

AN ABSTRACT OF THE THESIS OF

Robert Steven Roberts for the degree of Doctor of Philosophy
in Science Education presented on May 12, 1980
Title: Student Nonverbal Behavior and Frustration in Individual
Test Conditions

Abstract approved: _____

Redacted for privacy

Thomas P. Evans

The goal of the study was to investigate the relationship between nonverbal behaviors and frustration. As a part of the study an instrument for the observation of nonverbal behaviors related to frustration was developed.

The specific hypotheses investigated in this study were the following:

Hypothesis One - A valid and reliable instrument can be constructed for the systematic observation of student nonverbal behaviors related to frustration.

Hypothesis Two - There is a significant difference in the nonverbal behavior of students performing frustrating and nonfrustrating tasks.

Hypothesis Three - There is a clustering of nonverbal behaviors elicited in response to frustrating tasks; this clustering represents different student response modes to the frustrational stimulus.

In the pilot phase of the study, videotapes of students performing frustrating and nonfrustrating tasks were viewed and analyzed to identify nonverbal behaviors possibly related to frustration. The literature was reviewed to identify other behaviors that might be induced by frustration. The two major outcomes of the pilot phase were (1) the selection of the tasks used to induce frustration and

(2) an observational instrument that included a listing and definition of 24 behaviors to be viewed and tabulated, a record sheet, and instructions for the use of the instrument.

In the main study, a sample of 38 eighth grade students was videotaped during the administration of two individualized tests, the Peabody Picture Vocabulary Test (PPVT) and the mathematics subtest of the Peabody Individual Achievement Test (PIAT). Later, the behavior of each student was viewed and recorded using the observational instrument. Data were taken for two samples of behavior in each test, one sample for each test under non-frustrating conditions and one under frustrating conditions. The determination of whether the testing was frustrating or non-frustrating was made on the basis of the amount of success or failure in the student's responses. Non-frustrating conditions were seen at the start of the testing where students responses were usually all correct. Frustrating conditions were seen at the end of the testing where student responses were usually less than 20 percent correct.

The results of the study demonstrated that a valid and reliable instrument for the observation of student nonverbal behaviors can be constructed. Interobserver agreement was measured as the reliability of the instrument and found to be 0.88 overall. Also the instrument was sufficiently sensitive to detect the changes in nonverbal behavior that were hypothesized in hypothesis two.

The study did show that there was significant change in student nonverbal behavior from the non-frustrating to the frustrating portions of both tests. Further, there were strong parallels between the behavior changes seen in the two tests. The strongest behavior changes for both tests included longer times to answer questions, raising and lowering of the eyebrows, mouth twitching, mouth opening, and subvocalization.

The data from this study did not conclusively demonstrate the hypothesized clustering. Although some clustering was seen in some of the frustrational behavior data, there was a lack of consistency

in the clustering, and the sample size was considered marginal. Overall, the data did not clearly demonstrate the hypothesized modes of frustrational behavior.

Student Nonverbal Behavior and Frustration
in Individual Test Conditions

by

Robert Steven Roberts

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

June 1980

APPROVED:

Redacted for privacy

Professor of Science Education
in charge of major

Redacted for privacy

Chairman of Science Education

Redacted for privacy

Dean of School of Education

Redacted for privacy

Dean of Graduate School

Date thesis is presented May 12, 1980

Typed by Deanna L. Cramer for Robert Steven Roberts

TABLE OF CONTENTS

<u>Chapter</u>		<u>Page</u>
I	INTRODUCTION	1
	The Problem	3
	Definition of Frustration.	3
	Development of an Instrument	4
	Application of the Instrument.	5
	Outline of the Study Design	6
	The Pilot Phase.	6
	The Main Study	7
	Assumptions	8
	Limitations	8
	Delimitations	8
	Definition of Terms	9
	Organization of the Remainder of the Study.	11
II	LITERATURE REVIEW.	13
	The Literature of Nonverbal Communications and Nonverbal Behavior.	13
	Nonverbal Communications as an Inter- disciplinary Study	13
	Considerations in the Definition of Nonverbal Communications	15
	Nonverbal Communication Versus Nonverbal Behavior.	16
	Definition of Nonverbal	17
	The Shannon-Weaver Model of Communications.	19
	Communication Specificity	21
	Awareness and Intentional Control	22
	Psychophysiological Responses	24
	Major Research Studies and Strategies in Nonverbal Communications	25
	Early Research.	25
	Facial Communication of Emotion	26
	Studies in Psychiatry and Counseling.	27
	Ethological Approaches.	28
	Frustration Theory.	29
	The Rosenzweig Picture-Frustration Test.	30
	The Study of Barker, Diempe, and Lewin	31
	The Frustration-Aggression Hypothesis.	32
	The Frustration-Fixation Hypothesis.	32
	Frustration and Behaviorism.	33
	Common Frustrational Tasks	35
	Student Nonverbal Behavior and Frustration.	36
	Betts and Frustration in Reading Education	36
	The Study of Jecker, Maccoby, and Brietrose.	39
	The Studies of Allen and Others on Decoding Ability	41

Table of Contents -- continued

<u>Chapter</u>		<u>Page</u>
	Eggen's Study.	41
	Dittman and the Development of Nonverbal Behavior Skills.	42
	Summary of the Related Literature	43
III	DESIGN OF THE STUDY.	44
	Overview of the Study Design.	44
	Phase I - The Pilot Study Plan	45
	Phase II - The Main Study Plan	46
	The Pilot Study	47
	Definition of Frustration.	47
	Pