first problem solving session. The story, "Joe: A Case Study," which presented a behavioral problem, was used for the second session.

The results of using the analytic instrument to code the questions the students asked in these two sessions are reported in this chapter. The first analysis, dealing with the problem in physical science, is referred to as the "Ice Cubes." The second analysis, which coded the behavioral session, is called simply "Joe." The statistical treatment of the coding data is rather descriptive and is expressed in
tabular form as simple ratios of coding agreement among analysts. An interpretive commentary follows each analysis. A brief summary concludes the chapter.

The Ice Cubes

The typscript of the "Ice Cubes" inquiry session contained 46 student initiated entries. Most of these were questions. The analysts were asked to code each entry according to the category definitions specified in the analytic instrument. The coding symbols used are the first letters of the category names. The four trained analysts have been designated by the numerals 1, 2, 3, and 4. The investigator has been designated as 5. The composite item-by-item coding is presented in Table 1. The items for which there was not complete coding agreement have been noted with an asterisk (*).

It should be noted that all coders agreed on 36 of the 46 items. This yielded an interanalyst reliability ratio of .78. Four of the five analysts agreed to an additional 4 items which produced an agreement ratio of .87. And three of the five concurred on another 3 items which yielded a majority analyst reliability of .96.

1V=Verifying; S=Seeking; I=Inferring; C=Concluding; and F=Foil. See page 65 for a complete description of the analytic instrument.

2James D. Gammel (1), Jane Gammel (2), Judith Morris (3) and Victor Edmonds (4), all graduate students in the College of Education, The Ohio State University, Spring quarter 1971.
TABLE 1. ITEM-BY-ITEM ANALYSES OF THE "ICE CUBE" TYPESCRIPT

<table>
<thead>
<tr>
<th>Item No.</th>
<th>Analysts</th>
<th>Dif. Code</th>
<th>Item 1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Dif.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>S S S S S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S S S S S</td>
</tr>
<tr>
<td>2</td>
<td>S S S S S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>I I I I I</td>
</tr>
<tr>
<td>3</td>
<td>S S S S S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C C C C C</td>
</tr>
<tr>
<td>4</td>
<td>S S S S S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S S S S S *</td>
</tr>
<tr>
<td>5</td>
<td>C C C C C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V V V V V</td>
</tr>
<tr>
<td>6</td>
<td>C C C C C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V V V V V</td>
</tr>
<tr>
<td>7</td>
<td>S S S S S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V V V V V</td>
</tr>
<tr>
<td>8</td>
<td>S S S S S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S S S S S</td>
</tr>
<tr>
<td>9</td>
<td>S S S S S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V V V V V</td>
</tr>
<tr>
<td>10</td>
<td>S S S S S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V V V V V</td>
</tr>
<tr>
<td>11</td>
<td>I S I S I</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>I I I I I</td>
</tr>
<tr>
<td>12</td>
<td>I C I I I</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S S S S S</td>
</tr>
<tr>
<td>13</td>
<td>C C C C C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S S S S S</td>
</tr>
<tr>
<td>14</td>
<td>V V V V V</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F F F F F</td>
</tr>
<tr>
<td>15</td>
<td>I I I F F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S S S S S</td>
</tr>
<tr>
<td>16</td>
<td>S S S S S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F F F F F</td>
</tr>
<tr>
<td>17</td>
<td>S S S S S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F F F F F</td>
</tr>
<tr>
<td>18</td>
<td>C F F - C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S S S I I</td>
</tr>
<tr>
<td>19</td>
<td>S S S S S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S S S S S</td>
</tr>
<tr>
<td>20</td>
<td>S S S I S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>I C I C C</td>
</tr>
<tr>
<td>21</td>
<td>S S S S S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>V V V V V</td>
</tr>
<tr>
<td>22</td>
<td>V V V V V</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F F F F F</td>
</tr>
<tr>
<td>23</td>
<td>S S S S I</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C C C C C</td>
</tr>
</tbody>
</table>

A superficial examination of the complete coding record revealed that the lack of coding consistency was evenly spread among all of the analysts. Even the investigator's classification, supposedly the standard for the instrument, stood out as an example of a coding dissonance. Four of the analysts viewed Item 23 as a Seeking question, but the investigator considered it an inference.

To begin the systematic analysis of the instrument's reliability in this instance the number of items coded for
each category for each analyst was tallied. The category totals for each analyst are presented in Table 2.

TABLE 2. COMPOSITE ANALYSIS OF CODING BY CATEGORY (THE "ICE CUBES")

<table>
<thead>
<tr>
<th>Categories</th>
<th>Analysts</th>
<th>V</th>
<th>S</th>
<th>I</th>
<th>C</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7</td>
<td>24</td>
<td>5</td>
<td>6</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>24</td>
<td>3</td>
<td>8</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>23</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>22</td>
<td>4</td>
<td>7</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>22</td>
<td>5</td>
<td>7</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

One analyst (No. 4) did not code one item, which means his total of entries classified was 45 whereas the other analysts totaled 46 items. At first glance this table of composite tallies appears to show a fair agreement among analysts. Note the small ranges for each category. The range never spanned more than 4 numbers, which is deceiving. The overall agreement among all analysts for this session was only .78.

When agreements between analysts are compared, however, the ratios improve considerably. Table 3 presents the ratios of agreement between analysts. The ratio of interanalyst agreement was calculated by using the following equation:
This mathematical treatment of data was selected since the ratios derived were representative of the actual, observed item agreements between analysts. The broad categorical agreements between analysts, which are not restricted by actual item agreements, yield higher ratios of agreement which are deceiving and unrealistic.

**TABLE 3. RATIOS OF INTERANALYST AGREEMENT: ALL CATEGORIES ("THE ICE CUBES")**

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>.91</td>
<td>.85</td>
<td>.89</td>
<td>.89</td>
<td>X</td>
</tr>
<tr>
<td>4</td>
<td>.91</td>
<td>.89</td>
<td>.85</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.93</td>
<td>.87</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.85</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Range .85 - .93  
Mean .88  
Median .89

When the higher agreement ratios between analysts (.85 - .93) are compared to the overall ratio of agreement (.78) a distinct disparity is found. It seems as if the "whole is much less than the sum of its parts." This
disparity could be explained, however, by again examining Table 1. Coding differences were slight and not consistent. This overlapping of error seemed to maximize the agreement between two analysts but minimized agreement between all analysts. This tendency for greater agreement can be further illustrated by comparing the number of items agreed upon by four of the five coders (40/46)—that is, an agreement of .87—to the mean of all agreements between two analysts (.88). Very simply stated the investigator found that the majority of coding errors were few between any two analysts but rather abundant when all analysts were compared.

The reliability of interanalyst agreement within each coding category was examined next. The same equation was used to derive the ratios. The interanalyst consistencies for the Verifying and Seeking categories are presented in Tables 4 and 5.

| TABLE 4. RATIOS OF INTERANALYST AGREEMENT: VERIFYING CATEGORY |
|-----------------|-----|-----|-----|-----|-----|
|                 | 1   | 2   | 3   | 4   | 5   |
| 5               | 1.00| .92 | .93 | 1.00| X   |
| 4               | 1.00| .92 | .93 | X   |
| 3               | .93 | .86 | X   |
| 2               | .92 | X   |
| 1               | X   |
| Range           | .86 | 1.00|
| Mean            | .94 |
| Median          | .93 |

| TABLE 5. RATIOS OF INTERANALYST AGREEMENT: SEEKING CATEGORY |
|-----------------|-----|-----|-----|-----|-----|
|                 | 1   | 2   | 3   | 4   | 5   |
| 5               | .95 | .95 | .93 | .91 | X   |
| 4               | .91 | .95 | .93 | X   |
| 3               | 1.00| .94 | X   |
| 2               | .89 | X   |
| 1               | X   |
| Range           | .89 | 1.00|
| Mean            | .94 |
| Median          | .94 |
The coding consistency for all pairs of analysts for both the Verifying and Seeking categories was rather high. From this the investigator inferred that both these categories possessed specifications that were discriminating, functional and operational.

The interanalyst ratios of agreement for the Inferring category, however, were not very good. These ratios are presented in Table 6.

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>.60</td>
<td>.50</td>
<td>.60</td>
<td>.67</td>
</tr>
<tr>
<td>4</td>
<td>.44</td>
<td>.57</td>
<td>.44</td>
<td>X</td>
</tr>
<tr>
<td>3</td>
<td>.80</td>
<td>.25</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.50</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Range .25 - .80
Mean .44

From the range of ratios alone—.25 to .80—it was inferred that the coding of Inference questions was the most difficult task for the analysts. Each coder was asked about his interpretation of an Inference. There were subtle differences of what constituted an Inference. Each interpretation, though different, was well grounded in logic.

Ten times, at least one coder disagreed on the classification of an item (Table 1). For eight of these item
disagreements at least one coder used the Inference category. This was interpreted to mean that the specifications of this category were ambiguous. The discrimination power of this category was not equal to the discrimination of the Verifying and Seeking classes. The mean value of agreement between paired analysts was .44. This can hardly be interpreted as being significantly reliable. A random agreement would be expected to approach .50.

Tables 7 and 8 present the interanalyst reliabilities for the Concluding and Foil categories. For both of these classes the interanalyst agreement ratios approached the investigator's level of expectation.

TABLE 7. RATIOS OF INTERANALYST AGREEMENT: CONCLUDING CATEGORY

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>.92</td>
<td>.80</td>
<td>.83</td>
<td>.86</td>
<td>X</td>
<td>5</td>
<td>.89</td>
<td>.80</td>
<td>.80</td>
</tr>
<tr>
<td>4</td>
<td>.77</td>
<td>.93</td>
<td>.83</td>
<td>X</td>
<td>4</td>
<td>.89</td>
<td>.80</td>
<td>X</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>.91</td>
<td>.77</td>
<td>X</td>
<td>3</td>
<td>.89</td>
<td>1.00</td>
<td>X</td>
<td>2</td>
<td>.89</td>
</tr>
<tr>
<td>2</td>
<td>.72</td>
<td>X</td>
<td>2</td>
<td>.89</td>
<td>X</td>
<td>1</td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Range .72 - .93
Mean .83
Median .83

TABLE 8. RATIOS OF INTERANALYST AGREEMENT: FOIL CATEGORY

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>.89</td>
<td>.80</td>
<td>.80</td>
<td>1.00</td>
<td>X</td>
<td>5</td>
<td>.89</td>
<td>.80</td>
<td>.80</td>
</tr>
<tr>
<td>4</td>
<td>.89</td>
<td>.80</td>
<td>X</td>
<td>3</td>
<td>.89</td>
<td>1.00</td>
<td>X</td>
<td>2</td>
<td>.89</td>
</tr>
<tr>
<td>3</td>
<td>.89</td>
<td>1.00</td>
<td>X</td>
<td>1</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Range .80 - 1.00
Mean .88
Median .89

At first the overall consistency of coding by all analysts (.78) was regarded as a disappointment. A much higher ratio of agreement had been expected. After
appraising the coding agreements for the Inference category, however, the overall .78 ratio was gratefully accepted.

Although the overall reliability did not meet the standards of the investigator, he found that the instrument was indeed useful in drawing inferences about patterns of inquiry. Many educators expect the problem solving process to follow a logical sequence—that is, set the problem, verify the facts, seek new information, infer or hypothesize, and conclude. The process of inquiry with the "Ice Cubes," however, did not follow this textbook scientific method. In general Verifying questions filled the middle of the inquiry with one verification nearly ending the session. Conclusions, normally expected to terminate a session, were interjected quite early (items 4 and 5) as well as during the middle and at the end of the session. Questions that sought bits of new information (Seeking) dominated the data. Virtually one-half of the questions asked were coded as Seeking questions. These few observations of question patterns will be discussed further after the data for "Joe" have been presented.

Joe

The typescript for this inquiry (behavioral science) contained 60 student initiated entries. The same analysts were asked to interpret each item and code it according to the category definitions specified in the instrument. Each
analyst was reminded to code every entry. The composite item-by-item coding is presented in Table 9.

TABLE 9. ITEM-BY-ITEM ANALYSES OF THE "JOE" TYPESCRIPT

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>2</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>3</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>4</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>5</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>6</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>7</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>8</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>9</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>10</td>
<td>S</td>
<td>V</td>
<td>V</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>11</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>12</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>13</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>14</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>15</td>
<td>S</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
</tr>
<tr>
<td>16</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>17</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>18</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>19</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>20</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>21</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>22</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>23</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>24</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>25</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>26</td>
<td>S</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>C</td>
</tr>
<tr>
<td>27</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>28</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>29</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
</tr>
<tr>
<td>30</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
</tbody>
</table>

A quick appraisal of the composite coding record indicated that Seeking questions again dominated. Very few items were coded as Inferences, Foils and Verifications.
All coders agreed on 56 of the 60 entries which produced an overall coding consistency of .93 (compared to .78 with "Ice Cubes"). The agreement ratio improved to .97 when agreement of only four of the five coders was considered. And a perfect 1.00 correlation was obtained if only a majority agreement was considered.

The investigator certainly welcomed the improvement in overall consistency in coding. Three factors probably contributed to this improvement: (1) the typescript contained more items; (2) the types of questions may have been easier to classify; and (3) the coders had become more skilled with the instrument.

The category coding totals for each analyst are detailed in Table 10.

TABLE 10. COMPOSITE ANALYSIS OF CODING BY CATEGORY ("JOE")

<table>
<thead>
<tr>
<th>Analysts</th>
<th>V</th>
<th>S</th>
<th>I</th>
<th>C</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>46</td>
<td>2</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>44</td>
<td>3</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>44</td>
<td>3</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>46</td>
<td>1</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>44</td>
<td>2</td>
<td>8</td>
<td>2</td>
</tr>
</tbody>
</table>
The observation that Seeking questions seemed to dominate the coding record was more than verified. The investigator expected this category to be the statistical mode if the coding classes were ranked. However, the overwhelming dominance of Seeking questions was not anticipated. Since the conditions for both inquiry sessions were essentially the same (same students, instructor, and time period) the investigator hypothesized that the problem type—that is, behavioral or physical—may have influenced the percentage of Seeking questions asked. A similar relationship perhaps would be found with other categories as well. The possibilities for correlations, however, were not investigated nor considered findings in this study.

Interanalyst reliability ratios were calculated as before and appear in Table 11.

| TABLE 11. RATIOS OF INTERANALYST AGREEMENT: ALL CATEGORIES ("JOE") |
|-------------------------|----------------|----------------|----------------|----------------|----------------|
| 1 | 2 | 3 | 4 | 5 |
| 5 | .95 | .98 | .98 | .97 | x |
| 4 | .97 | .95 | .95 | x |
| 3 | .98 | 1.00 | x |
| 2 | .97 | x |
| 1 | x |

Range .95 - 1.00
Mean .97
Median .97
These ratios of coding consistency between analysts were far better than the agreement ratios for the first coding effort. With the "Ice Cubes" the agreement ratios ranged from .85 to .93. For "Joe" the reliability range was .95 to 1.00. When the mean of agreement between analysts was compared with the overall ratio of agreement for each session a similar improvement was found.

<table>
<thead>
<tr>
<th>Overall Ratio of Agreement</th>
<th>Ice Cubes</th>
<th>Joe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Ratio of Agreement</td>
<td>.78</td>
<td>.93</td>
</tr>
<tr>
<td>between Analysts</td>
<td>.89</td>
<td>.97</td>
</tr>
<tr>
<td></td>
<td>.95+</td>
<td></td>
</tr>
</tbody>
</table>

The investigator believed that further statistical comparisons would be superfluous and of questionable merit. It had been demonstrated that both the overall coding consistency and the interanalyst reliability ratios for "Joe" were much higher than were the ratios for the "Ice Cubes."

The interanalyst reliability for the Verifying and Seeking categories are presented in Tables 12 and 13.

**TABLE 12. RATIOS OF INTERANALYST AGREEMENT: VERIFYING CATEGORY**

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>.86</td>
<td>1.00</td>
<td>1.00</td>
<td>.86</td>
</tr>
<tr>
<td>4</td>
<td>1.00</td>
<td>.86</td>
<td>.86</td>
<td>X</td>
</tr>
<tr>
<td>3</td>
<td>.86</td>
<td>1.00</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.86</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Range .86 - 1.00
Mean .92

**TABLE 13. RATIOS OF INTERANALYST AGREEMENT: SEEKING CATEGORY**

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>.98</td>
<td>1.00</td>
<td>1.00</td>
<td>.98</td>
</tr>
<tr>
<td>4</td>
<td>.98</td>
<td>.98</td>
<td>.98</td>
<td>X</td>
</tr>
<tr>
<td>3</td>
<td>.98</td>
<td>1.00</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.96</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Range .96 - 1.00
Mean .98
Although the ratios of reliability are relatively high they did not correspond with the near identical relationship of the Verifying and Seeking categories in the "Ice Cubes" (Tables 3 and 4). The investigator considered these reliability ratios to be positive indicators of the discrimination power of both categories.

The investigator then calculated the agreement ratios among analysts for the Inference category. In one comparison the observers matched perfectly (2 and 3); in another comparison, however, they did not match at all (1 and 4). Table 14 presents the findings.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>.50</td>
<td>.80</td>
<td>.80</td>
<td>.67</td>
<td>X</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>.50</td>
<td>.50</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.80</td>
<td>1.00</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.80</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Range 0 - 1.00
Mean .56
Median .67

Obviously coding reliability was not characteristic of this category. Low agreement ratios were also reported in the analysis of the Inference class of questions with the "Ice Cubes."
At first the investigator thought that the small number of Inference questions might have contributed significantly to the low incidence of agreement. One analyst selected only 1 entry as an Inference; two analysts designated 2 items as Inferences; and one analyst selected 3 entries as Inferences. The chances that coders would agree exactly on which items of 60 entries were Inferences might have been small. However, that logic was merely mathematical and was rejected as inappropriate for the study. The Foil category was used only twice by each analyst in coding the 60 items, and in all cases each analyst coded the same two items as Foils. This yielded a perfect agreement. Therefore, the investigator concluded that the coding of Inferences could not be done reliably and consistently by analysts using this instrument and trained as they were. Either the category specifications were ambiguous and lacked discrimination power or the training program was inadequate. Both the category description and the training program should be re-evaluated before a similar study is undertaken.

To complete the presentation of data for the "Joe" session the ratios of agreement for the Concluding and Foil categories are detailed in Tables 15 and 16.
Again very high ratios of consistency are in evidence. The investigator believed that these two categories, Concluding and Foil, were viable and functional classifications which were useful in designating types of student verbal interaction during problem solving.

**Summary**

In this study the validity and reliability of the analytical instrument were not firmly established. Based on the findings of this study the investigator believed the instrument should be revised before its validity and reliability are remeasured. Four of the five categories appeared to be appropriate and functional. The power of discrimination of the Inference category, however, was questionable, and it is this category that should be revised.

Although the instrument was far from perfect the investigator did find it to be useful for classifying the

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>.93</td>
<td>.93</td>
<td>.93</td>
<td>1.00</td>
<td>X</td>
</tr>
<tr>
<td>4</td>
<td>.93</td>
<td>.93</td>
<td>.93</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.00</td>
<td>1.00</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Range .93 - 1.00
Mean .96
types of student questions and statements that were verbalized during a problem solving interaction. Since no similar analytic system was found, the data derived from this study could not be compared to a standard.

The investigator assumed the instrument possessed a certain reliability that discriminated among the types of student questions and statements. Although the scope of this study had been narrowed to develop an analytic instrument and test it, some brief observations were made.

Considering his coding report the investigator noted that Concluding statements, which were normally expected at the end of an interaction, often appeared early in the session. This was not expected. Questions that sought new bits of information (Seeking questions) clearly dominated both inquiry sessions (physical science and behavioral science). Seeking questions always opened a session; this was expected. Seeking questions also came in uninterrupted batches early in the sessions. When the pattern of Seeking questions was broken, Inferences and Conclusions began to appear in clusters. Verifying questions only twice followed a conclusion; at all the other times they followed a Seeking question. The entries that were coded as Foils generally released a tension or broke a line of reasoning. Most Foils (and there were only a few) occurred during the last third of the interaction.
The above mentioned observations and comments were findings. Further expansion, however, would have been derived inferences of questionable value and inappropriate for a chapter reporting the data. Some of these implications are discussed in the final chapter.
CHAPTER V
SUMMARY, FINDINGS, INFERENCES,
IMPLICATIONS AND RECOMMENDATIONS

The major task in this investigation was to develop an instrument that could be used to discriminate and classify the questions students asked during problem solving. Essentially this study described the development of that instrument and reported what happened when the instrument was used. This final chapter contains a summary of the study, reports the findings, discusses the inferences that were drawn, suggests some implications that seemed appropriate, and lists some recommendations for further study.

Summary

This investigation dealt with student questions. For the most part student questions have been ignored in educational research. The few studies that have dealt at all with student questions revealed that students ask few questions in normal classroom interaction. Studies of classroom questioning almost always dealt with teacher questions and usually demonstrated the pedagogical importance of asking "better" questions. This study, however, focused only on student questions.
To obtain sufficient data (student questions) for analysis the investigator examined the verbal interaction in problem solving situations—that is, inquiry sessions in which student questions were central to and provided the momentum for the discussion. Since these rather contrived situations did not represent general classroom interaction the scope of generalization was severely limited.

The investigator tried to classify student questions by using the already proven category systems that were well known in education. All these observational systems and cognitive systems were rejected as being unsuitable and inappropriate for analyzing student questions.

By closely examining the student entries in typescripts of several inquiry sessions the investigator inductively derived a suitable classification system. The five categories—Verifying, Seeking, Inferring, Concluding, and Foil—were then validated by classifying all student initiated comments from many problem solving interactions.

This system was then taught to four analysts. The trained analysts were asked to code two typescripts of inquiry sessions conducted by the investigator with students at The Ohio State University. One inquiry session focused on a discrepant event in physical science; the other dealt with a problem of human behavior. The mean coding agreement among analysts for the physical science was .88 and for the behavioral session it was .97. In general the consistency
of coding among analysts for Verifying, Seeking, Concluding, and Foil entries was acceptable. The interanalyst reliability for coding Inferring items, however, was erratic, ranging from no agreement to perfect agreement between two analysts.

The way questions were coded using this system could not be adequately compared with another coding method. No similar classification system existed.

Findings

Broadly considered the instrument developed for this study appeared to be useful. The investigator as well as trained analysts were able to classify and discriminate among the questions students asked during the problem solving sessions that were examined. The complete analytic system seemed to be all inclusive and seemed to possess mutually exclusive categories. The validity and reliability of coding was deemed adequate for four of the five categories (Verifying, Seeking, Concluding, and Foil). The discrimination of Inferring questions was less than adequate. The coding of an Inference depended a great deal upon the analyst's interpretation of the item in its context. Even though all classifications related to context the investigator found that the other categories were more specific and descriptive.
The following general observations applied to all analyses, no matter whose coding record was used. In all sessions Seeking questions initiated the inquiry and were the dominant question type. Both of these findings were anticipated.

Questions of a probing nature usually initiated and always dominated the inquiry sessions conducted by Suchman. He found that autonomous learning depended upon the learner's freedom to inquire which the investigator has interpreted to mean the opportunities for Seeking information.

In this study the investigator expected Concluding statements to occur toward the end of a session. Often, however, conclusions were drawn early and midway through the inquiry as well as at the end of the session.

The number of Inferring questions was always fewer than the number of Conclusions. This was not expected. Suchman and others have found that a conclusion usually couples more than one inference into an explanation.

Another finding in this study was that more Inferring questions were asked during the sessions that dealt with discrepant events in physical science than were asked during the sessions that focused on behavioral problems.

Verifying questions most often followed Seeking questions. Only once did a Verifying question follow a Conclusion.
Foils were few and usually occurred near the end of a session. This category type was also the most easily identifiable, even out of context.

The investigator as well as the analysts believed that a basic pattern of questioning would emerge from the codings. Very simply stated the expected strategy would have begun with a few Verifying questions that established the problem set. A large number of Seeking questions were expected next. These, it was believed, would have been followed by one or two Inferring questions. After a very few Seeking or Verifying questions a Conclusion was anticipated. This basic syllogism, however, did not emerge. In fact, precise descriptive strategies of student questioning were not actually documented in this limited study.

Inferences

Traditionally final chapters present conclusions. Because of the nature of this study and the results obtained the investigator believed that drawing inferences would be more appropriate than drawing conclusions.

The fact that no precise strategies of student inquiry emerged from the analysis of the coded sessions was an initial disappointment. The investigator inferred, however, that the students did inquire and solve problems in different ways even though these ways were not precisely described when this analytic system was used.
If the validity and reliability of the analytic system can be improved certain descriptive strategies of student questioning may become more evident. On the basis of this limited study the investigator has inferred that student questions are a fruitful source of data from which to analyze patterns of student learning during problem solving activities. Primarily this inference has been drawn from the investigator's analysis of the proportions of each question type and the sequences of question clustering that were obvious.

Another inference that has been drawn is that an instrument can be developed that is useful in cataloging the elements of student interaction during problem solving. Broadly considered the five descriptive categories of the instrument seemed relatively appropriate. The investigator believes that the power of discrimination of the instrument would not be improved appreciably if far more descriptive categories were added. A few modifications and refinements, however, are definitely in order.

The results of this limited investigation seem to indicate that descriptive research and systematic analyses of student interactions may prove to be more fruitful than the exclusive analyses of teacher behavior have been in the past. This inference is based upon the investigator's perception that a shift in educational emphasis is being made from teacher centrality to student autonomy.
The investigator also inferred that some students may have benefitted from the group interaction during the problem solving. Others probably did not. It was inferred from the items that were classified as Foils that at least some students' involvement in group problem solving had ended. It should be noted that other students entered and left the interaction nonverbally, but their behaviors were never recorded.

The investigator also inferred that those students who did verbally participate in the problem solving asked questions which were mainly convergent in nature. Perhaps the only verbal expressions of divergent thinking may have been coded as Foils. From this the investigator further inferred that a student's productive divergent thinking probably evolved from his having productive experiences with convergent thinking.

Since the student interaction examined in this study was taken from records of contrived situations and since only the verbal dimension of the interaction was analyzed, the inferences that have been drawn are very tentative in any broader context.

**Implications**

Although this study was limited in scope some implications for the larger field of education were drawn. The investigator believed that the instrument developed for this
study was useful in degree. With it an observer or analyst is forced to focus on the learners in the classroom. This shift of focus may assist researchers in finding some clues about student learning during problem solving.

The logic of instructional agendas—that is, teaching sequences, textbook sequences, and lock-step curricular sequences—may have to be reappraised if learning through inquiring continues to be an educational objective. The teaching logic of these ordered presentations may be somewhat dissonant to the learning logic of students engaged in problem solving. To be more specific, many teaching methodologies may have to be modified or even replaced. In this teachers should acquire practical methods which facilitate learning rather than methods which control classroom behavior and learning.

Another implication drawn from this study relates to divergent thinking. Presently one emphasis in education is to encourage students to practice divergent thinking. Divergence has been associated with creativity and at present seems to be more prized than convergence. The processes of education— inquiry skills, critical thinking, problem solving, autonomous learning and the like—are also being emphasized. These processes, however, are more convergent in nature than they are divergent. Factors of creativity in reasoning can apply to convergent thinking as well as to divergent thinking. Creativity in convergence, though, may
be less easily identifiable. The findings of this study clearly illustrated that as a tactic convergence far outstripped divergence in the questioning of students engaged in problem solving. From this the investigator determined that educators may be underestimating the value of convergent thinking or may be incorrect in coupling creativity only with divergent thinking.

The last major implication drawn from this study deals with observational and analytic systems. In general the well known, reliable, and well validated category systems used in educational research have cataloged teacher behaviors almost exclusively. The question analysis system developed for this study was devoted only to cataloging student verbal behaviors. In the mind of the investigator the system proved useful by organizing student data into a rough but meaningful reference frame. Further refinement of the instrument's classifications, especially the Inferring category, may yield a coding screen that will be useful for analyzing student learning behaviors. The refined instrument or another system then might be used by teachers to collect data which would provide the basis for extensive analysis of student verbal behaviors.

In any case the investigator believes that observational and analytic systems should be developed that focus on student behaviors as well as on teacher behaviors. Classroom interaction can be viewed from at least three vantage
points—from a teacher's position, from a student's position, or from a flexible interactive position.

The investigator believes that this study provided a narrow glimpse of how students learn in problem solving situations. When this fleeting glimpse widens to a broader look educational research may well readjust its instrumentation—focusing on students, teacher, and group interaction—to yield information that will be extremely useful in understanding the learning-teaching process.

Recommendations

On the basis of both this study and his previous experiences with inquiry development programs the investigator recommends the area of student questioning to educational researchers. Specifically this system should be refined or another should be developed that will be both useful in analyzing student questions and reliable in all of its category discriminations.

Student-teacher interactive styles and student problem solving patterns should be described and analyzed. Statistical comparisons of student achievement would be in order for this type of an investigation.

Pedagogical styles and methodologies could be compared according to the numbers and types of student initiated questions. These comparisons may reveal significant benefits in student learning accruing from certain pedagogical styles.
An investigation might be made to determine whether distinct patterns of student questioning are representative of certain grade levels or are related to specific curricular areas.

Another fruitful study might examine the nonverbal dimensions of student interaction with the verbal dimensions of student questioning. This aspect of communication has been overlooked far too long.

The final recommendation the investigator would offer deals with the development of analytic systems. The investigator has found that the systems he examined and the system he developed were objective in their intent. Realistically, however, there were ambiguities in category specifications and author biases were evident in most. A study might be made of several of these systems to determine to what degree interpretative subjectivity and personal inference are a part of objective instrumentation.
APPENDIX A

JOE: A CASE STUDY
When they finished bolting lunch the boys of Cabin A hurried to the stable. Pete, their counselor, arrived at 12:30, surveyed the group and asked, "Where's Joe?" As a group they shrugged. Vince said, "I don't know why he's always late. Can't we go without him?" A chorus of yeas seemed to express a unanimous feeling. Pete faced the boys, then glanced towards the mess hall. He slammed his fist into his hand and started to say..."Maybe we ought to..." Then in midsentence he changed his direction. "Mike, you're fast. Run up to the mess hall and see if Joe's still there. We'll wait just till you get back."

Mike took off at full stride. The boys mumbled and grumbled, but Pete didn't seem to hear. In less than five minutes Mike and Joe were trotting down the hill. Pete looked at his watch--12:45. The boys quickly mounted and struck off in a single file toward the western ridge. After about a half an hour of steady riding Pete reined up in a clearing. He tied his horse to a branch and hollered for everyone to do the same. He told Joe twice. The boys clustered around their counselor and he said they could explore the meadow and the stream that edged it. "Perhaps," he said, "you'll find something for the nature exhibit that would clinch first place for Cabin A." He warned them to pair off, stay in sight and be back by 2:30. He talked directly to Joe and asked him if he understood. Joe nodded. Again talking directly to Joe he reminded the group that the western slope caves were off limits.

The boys immediately scattered and began treasure hunting. The nature exhibit was very important since the competition for best cabin was very close. C Cabin had a 2 point lead over F Cabin. A Cabin was only 1 point behind F Cabin. First place in the nature exhibit was worth 5 points which would vault Pete's group into the top position with less than one week to go in camp.

By 2:20 most of the boys had returned to the horses. Tony's find, a deer skull with antlers still attached, seemed to evoke the most admiration. Pete blew a long blast on his whistle and George and Ben picked up their pace to rejoin the group. Pete surveyed the boys. "Where's Joe?" was his comment. Vince said, "Don't look at me, I haven't seen him." Ben suggested leaving without him to let him find his own way back. Pete looked at his watch--2:40. A Cabin had its swimming test at 3:30. Everyone might pass the life saving test--more points for the cabin. Pete looked around. No sign of Joe. Pete finally told Vince to take the boys back to the stable and to get ready for swimming. Pete would look for Joe and then be right along.

The boys quickly mounted and started back. Pete started off looking for Joe. He scoured the meadow, but there was no sign of the boy. At 3:15 he was about ready to go back for help when he saw Joe running toward him. Joe was scratched, shirtless but smiling broadly as he stumbled down the off-limits area of the western slope. "Where have you been?" Pete screamed. "Don't you know that's off limits?" Joe's smile faded only a bit as he unwrapped his shirt which bore his treasures. Pete looked at the contents and nodded. "Let's go, we're late enough already." The two mounted up and headed back to camp.

......inquiry in the behavioral sciences....Why is Joe always late?
APPENDIX B

BEHAVIORAL SCIENCE PROBLEM SOLVING SESSION WITH SELECTED EDUCATION 435 STUDENTS: CONDUCTED BY ALLAN WALDREN--"Joe: A Case Study"
S: Does he have a mother and father?
AW: He lives with his mother and two younger sisters.

S: Is his father dead or are they divorced?
AW: No one knows whether he dead or not, and they're not formally or legally divorced. The father hasn't been around for about 3 years.

S: Is he late for other things aside from just camp?
AW: Yes. Late to school, late for work, late for most things--about 80 per cent of the time he's late.

S: Is he a loner? You know, no friends?
AW: No, he has friends, but much of the time he is by himself.

S: What kind of grades does he get in school? What's his best subject?
AW: Ah, it's hard to pick a best subject and his grades are poor. Several failures.

S: Have they ever tested his hearing?
AW: Yes.

S: How did it check out?
AW: He's almost totally deaf in the right ear. Other's O.K.

S: How's his mother handle his lateness?
AW: What specifically do you mean?
S: You know, like if he's late for dinner. Does she ignore it?
AW: Usually she verbally reprimands him.

S: Nothing more than that? That's all?
AW: Yes, he's not usually late for his mother.

S: How old is he?
AW: Thirteen

S: Has he been checked to see if he's retarded?
AW: There's two IQ score on him.

S: What are they?
AW: First one was 87 and the other one 73.

S: How old was he for these tests?
AW: For the first one, 87, he was ...let's see 8. And for the other one he was 12.

S: So, in other words he's regressing, IQ-wise that is?
AW: It seems so.

S: What's his present reading ability?
AW: Now, about first or second grade.

S: Has he any police record?
AW: Yes.

S: For what?
AW: Breaking and entering, burglary, grand theft--auto, attempted rape...

S: Hey, he's not second grade material there!

S: What kind of a camp is this?
AW: Kind of a boys rehabilitat ing camp.
S: Can he tell time?
AW: Yes, I think so.
S: What's his explanation when you ask him why he's late?
AW: He usually just shrugs.
S: Is his mother overprotective?
AW: I don't think I can say for sure, but I'd guess no.
S: How's he feel about his sisters?
AW: I can't really say since I'm not Joe.
S: How old are they?
AW: Eight and six.
S: Could this lateness be a way to get the attention he doesn't get at home?
AW: Well, that's sort of a theory isn't it? I can't comment on it since I'm not Joe. How could you test it out?
S: Has any older male ever taken an interest in him?
AW: A parole officer.
S: What was the relationship with the parole officer? His comments?
AW: Generally the officer said Joe was surly.
S: You say he's been arrested 4 times?
AW: Well, actually 5 times. Convicted 3.
S: How many times has he tried...ummm...maybe you wouldn't know, but I'd like to get a fix on his street intelligence. How many times has he tried things and not been caught?
AW: Don't know for sure, but the guess is about 15.
S: Well, in other words he's hitting about 667?
AW: I guess that's right.
S: Is he a discipline problem in school?
AW: He has an extensive discipline record at school.
S: And at home?
AW: His mother thinks he's a pretty good boy. He does take care of his sisters and is the male figure in the home.
S: Well, what kind of a discipline problem is he? In school I mean?
AW: He been referred to the attendance officer often for his lateness. He's been surly in class according to teachers. "Very uncooperative" is what they say. He's failed to do most of his work.
S: What grade is he in?
AW: Third.
S: And he's thirteen?
AW: Yes.
S: How many times's flunked?
AW: He's been retained twice.
S: Oh, come on now. What happened during the rest of the time?
AW: Well, he's only been in American schools for about 5 years.
S: Where's he from, originally?
AW: Mexico.
S: Does he have a decent command of the English language?
AW: He speaks English.
S: Do the people in camp know about his hearing problem?
AW: Yes.
S: Yah, but in the story Pete shouts. Do we know whether he could be heard?
AW: No we don't know whether Joe heard him.
S: Is his hearing in the better ear normal or approaching normal?
AW: Yes.
S: Then I don't think he'd have that much problem with his hearing. I mean he'll have trouble localizing sound, but he should have trouble just hearing unless someone talked into his dead ear.

AW: O.K. That's your professional opinion. Remember Patty's in speech and hearing therapy.

S: Yah, that means she doesn't know anything, right? 

AW: O.K. That's your professional opinion. Remember Patty's in speech and hearing therapy.

S: Where's he live?

AW: Los Angeles.

S: In a Chicano area?

AW: Yes.

S: Do you have any social-emotional balance test data?

AW: No, I don't.

S: What was the stuff he brought back?

AW: Indian arrowheads.

S: I was wondering what were the nationalities of the other children in the camp. Because if he's a one of a kind that might, you know, account for this.

AW: In his cabin there's another Chicano and there are 4 blacks. The others are a mixture of Heinz 57 varieties, a little of everything.

S: Well, I think then he's trying to get attention. The tardiness is not really a problem, it's a symptom.

AW: O.K. Any other ideas?

S: Yes, I think he's also rebelling. That it also might be...well, he seems hostile as well as withdrawn.

AW: O.K.

S: I think he's frustrated due to his school ability. Because like it says, 'C Cabin had a 1 point or whatever lead over F Cabin, and A Cabin was one point behind F Cabin.' He wouldn't be able to figure that out too fast. I can't.

S: What do you win if you're best cabin?

AW: An extra week in camp.

S: Who won?

AW: Joe's cabin.

S: Were the arrowheads disqualified? Or did they enter them?

AW: No, they were entered. Only Pete and Joe knew where they're from.

S: Which cabin won the nature thing?

AW: Joe's cabin.

S: Probably Joe's behavior, if you're bothered by that stuff, was caused by a lot of things.

S: Yah, probably not one or two things contributed to it, but we'll never be able to know for sure with this kind of data.

AW: O.K. Need any other information?

S: Naw.
APPENDIX C

PHYSICAL SCIENCE PROBLEM SOLVING SESSION WITH SELECTED EDUCATION 435
STUDENTS: CONDUCTED BY ALLAN WALDDREN--"The Ice Cubes"
S: Is this a trick or something? Like trick photography?
AW: No, it's straight cinematography with some time delay, but no tricks.
S: Were both those ice cubes regular ice cubes?
AW: Both cubes were frozen water taken from the same ice cube tray.
S: Were the liquids the same in both glasses?
AW: Not exactly.
S: Hah! was there alcohol in one of them? The one on the right?
AW: Yes.
S: OK. Then the cube sunk in the right glass because alcohol isn't as dense as water. Right?
AW: That's your theory then, that alcohol is less dense than water and therefore won't support the ice cube. OK. That's a theory and remember I'll try not to comment on your theories. I'll just ask if you want any other information. Do you?
S: Nope. That explains it.
....(long pause)....
AW: Does everybody feel that explains it?
(inaudible)
AW: OK. Then...
S: Why didn't the cube stay on the bottom of that, that right hand glass?
S: Hey, yah!
AW: What do you think? Need any information?
S: Yes, was the temperature of the alcohol higher than the water?
AW: A little.
S: How much?
AW: Oh, let's see...about 10 degrees Centigrade.
S: So then it melted the cube faster, right?
AW: Yes.
S: Well, then that cube got smaller and didn't stay sunk.
AW: You're saying the smaller the cube is the better buoyancy it has?
S: Yah.
AW: Well, I could put a large cube in the right hand glass and it would float at the same level. How would you account for that?
S: Well,....it would?
S: Yah, you know when you add fresh ice to a drink it sure doesn't sink.
S: Is that alcohol the same as say, gin or vodka?
AW: Yes, except it's a little purer you know a higher proof anyway.
S: What is proof?
S: Real pure alcohol is 200 proof. Most of the stuff you drink is 80 proof or so. That makes it 40 percent pure. The rest
color and water. Right?

AW: That's as good a description of proof as I know.

___ S: Then the liquid in the right hand glass was pure alcohol?

200 proof?

AW: No, it was 160 proof and only 40 percent of that liquid was

alcohol.

___ S: Then there was something else in there?

AW: Yes.

___ S: What?

AW: Water.

___ S: That's all? Water and alcohol?

AW: Right.

___ S: Yah, but then when the ice cube melted it upped the amount of

water, didn't it?

AW: Yes, the amount of liquid water increased in both glasses as

the ice cubes melted.

...(pause)....

___ S: Doesn't alcohol evaporate?

AW: Yes.

___ S: Well, then maybe the alcohol evaporated and then the glass

had mainly water in it.

AW: Is that your explanation. That the alcohol evaporated and

left mainly water?

___ S: Could be...yah...OK. That's it.

AW: OK. I'm not sure how much alcohol evaporated. Some, no doubt,

but I don't think it was an appreciable amount.

___ S: So, well, then I'm wrong, right?

AW: Well, I can't say that it's wrong; just that from the informa-

tion I have not much alcohol evaporated.

___ S: Do alcohol and water mix?

AW: Yes.

___ S: And those cubes were regular ice cubes?

AW: Yes.

___ S: Then the... I mean when the ice cubes melted the water just

mixed in with the other water?

AW: Yes.

___ S: How big was the cube at the end?

AW: About half as big as when it was put in at the start.

___ S: You know that one cube was whiter, wasn't it?

AW: I don't know what you mean.

___ S: You know, it was cloudy like.

AW: OK. Yes, it was, the one in the right hand glass. Are you

thinking the one cube was different from the other and that

caused the funny action?

___ S: Well, could it?

AW: What do you think? Or I should say how can you find out?

___ S: That cloudiness could have been...like...oh...air bubbles.

AW: Do you want to ask about that?

___ S: Yes, were they air bubbles?

AW: Probably some of the cloudiness was due to air trapped in

the frozen water, but most of it, I think, was due to the

crystal formation. Monoclinic crystals, long needle like
crystals that water forms when it freezes.

S: Hey, this is education not science class.
AW: Sorry.

S: Was there more liquid in one glass than the other?
AW: A little, not very much.

S: Would that have been...no...
AW: Let me say this. If the amounts of liquid in both glasses had been exactly the same, the same thing would have happened.

S: Oh.

S: If you poured both glasses together would two layers have formed?
AW: Yes.

S: Well, what's that get you?
S: Well, then the ice cube would float on top of the one layer and underneath the other.
S: How's that? He said alcohol and water mix.
AW: They do, but they also separate.
S: Yah, all bar tenders know that.
S: Well, then...yes, as the water, I mean ice cube melted the water amount increased and the ice cube floated on the layer, the heavier one. Isn't water heavier than alcohol?
AW: Yes it is.
APPENDIX D

ANALYSTS TRAINED IN THE STUDY
ANALYSTS TRAINED IN THE STUDY

Victor E. Edmonds--Doctoral Candidate, Curriculum and Foundations Faculty, The Ohio State University. Mr. Edmonds has taught both high school and college classes and served as a guidance director for the Ohio State Youth Commission.

Judith Morris--Graduate Student and Teaching Associate, Curriculum and Foundations Faculty, The Ohio State University. Miss Morris has taught secondary and college classes and served as an Assistant to the Director of the American Association of Colleges for Teacher Education.

Jane Gammel--Supervisor of Student Teachers, Faculty of Early and Middle Childhood Education, The Ohio State University. Mrs. Gammel has taught elementary school and was an Assistant to the Director of the Teacher Corps, University of Kentucky.

James D. Gammel--Doctoral Candidate, Faculty of Educational Development, The Ohio State University. Mr. Gammel has taught junior high and college classes and served as assistant principal of a junior high school.
APPENDIX E

INQUIRY SESSION WITH J. RICHARD SUCHMAN AND 6th GRADERS
"THE KNIFE," Film 15
INQUIRY SESSION WITH J. RICHARD SUCHMAN AND 6th GRADERS
"THE KNIFE," Film 15

JRS: O.K. Any theories? Yes?
S: Was the knife made out of copper?
JRS: Partly.
S: Was it an alloy of any sort, maybe?
JRS: Yes.
S: Was it something with copper and tungsten?
JRS: No.
S: Well, in that basin, was it a chemical or water?
JRS: You mean something different from water? You could call water a chemical too, I should think.
S: Yeah.
JRS: Well, do you want to ask me about one or the other?
S: Is it water?
JRS: Yes.
S: Is it supposed to be copper, or since the blade is partly copper, would the other be sulfur? Not sulfur, but,--yah, sulfur?
JRS: No.
S: Is the other iron?
JRS: No.
S: Well, steel?
JRS: To be exact the blade is made of copper and steel.
S: Was the gas jet the same jet that was used in the other film?
JRS: Yes.
S: So that it was the same temperature?
JRS: Approximately.
S: Did the blade bend four inches?
JRS: When you say four inches, you mean off the ah.....if you measured the distance that the end moved down from where it was to start with?
S: Say the blade is straight. Well, when it came down like this, was it four inches here?
JRS: Oh, maybe. Somewhere around that.
S: Was the water cold?
JRS: What do you mean by cold?
S: Well, was it below freezing?
JRS: No.
S: Was the water hot or warm?
JRS: No.
S: Then it was cool?
JRS: I would say so.
S: Was the water chilled below the temperature of the room?
JRS: Ah, probably a little less than the room temperature, but not much less.
S: In other words, if the temperature in the room was approxi-
mately eighty degrees, the water would be seventy-five?

JRS: Somewhere around that. Maybe even sixty-five degrees.

S: You say that this was the same flame as the last one. Then you said it was kind of a gas mixture again?

JRS: Right.

S: Would there have been a reaction between the material in the blade and the chemical in the water?

JRS: Reaction? You mean a chemical reaction?

S: Yes.

JRS: No.

S: Is this blade or knife or whatever you want to call it the same strength as an ordinary kitchen knife would? You know if you could hit it against the kitchen table, it wouldn't bend or something?

JRS: No, I would say it wouldn't have quite that much strength.

S: Now, was the part that was the copper and steel part—the part that we say saw, the blade—was that steel?

JRS: Partly.

S: In other words, it was...ah...it was half and half?

JRS: Yes.

S: Was it mixed together?

JRS: Oh, no.

S: Then there was more steel that was in it, than copper?

JRS: No.

S: It was about the same?

JRS: Uh-huh.

S: What was the handle made of?

JRS: Ah...you want to ask a yes-or-no question?

S: O.K. Was the handle made out of wood?

JRS: Yes.

S: Was there any platinum in the metal composition of this knife?

JRS: Not that I know of. I'm not certain about whether there is platinum in steel.

S: O.K. Was the blade heated to about fifteen hundred degrees?

JRS: Oh, no. The blade didn't get that hot, probably.

S: A thousand?

JRS: I kind of doubt if it even got that hot.

S: Five hundred?

JRS: Possibly.

S: Was the water in the basin salt water or chlorine? I mean was it salt water?

JRS: No.

S: Was it ordinary tap water?

JRS: Yes.

S: Was the knife actually melting?

JRS: No.

S: Was the heat just making it bend a little, I mean?

JRS: Well, that's something of a theory, isn't it?

S: Yeah.

JRS: The idea of heat making it bend. See what you can do with that theory. Does it suggest any experiments you could try to test it?

S: I'll think a while.
S: Was the pressure in the flame lower than the room pressure?
JRS: What do you mean?
S: 14.7 pounds?
JRS: No.
S: Was the room temperature high?
JRS: High?
S: Yes.
JRS: Well, there was a film being made and there were floodlights so it got kind of warm, but no more than eighty degrees I would guess.
S: I noticed that the knife blade was really polished. Would it make any difference if it was rusted?
JRS: If some oxide or some rust had formed, I think it still would behave pretty much the same way.
S: Would the time in the fire—the blade in the fire—was that a major factor?
JRS: Well, that's kind of a theory too. See if you can test that for an experiment or some kind of data gathering.
S: Say the blade was in there for fifteen seconds. Would it have bent further if it was in there for thirty?
JRS: Yes.
S: Was the flame...would the gas coming into it...would that have anything to do with a ...not a chemical reaction, something like it..in the air, that's a sort of..the flames you know, spreading something, fumes, something into the air that made a chemical change when it got in the air--now would that make any difference?
JRS: O.K. Now you've got another theory. You know you can test that theory. You could do some experimenting to see whether that does make a difference. O.K.
S: If she had held the knife differently, would it have changed the way it bent?
JRS: Well, now why don't you suggest a different way of holding it and then see what happens?
S: Well, if she had held it--well, like this--and put it in the fire, would it have bent up or down?
JRS: All right now. It all depends. You haven't told me enough to ...ah...about how she was holding the blade.
S: Well, she sort of had it in her fist and her--top of the hand was up.
JRS: O.K.
S: Then she put it in the flame and it bent up. Right?
JRS: Right.
S: Now if she turned it over and did the same thing, it would bend down....if she did it all over again. Wouldn't it?
JRS: Yes it would.
S: If the blade was held in the water longer than it was, would it have changed its reaction? Like get straighter?
JRS: No. If it was held in the water longer, it wouldn't have made any difference.
S: O.K. I pass.
APPENDIX F

DEMONSTRATION WORKSHOP: INQUIRY DEVELOPMENT PROGRAM
"DRINKING BOILING COFFEE"--Film 16
S: Is that real coffee?
T: Yes.
S: Didn't he burn his mouth?
T: Not really.
S: Could you, I mean, would you drink that?
T: Yes, in fact I'd like a cup of coffee now.
S: Were those some special gloves? I mean like to hold hot stuff?
T: They were asbestos gloves.
S: Did he have to wear them?
T: What do you think?
S: Well, I guess so.
T: I don't want to mislead you. What I meant was you could have asked me if he would have burned his hands if he hadn't used them, and then I'd have said no.
S: Oh.
S: Well, then if it didn't make any difference with the gloves, why'd he wear 'em?
T: That's hard for me to say for sure since I wasn't there when they made this movie. What do you think?
S: Probably just to mess us up.
T: That could be.
S: Was that a hot plate?
T: That he put the beaker on?
S: Yeah.
T: No, it wasn't.
S: Was it hot, though?
T: No, it wasn't. I mean just about room temperature.
S: Well, how'd the coffee boil then?
S: Was that steam coming off the coffee?
T: It was water vapor, which most people call steam, yes.
S: Would it have boiled if he didn't put that jar over it?
T: The bell jar? No it wouldn't.
S: Well, then was it the jar that did it?
T: Can you tell me what you mean?
S: Well, if the jar made it boil, that's just what you said.... then well, what... oh, I don't know.
S: How hot was the coffee when he drank it?
T: I don't know the exact temperature, maybe 150, 160 degrees.
S: Is that the regular temperature for coffee?
T: That's what I mean. It was about the temperature one drinks coffee at.
S: Was it hotter first or last?
T: About the same.
S: But it boiled?
T: Is that a question?
S: Well, it did, didn't it?
T: Yes, the coffee boiled. It boiled twice actually.
S: Isn't that why he used the gloves?
T: I'm not sure why he used the gloves.
S: You mean he didn't need them?
T: Let's say the experiment would have worked the same way if he hadn't.
S: Is it pressure that has something to do with it?
T: Now, then, that's kind of a theory isn't it? What kind of questions can you ask me to find out?
S: Was... umm... was the pressure the same in the jar?
T: No.
S: It wasn't?
T: Right.
S: Well, what was it?
T: When?
S: When the man put it in the jar?
T: The air pressure when the man put the coffee under the bell jar was about 15 pounds per square inch.
S: What was it when he took it out?
T: About 15 pounds.
S: Well, then....
S: Did the pressure go up while it was boiling?
T: No.
S: Down?
T: Yes.
S: What went down?
S: The pressure in the jar.
S: Is that what she said?
T: That's what I meant to say.
S: The pressure went down when the coffee boiled?
T: Yes.
S: How can that be?
T: It did.
S: Was that some special thing he put it on?
T: I guess you could call it a special thing.
S: What was it?
T: A vacuum pump.
S: That pulled the air out?
T: Yes.
S: Does something boil easier when the pressure goes down?
T: Yes.
S: Oh.
S: And the temperature stayed about the same all the time?
T: Yes.
....(long pause)....
T: Do you have explanations now for how the man could drink the boiling coffee?
S: Yah, it really wasn't that hot.
APPENDIX G

DEMONSTRATION WORKSHOP: INQUIRY DEVELOPMENT PROGRAM
"The Pulse Glass"
DEMONSTRATION WORKSHOP: INQUIRY DEVELOPMENT PROGRAM
"The Pulse Glass"

S: Is that water in there?
T: No, it isn't water.
S: Would it work if it was water?
T: Let's see, I think it would but it'd take a long time.
S: Can I hold it once?
T: OK.
S: Hey, look at it now!
T: Does that tell you anything?
S: Just that it works for me too.
T: Do you have any other questions?
S: Is there some special chemical in there?
T: It's alcohol that has been colored red.
S: Is this like those things in stores...umm...you know love meters?
T: I haven't seen those so I really couldn't...wait, yes, my partner says yes.
S: Well, if it is, then the alcohol goes up faster if you squeeze it.
T: Do you want to squeeze the pulse glass and see if it goes faster?
S: Does it?
S: No, just the same.
S: Let me try.
T: Do any of you have an explanation yet? Yes?
S: Well, I think this is a trick thing and if you hold it in your hand the stuff goes in the other side.
T: Let me see if I understand what you said. You think it's a trick device that works by holding one ball in your hand. All right. Can you guess what would happen if you held both balls, one in each hand?
S: Umm, would it maybe break the glass?
T: Do you want to try it?
S: Does it work by heat?
T: Well, now that's kind of a theory isn't it? Can you ask me questions that would basically test out that theory?
S: Yah, would it work faster if it was hot?
T: Exactly what do you mean? If what were hot?
S: You know, like if your hand was real hot or if you put the ball next to a fire?
T: Yes, the liquid would move faster.
...(long pause)...(inaudible dialog)....
T: Everybody has an explanation? Anybody need any more information?
...(long pause)....
T: All right, would you tell me your explanations?
S: Like he said, the heat makes the alcohol go over there.
T: All right, any other ideas?
S: Does this have any use?
T: I'm sorry, what do you mean "any use?"
S: You know, can you do anything with it?
T: Well, let's see.... Well, what do you think?
S: I don't know. It's just a game or something?
T: What do the rest of you think? Any ideas?
APPENDIX H

INQUIRY SESSION WITH 7TH GRADERS, NORTON MIDDLE SCHOOL
"Joe: A Case Study"
INQUIRY SESSION WITH 7TH GRADERS, NORTON MIDDLE SCHOOL
"JOE: A CASE STUDY"

AW: Why do you think Joe's late?
S: Cause he's punky.
S: Cause he's got stuff to do.
AW: OK, cause he's punky or got stuff to do. What else?
S: He's (inaudible)
AW: I'm sorry, I didn't hear.
S: He's mischievous.
AW: OK, he's mischievous.

...(long pause).....
AW: Any other ideas?
S: He's a goof off.
AW: A goof off?
S: Yes, that's right, he's a goof off.
AW: Alright, any other ideas?

...(long pause).....
AW: Do any of you need other information? (nods) Well, I have some information about Joe and I'd be glad to let you know what ever I can about him if I have the information or if I know. But I can only give you hard data. Do you know what I mean by hard data?
S: Just the facts.
AW: Right! Since I'm not Joe, I can't tell you how he felt or what he was thinking. The same applies to Pete or Ben or anybody else. OK? What questions do you have?
S: What kind of kid is he?
AW: I can tell you some things, but what do you want to specificall­
S: Does he get in trouble?
AW: Do you mean like in jail?
S: Yah.
AW: Yes, then Joe has gotten in trouble.
S: Does he mess around and does he...is he not in the right place at the right time?
AW: Yes, I guess you can say that's happened to Joe. Does that help you?
S: Yah, a little bit.
AW: Anything else?
S: Does he have a Mom and Dad?
AW: He lives with his mother and...
S: Where's his dad?
AW: No one knows. He left home and hasn't been seen.
S: Why?
AW: No one knows for sure except his dad. Remember I can only give you hard data.
S: When did his dad leave?
AW: About three years ago.
S: Does Joe have any brothers and sisters?
AW: Two younger sisters.
S: What's his last name?
AW: Moreno.
S: Did he ever get in bad trouble?
AW: What's bad trouble? Jail?
S: Yah.
AW: Um hum.
S: What for?
AW: Stealing a car...
S: How old is he?
AW: Thirteen.
S: And he was driving a car?
AW: I didn't say he was driving. I said he stole one and got caught.
S: Was that all?
AW: No, he was convicted for selling stolen goods, radios and television sets and...
S: How many times was he arrested?
AW: Three times.
S: Has he been in jail? And how long?
AW: The longest he's been in jail is 1 month. He's been sentenced to 6 months but put on probation.
S: Who's car was it?
AW: I really don't know.
S: How do you know this stuff?
AW: This story was taken from a case history that I had access to. What you have read in that one page was extracted from that longer one. I couldn't put the whole thing on one page.
S: Where'd did this all happen?
AW: California.
S: What's he look like?
AW: What's he look like? Well, he's about 5' 3" tall, dark skin, black hair, brown eyes, I guess that's about it.
S: I'm 5' 4", am I like him?
....(laughter)....
AW: I guess just an inch or so away.
S: What's his full name, his real full name I mean?
AW: Jose Angelo Moreno.
S: In the story was he in some kind of cave or something? No, wait, was this some kind of special camp?
AW: Special camp? Yes it was a camp in which everybody there had a previous jail record and it was run basically to rehabilitate boys.
S: Who runs this camp?
AW: The state of California and one of the large service clubs... I think like the Lions or Kiwanis. They help with the finances. Raise money.
S: Did he want to go?
AW: Did he want to go? Well, he was sent there by the court. By a judge.
S: Part of the reason he's late is that he doesn't care, he doesn't want to go in the first place.

AW: The reason he's in the camp is that the judge told him he'd either go there or to jail.

S: Well, that's a big choice!

S: Do they get locked up when they go to sleep?

AW: No, they sleep in cabins and there are no locks on the doors.

S: How many in a cabin?

AW: In most cabins there are twelve boys. In Joe's there are thirteen.

S: Do the other boys like him?

AW: I can't really tell you for sure since I'm not any of the other boys.

S: Well, I don't think they really liked him or they wouldn't have said, 'Oh, let's leave without him,' and stuff.

S: He doesn't hang around with them does he?

AW: No, there's not much of a record of his hanging around with any of the boys.

S: Does he have any friends in the camp? Does he know or did he know anybody there?

AW: He knew one boy in another cabin before. As far as I know he doesn't or didn't know any of the boys in his cabin before.

S: How long has he been in the camp?

AW: About four days.

S: How long will they be staying?

AW: Another week. Unless, of course his cabin would win the best cabin award, and then they have an extra week.

S: Have his sisters gotten into any trouble?

AW: No.

S: Does his mother ever pay any attention to him?

AW: What do you mean by pay attention?

S: Does she care about him? Or does she just let him run around?

AW: Well, the record I have about this shows that she cooks meals for him, washes his clothes, talks to him, goes to school for conferences about him, and stuff like that.

S: Does he have any brothers?

AW: He has no brothers.

....(long pause)....

AW: Is this an easy problem?

S: Uh uh well, maybe people have always waited for him and he's not used to getting blamed for being late.

S: Are we trying to find out why he's late?

AW: Right.

S: If they'd have left him a couple of times he'd probably be on time.

S: Maybe he had something to do.

S: All the time?

S: Yah.

S: Maybe he wanted to be late. Maybe he was going to run off.

AW: Could be. You think maybe he wants to run away or be off by himself.
...(long pause)....

AW: Any other information you want?

S: Where's he live?

AW: When he's not in this camp, Los Angeles.

S: Does he go to a regular school?

AW: If you mean one like this one, yes.

S: What grade?

AW: Fourth grade.

S: Has he failed?

AW: Yes, twice.

S: Is he a behavior problem?

AW: In school?

S: Uh-huh.

AW: Yes, he's been regarded as a behavior problem in school.

S: Like why?

AW: Often because he's been late.

S: Does he live in a very good area?

AW: How do you describe a very good area?

S: Does he live in the ghetto?

AW: Yes.

S: Isn't that a good area?

AW: It all depends on your notion of good.

S: Does his Mom work?

AW: Yes, she has two jobs. Works in a laundry 5 and a half days a week, and she cleans offices 3 nights a week.

S: Does she get any sleep?

AW: Not a lot.

S: I think he's late cause he wants to be alone, by himself.

S: I doubt that!

AW: That's neat. You disagree. What do you think?

S: I don't know yet.

AW: Well, could you test out your theory if you had some more information?

S: I don't know.

AW: If you thought he wanted to be alone all the time, you could find out if he was alone most of the time.

S: Well, was he?

AW: Nope.

S: Did he have to take care of the kids, his sisters, I mean?

AW: Yep, he often had to take care of them.

S: Was he with anybody when he was late all the time?

AW: Sometimes.

S: Who was he late with at home?

AW: Mostly one or more of a group of guys he often hangs around with.

S: His friends?

AW: I guess.

S: How old are they?

AW: All of 'em are older than Joe, like 2 to 5 years older.

S: Does he take drugs?

AW: There's no record of his taking drugs.

S: I bet he does, he just never got caught.

AW: Maybe, but don't forget I can only give you factual data I have.
S: He might be late just to get in trouble.
S: Yah, get attention.
S: Maybe he doesn't get enough attention at home, so he tried to get attention by doing other things.
AW: OK, that's a possible explanation. Any other ideas? Anything else you want to know?
S: Are his sisters very smart?
AW: The only information I have like that is their IQ's are over 100, and they're in the grade they're supposed to be in for their age.
S: Did he get along with 'em?
AW: He seems to get along with them. There's nothing in the record to show he doesn't.
S: Maybe he has to take care of them and it's too much responsibility. Is it?
AW: Might be, that's a possible explanation. How can you test it out?
S: I don't know. That's just my guess.
S: What's Joe's IQ?
AW: It's been measured twice. First time 83 next time 78.
S: Oh, that's icky.
S: What grades are his sisters in?
AW: One's in third and one's in first.
S: And Joe's in fourth?
AW: Right.
S: What does his mother say about all this? His failing and trouble?
AW: What does she say... there's not too much in the record of her saying too much.
S: Does she care about him?
AW: She seems to.
S: Like how?
AW: She's gone to school for conferences, and things like that. She asks him about school, asks him if he has homework.
S: Does he have homework?
AW: He may, but he doesn't seem to do much at home.
S: Doesn't he ever see his father?
AW: He hasn't for three years.
S: Does he have any pets?
AW: Not that I know of.
S: Is Joe normal?
AW: What do you mean normal?
S: Oh, like does he or is he short or real tall, or, no you told us. Is he all right, though, like you know, normal?
AW: Well, normal except for hearing.
S: Oh.
S: Is he deaf?
AW: Well, he's partially deaf in his right ear.
S: Since birth?
AW: No, it happened about three years ago.
S: Is that, I mean, oh I don't know, I pass.
S: How did he get deaf?
AW: The medical records show that his ear drum was damaged by a blow to the head. It is reported that his father cuffed
him along side the head several times causing the ear damage.

S: Is that, I mean did his father leave home then?
AW: Right after this incident his father left the family.
S: Did he mean to do it?
AW: Remember I can only give you hard data answers. I don't know what his dad was planning or thinking.
S: Did his dad beat up on him often?
AW: There are a few records of his dad hitting Joe.
S: Was he late before his dad hit him?
AW: Not as often.
S: Oh, oh.
S: Oh, oh what? What does that mean?
AW: All right I give up, what does that mean?
S: Well, ...oh, I don't know.
S: Did his dad use dope or drink a lot?
AW: There's no record of dope, but he was known to have drunk a lot.
S: Did his dad get drunk often?
AW: There's quite a few arrests for it I guess about 15 or so.
S: Does Joe drink?
AW: It's never been noted that he did.
S: Does he smoke?
AW: Yes.
S: A lot?
AW: I don't know if it's a lot, but he's known to smoke.
S: Was he ever smoking when he was late?
AW: Yes.
S: Aha, he was catching a smoke. Was he smoking when he was late this time?
AW: Which time?
S: You know when they were all going back to go swimming?
AW: I really don't know.
S: Was he smoking in the mess hall?
AW: I don't know that either.
S: Does he know how to tell time?
AW: It seems that he knows that.
S: Does he have a watch?
AW: No.
S: Well, then how would he have known that it was time to come back? You know 2:30 or whatever he was supposed to come back.
AW: How would he? I don't know what do you think?
S: Didn't Pete blow the whistle?
AW: Yah, but maybe he didn't hear it.
S: That's a possibility.
AW: Well, who was he with? Anybody?
S: Yah, but you said they were supposed to pair up.
AW: Right but there were 13 in this cabin. It's kind of hard to pair up 13.
S: Well, why didn't Pete go with him?
AW: I don't know, I'm not Pete.

...(long pause and then an inaudible dialog)....
AW: Well, let's see how many different explanations do we have?
S: Well, I think he couldn't hear the whistle or he didn't hear the whistle and he didn't care anyway. He was probably or didn't want to go swimming anyway.
S: We don't know that. Did he like to go swimming?
AW: I can't say for sure, but he did go everytime the cabin had it's swimming time.
S: Hey, what was he doing when Pete found him? What did he have?
AW: Well, he was coming down the slopes and had a shirt full of old arrow heads and stuff like that.
S: Where's he get them?
AW: In a cave up on the slope.
S: That's probably what he really liked, exploring around I mean.
AW: OK.
S: Yah, but every time he's late he doesn't go exploring for arrow heads.
S: Yah, but maybe he's doing neat stuff.

...(inaudible dialog).....
AW: Well, let's wrap this one up. What do you think? Why's Joe late?
S: It could be a lot of reasons, and we'd never know for sure.
AW: Such as?
S: Well, maybe he's just a loner. Or maybe he wants to be late.
S: Yah, or maybe he didn't hear or couldn't tell time for sure.
S: This is silly, we'd never be able to know for sure.
AW: OK, your guess or guesses could be and probably are as good as anybody elses.
S: Did anybody ever ask Joe why he's late?
AW: Yep.
S: And what did he say?
AW: The record I have shows he just shrugged his shoulders most of the time or said, "I dunno."
S: That's what I do.
AW: I notice that the bell's going to ring, so I just want to say thank you for playing this little game with me.
S: Well, why was he late?
AW: I really don't know. All of the theories you had sound pretty interesting and may be the explanation. Are you happy with your explanations?
S: Yah, I got it.
AW: What's your final hypothesis?
S: You know, he wanted attention, he's a natural loner, he didn't hear and he really didn't want to hear, and that's about it.

....BELL.....
AW: Thanks.
S: Yah, but why was he really late all the time?
AW: Well, no one can say for sure. What do you think?
S: He just wanted to be naughty.
S: Oh, come on Sherry, that's dumb.
AW: Maybe that's a good explanation, how could we find out?
S: You couldn't, that's why it's dumb.
AW: Was this a dumb thing to do?
S: Oh, no, I guess not, but we never found out anything.
AW: Didn't you?
S: Yah, we found a lot of stuff.
S: Hey, was Joe white?
AW: Not exactly.
AW: What do you mean?
AW: He was Mexican American.
S: Oh.
APPENDIX I

INQUIRY SESSION CONDUCTED BY DONALD TEFFNER WITH 7TH GRADERS,
NORTON MIDDLE SCHOOL--"Joe: A Case Study"
INQUIRY SESSION CONDUCTED BY DONALD TEFFNER WITH 7TH GRADERS,
NORTON MIDDLE SCHOOL—"Joe: A Case Study:

DT: Now I'll try to answer any question you ask except for questions that ask me to tell someone's feelings or to tell you if your right about an explanation. OK? Fire away.

S: Did they make it back to camp for swimming?
DT: No, they did not.
S: Did he get in trouble for being in the off limits area?
DT: Yes, he did. He was denied some free time privileges for the rest of the camp session.
S: How long was the session?
DT: The sessions are for two weeks, unless you win best cabin award then you get an extra week.
S: Who won it, or do you know?
DT: A cabin did win and got an extra week.
S: Did they win for the ah...nature thing?
DT: They did win the nature exhibit.
S: With what exhibit?
DT: With the deer skull and with what Joe found and some other stuff.
S: What did Joe find?
DT: A whole bunch of Indian arrow heads and things like that.
S: Where?
DT: In the caves on the off limits area.
S: What did he do in the lifesaving thing?
DT: Joe did not pass lifesaving.
S: Did the rest of 'em?
DT: All but two of them.
S: How could they win first prize with the deer head and the arrow heads if the arrow heads were from out of bounds?
DT: Pete entered the arrow heads in the exhibit but didn't mention where they came from.
S: (snickers)
DT: What else do you need to know?
S: What do you think right now? Why was Joe late?
DT: OK. Any other ideas?
S: How old are the boys?
DT: Ages range from 12 to 16.
S: How old's Joe?
DT: 13.
S: How old's his counselor, what his, yeah Pete?
DT: 18.
S: Well, no wonder he didn't like him.
DT: Why's that?
S: Oh, you know.
DT: What's your idea?
S: No wonder Joe got in so much trouble...ha ha.
DT: You mean because of that old guy Pete?
....(laughter)....
S: How did Joe get to camp? I mean did he want to come?
DT: He was brought to camp by a court bailiff late on a Tuesday afternoon.
S: Why was he?
DT: Because the judge told him to either go to camp or to jail.
S: Why was he in court? What did he do?
DT: Stole a car.
S: At age of 13?
S: Well, maybe that's why he's always late.
DT: He's out stealing cars?
....(laughter)....
DT: All the other boys in the camp arrived on a Sunday and Joe arrived on a Tuesday afternoon.
S: Did he like the boys?
DT: He only knew one of the boys from before. And he seemed to like him. Also remember I can't tell you how Joe felt about things since I'm not Joe.
S: Who was this boy?
DT: That Joe knew? ...a boy in F cabin.
S: Could they change cabins?
DT: No, they were assigned to a cabin for good.
S: Did anybody know Joe's been in court?
DT: Yes, in fact everybody in the camp had a jail record of some kind.
S: Everybody, even the you know Pete?
DT: Everybody from the camp director on down.
S: What kind of a camp was this? I mean why'd they have it?
DT: To help rehabilitate boys with police records.
S: What all did they do in camp?
DT: You mean the activities?
S: Yah.
DT: Well, everything was done by a cabin unit. There was swimming, riding, games, handicraft, camping, canoeing, hiking, archery, all those sort of things.
S: Were they locked up at night?
DT: No.
S: Was Joe liked by the guys?
DT: I can't tell you cause I'm not the other guys. I might be able to give you information though, that would help you decide.
S: Did they, well, umm...did they all hang around together?
DT: Like I said, everything is done together, in a cabin unit. They eat together, swim together, like that.
S: Was he the only one who was ever late for anything?
DT: No, but he was late more than he was on time.
S: Did he have a watch?
DT: No.
S: Well, then how would he know what time it was?
DT: That I don't know. But the others didn't have watches either, except for Ben, and he was almost late, too.

S: Maybe he just didn't like to be on time.

DT: Now, that's an explanation or the start of one. How could you find out or prove it to yourself?

S: I don't know, maybe ask him.

S: Yah, what did he say when they asked him why he was late?

DT: He just usually shrugged or shook his head.

S: Well, if they left without him a couple of time maybe he wouldn't be late anymore.

S: Yah, like if they said you be here at such and such a time or we're going without you.

DT: OK. Then you're suggesting a way that might cure him of his lateness.

S: Yah, that might work.

DT: How could we find out? Or do you want to ask if anybody ever tried that with him.

S: Did they?

DT: Yes.

S: And he still was late after that?

DT: Mostly.

S: Did he just like to be late?

DT: That's a possible theory isn't it?

S: I don't know. I mean maybe he just wanted to be late; so he was.

DT: How can you find out? What questions could you ask that would help you decide?

S: Well, was he on time for anything most of the time? You know, for stuff he liked to do?

DT: Yes, he seems to have been on time when he met his friends after dinner at home.

S: Who were his friends, what did they do?

DT: Oh, boys a few years older than he was. And, ah, oh they mostly played pool or just hung around the neighborhood.

S: Like a gang?

DT: Yes, sort of.

S: How much older?

DT: Oh, between 3 to 5 years older.

S: Are you going to tell us when we figure out why?

DT: No, not really.

S: Then why are we doing this? I mean what good is this?

DT: Well, I'm helping out a friend and I also thought you might like to try to figure out this problem on your own. Would you rather I told you the reason or reasons?

S: Yah.

DT: What if I said no one really knew and this was your chance to do some detective work and explaining without anybody telling you you're right or you're wrong.

S: I don't know. We don't do this kind of stuff much here.
APPENDIX J

INQUIRY SESSION WITH 7TH GRADERS, NORTON MIDDLE SCHOOL
"JOE: A CASE STUDY"
INQUIRY SESSION WITH 7th GRADERS, NORTON MIDDLE SCHOOL
"JOE: A CASE STUDY"

AW: 0.K. Now can you figure out why Joe's late?....I've got, incidentally, a lot of data about Joe that you don't have. So if you ask me questions I can maybe give you some of the data you'd need to figure out why he's late so much. What do you want to know?

S: What kind of camp is this? Is it a special camp?
AW: Yes, it is a special camp. Everybody there has been arrested at least once.

S: Everybody?
AW: Yep, from the camp director on down.

S: Why?
AW: Why? Well, I guess for this kind of a camp--to rehabilitate boys--they figure that everybody there should know first-hand-wise what it's like to have been arrested. Is that what you mean?

S: I guess so.

S: What's Joe's family life like?
AW: Well, he lives with his mother and two younger sisters in an apartment.

S: Where does he live?
AW: In Los Angeles.

S: No, I mean does he live in the ghetto?
AW: Yes, I think you could call the area where he lives a ghetto.

S: Does everybody in the camp come from the ghetto?
AW: No, not everybody.

S: Do you mean there's rich kids there too?
AW: Well, if you mean are there boys there from rather well-to-do families, then yes.

S: What did Joe do? I mean what was he arrested for?
AW: Well, he was arrested for stealing a car, selling stolen merchandise and...

S: When did all this happen?
AW: Over the last two years.

S: Is he...well, is he different from the, you know other people?
AW: We're all a little different. What specifically do you want to know?

S: Well, was he....Oh I don't know.

S: Is he a real boy?
AW: Yes, he is.

S: Is this a real story?
AW: Essentially yes. It's a boiled down version of a part of a day in the life of a real boy.

S: Does Joe like camp?
AW: Again I don't know because I'm not Joe.

S: Do they go swimming, canoeing and then into town?
AW: Well they go swimming and ah...have some boating instruction a couple a times a week, but there's no real town nearby.

S: Is Joe an athlete?
AW: A professional athlete?
S: No. I mean is he good in sports?
AW: Well, he plays some sports, but I don't know how good he is according to your standards.
S: Did he play baseball?
AW: Yes.
S: What'd he play? What positions?
AW: Let's see. Right field, I think.
S: Was he average?
AW: In what way do you mean average?
S: You know, average?
AW: No, I'm afraid I don't know what you mean. Specifically what is it you want to know?
S: Well, what good are you? Ha, ha.
S: Is Joe a sissy?
AW: Well, how do you define a sissy?
S: You know, runs away from stuff?
AW: Well, according to what I remember from the record of this, Joe hasn't been noted for running away from anybody.
S: Was he tough, then?
AW: Do you mean, did he start fights or get into fights?
S: Yah.
AW: Yes he's been in fights.
S: Did he win 'em?
AW: More than he seems to have lost.
S: Then, he was a good fighter?
AW: I guess he was fair to middlin.
S: Did he fight in camp?
AW: Yes.
S: Who with?
AW: Let's see....Joe's fought with Ben, Jack, Bill, George and even with Pete. And then a couple of guys from another cabin.
S: He fought with Pete?
AW: Yep.
S: How old is Pete?
AW: Eighteen.
S: And Joe?
AW: Thirteen.
S: Was Joe big?
AW: About five feet four inches. No wait, about five two.
S: How about Pete?
AW: About six feet.
S: Who won?
AW: I don't really know. You mean in the fight with Pete? I don't know really.
S: Did he have a knife?
AW: No neither of them had a knife.
S: Are Pete and Joe related? Like cousins?
AW: Not that I know of.
S: Where's Pete from?
AW: Los Angeles.
S: Same as Joe, then?
AW: Right.
S: Was everybody from Los Angeles?
AW: No, but they were all from California.
S: What did Pete do...I mean for his jail stuff?
AW: Why was Pete arrested? Or I should say for what charge was he convicted?
S: Yah.
AW: Burglary and assault.
S: How long ago?
AW: I guess about 3 or 4 years ago.
S: Is he straight now?
AW: Yes.
S: Does Joe have a girl friend?
AW: He's been out with a couple of girls, but not one particular one.
S: How old is he again?
AW: Thirteen.
S: How old are his brothers and sisters?
AW: He doesn't have a brother, but his sisters are 6 and 8.
S: How old is his mother?
AW: About 30 or 31 I think.
S: His dad?
S: His dad's dead. Remember!
AW: Now wait a minute. We don't know that his dad is dead. I said that Joe lived with his mother and sisters.
S: Where's his dad, then? In jail?
AW: I really don't know. He hasn't been seen in about 3 years.
S: Joe hasn't seen his dad in three years then?
AW: Not that I know of.
S: He might be planning to find him. .....What's his dad do?
AW: He was a farm worker and then sometimes a factory worker.
S: Maybe Joe's planning to search for him?
AW: Maybe.
.... (long pause)....
AW: You look like you've been thinking hard. Do you have an idea?
S: Oh, umm, no, not yet.
AW: O.K. Does anybody need any other information?
S: How old's his dad?
AW: Early forties, I think.
S: Wow!
LAUGHTER
S: And his mom's only 30?
AW: Right.
S: Does she go out with any guys?
AW: Occasionally. Once in a while she goes out on weekends.
S: Are they divorced?
AW: No Joe's mom and dad aren't divorced.
S: Oh, oh,...separated?
AW: I guess you could say they've been separated. But it's not a legal separation.
S: Has his dad gone off before?
AW: Yes.
S: But he came back, right?
AW: Right.
S: Where'd he go then?
AW: Mexico.
S: Mexico?
AW: Yes, Mexico.
S: Why'd he go there?
AW: I don't know for certain, but maybe he just went home. He was Mexican.
S: Was Joe's mom a Mexican?
AW: Yes, she's Mexican too.
S: Well, then, maybe...Were all, I mean were the other boys from Mexico originally.
AW: All the boys in camp? No. Some were.
S: Hey, did he speak English?
AW: Yes.
S: I mean did he have an accent?
AW: Somewhat. His native language was Spanish, but he spoke English too.
S: Well, did he talk good, I mean did he talk as good in English as he could do with Spanish? Or did he understand well?
AW: Spanish was his native language. Spanish was spoken in the home most of the time. English was spoken in school. I'll let you guess which he understood best.
S: Could he tell time?
AW: Yes, I think so.
S: Did he have a watch?
AW: No.
S: Well, then how could he know what time it was to be back at what ever time he was supposed to be there?
AW: I don't know, but some of the other boys didn't have watches either. In fact most of them didn't.
S: Were any of them late, ever. I mean a lot of times?
AW: Some of them showed up late, but never as often as Joe did.
S: Well, maybe he didn't care or maybe time didn't make that much difference to him.
AW: O.K. That'll be kind of an explanation or your theory. Can you figure out a way to test it?
S: Yah, was he ever on time for stuff?
AW: Yes.
S: Like what?
AW: Sometimes he was on time for school, most of the time he was on time for his jobs, and he always was on time when he met his friends at night.
S: He had a job. You mean he worked?
AW: Yes, he had a job after school.
S: Doing what?
AW: His first job was in a grocery store, then a hardware store, and now he works at a warehouse.
APPENDIX K

INQUIRY SESSION WITH 5TH GRADERS:
"Drop the Dollar"
INQUIRY SESSION WITH 5TH GRADERS:
"Drop the Dollar"

AW: Now, if you can catch this dollar between your finger and thumb you can keep the dollar.

S: Oh.

AW: Too bad. You were close.

S: Can I try it?

AW: OK, I'll try it once more and that's all.

S: Oh.

AW: OK. Now let's try to figure out why these two guys couldn't catch the dollar before it passed through their fingers.

...(Long interruption while the same experiment was done with everybody in the class).....

S: Would it work with just paper?

AW: Yes.

S: Does the size have something to do with it?

AW: What do you think?

S: I don't know for sure. If it was bigger could we catch it?

AW: Let's try it.

S: Yah, I got it.

AW: Does that help you?

S: If you told me...like on your mark, get set, go, would I be able to catch it?

AW: Let's try it.

S: Yeah.

AW: I imagine you want me to try it again for keeps, but I won't.

S: Did you make this up?

AW: No, I had it tried on me too.

S: Did you get it?

AW: No, I missed it too.

S: Do you know why?

AW: I think so.

S: Will you tell us?

AW: No, I think you can figure it out.

S: Does she know why?

AW: Your teacher, I think so.

S: Then she'll tell us when you go.

AW: Oh, I don't think you people should give up so easy. Is this that hard of a problem?

S: Did you hypnotize us or something like that?

AW: No.

S: Did you ever loose a dollar to anybody?

AW: Yes.

S: Who?

AW: Oh, a boy in another school, a long time ago.

S: Did you let him keep it?
Yep.

**S:** You know how to drop it, that's it, isn't it?

**AW:** Do you want to drop it for me or somebody else?

**S:** OK.

**AW:** Now do you know some special way of dropping it?

**S:** No, I just dropped it.

**AW:** Well, how do you explain it?

**S:** I dunno.

**AW:** What questions do you have that might help you out?

**S:** Could you practice and then get it?

**AW:** I think you could practice and have a better chance, but I don't know if you'd get it every time. You'd probably get it some times.

**S:** If you used real light paper, could you get it?

**AW:** You'd probably have a better chance with something like tissue paper.

**S:** Why's that?

**AW:** What do you think?

**S:** I don't know.

**S:** Because it falls slower?

**AW:** Sort of.

**S:** Is it because we don't know when you're going to drop it and we're slower?

**AW:** That's kind of a theory isn't it? How can you test it out?

**S:** I don't know.

**S:** If you had the dollar higher would it work?

**AW:** Let's try it. You show me what you mean.

**S:** Do I get to keep it?

**AW:** Not this time.

**S:** Oh, but you said....

**AW:** OK, then you got to let me try to get it back. The same way.

**S:** That's not fair.
APPENDIX L

INQUIRY SESSION WITH 9TH GRADERS
"The Rosenthal Effect"
AW: Has everybody finished reading over the paragraph? OK then let's look at the chart. Does this prove or show that some rats are much smarter than other rats?

S: Yah.

AW: Does the same thing apply to people?

S: (inaudible)

AW: This seems like a pretty good experiment doesn't it? I mean you know about experiments with controls and everything. But let me throw a monkey wrench into it now and tell you that all of these rats came out of the same litters. You know what litter-mates are?

S: Like brothers and sisters?

AW: Right.... So what I'm saying is that the "dumb" rats and the smart rats were actually all from the same parents. Now what does that do to this experiment?

S: There wasn't any difference then in the rats. They were all the same, treated just the same?

AW: Yes, they were all born at the same time and were fed the same way, lived in the same kind of cage, and stuff like that.

S: Maybe the mazes were just harder for some of them then?

AW: All the mazes were the same.

S: Maybe it was just the way it happened.

AW: This was tried again, and the same thing happened.

S: Well, I don't know. It just happened.

S: You say everything was just the same, no different food or anything?

AW: No, everything or all the rats had the same food, same kind of cages, same water, ran the same mazes.

S: I still don't get it.

AW: Well, let me try to explain it this way. If I took a set of identical twins and separated them--twin rats I mean. And I gave you one and I kept one and then we tried this experiment to see which one was smartest and which dumbest And your rat seemed to be smartest, but really wasn't any different from my rat, but you didn't know that he was my rat's brother--twin brother. Then we'd have the same kind of experiment that Dr. Rosenthal tried. OK?

S: Well, then my rat would just be the smartest rat and that makes sense.

AW: OK. But now let's try this with maybe 20 rats. I'll give you guys 10 and tell you they're from real smart parents. Then I'll give you people over here another 10 and tell you they're from dumb parents. Everybody who has a smart rat then
proves how smart his rat is and everybody with a dumb rat
proves how dumb his rat is. See what I mean? How come this
happens everytime?

S: You say if I had a dumb rat and I was told he was dumb he'd
be dumb?

AW: That seems to be happening in this experiment.

S: Is that like could it be done with dogs too?

AW: Probably.

S: Well, then maybe identical twins or whatever aren't really
identical. There's some difference.

AW: According to what I remember there is no genetic difference
in identical twins.

S: Well, then I don't know. I couldn't figure it out.

AW: OK. Does anybody have an idea?

....(long pause)....

AW: Well, would you like it if I said we people up here in the
front are the brains and the people over in the back are a
little short on brains?

S: Well, people are different. Some are smarter than others.

AW: OK. But how do you know that?

S: Well, look at the grades or how some people can learn to
do math real easy and some can't. And you know some people
are real good in baseball and other guys aren't.

AW: OK. But these people are all different genetically. Nobody
in this room is exactly the same as some other guy. But if
everybody was and then we split the room right down the middle
and found out that the smart side people all were smart and
the dumb side people were all dumb, then we'd have the same
kind of experiment as Rosenthal.

....(long pause)....

AW: Let me try it this way. Say I take 6 peas out of a pea pod.
Each one is exactly the same. I give you 3 peas and tell you
to plant them in a pot of pure sand. I then give you 3 peas
and tell you to pot them in pure sand too. Each day we water
the peas with the same amount of water, and both pots are put
right over there on the window sill. Each guy treats the peas
the exact way, OK? The only difference is that I tell you
you've got fast growing seeds, but I tell you you've got real
slow growing seeds. You don't know that all the peas came
out of the same pod. Well, after while your peas are about
up to the ceiling, and your peas have hardly sprouted. See
what I mean.

S: Would that really happen?

AW: I kind of doubt it with peas, but it happened with rats.

S: It just doesn't seem right. I mean the only thing different
is that you said one thing to her and another to me. Every­
thing else is the same.

AW: Right. Now that's puzzling. How can you explain the results?

S: Then you shouldn't tell anybody anything about anything and
just see what you get.

AW: OK. That's been done.

S: What happened then?

AW: Everything was just about equal. The rats all did about the
same. Hardly any difference.

S: They did it again with the same rats?

AW: No different rats, but also litter-mates.

S: That's weird.

AW: True. That was the problem Rosenthal tried to explain.

S: Did he get an answer?

AW: Kind of. That's what I'm trying to get you to figure one out too.

S: Just by telling people that they got smart rats they prove that they got smart rats, even though they don't?

AW: That's what seemed to have happened.

S: Well, that means it was a bad experiment. You shouldn't tell anybody anything if you want to find out something.

AW: You think that the experiment was goofed up by his telling then?

S: Yah, must have been something like that.

AW: OK. Any other ideas?
REFERENCES


Bartolome, P. I. "Teachers' Objectives and Questions in Primary Reading." Reading Teacher, 23 (October, 1969), 27-33.


Biddle, Bruce J. "Facets of Teacher Role Research," 1968. (Mimeographed.)


Butts, David P., and Raun, Chester E. "The Relationship Between the Strategies of Inquiry in Science and Student Cognitive and Affective Behavioral Change." Science Inservice Project Research Report No. 4, Research and Development Center for Teacher Education, University of Texas at Austin, n.d. (Mimeographed.)


---


Guszak, F. J. "Teacher Questioning and Reading." Reading Teacher, 21 (December, 1967), 227-34.


Houston, Victor M. "Improving the Quality of Classroom Questions and Questioning." Educational Administration and Supervision, 24 (January, 1938), 17-28.


Hunter, Elizabeth. "The Effects of Training in the Use of New Science Programs Upon the Classroom Verbal Behavior of First Grade Teachers as They Teach Science." Classroom Interaction Newsletter, 4 (May, 1969), 5-11.


"Illinois Studies of Inquiry Training." N. D. (Mimeographed and dittoed materials.)


Morse, Kevin R., and Davis, O. I., Jr. "The Questioning Strategies Observation System (QSOS)." Report Series No. 35. Austin: The Research and Development Center for Teacher Education, The University of Texas at Austin. (Mimeographed.)


Smith, B. Othanel, and Meux, Milton in collaboration with Coombs, Jerrold; Eierdam, Daniel; and Szoke, Ronald. *A Study of the Logic of Teaching.* Urbana, Ill.: University of Illinois Press, 1970.


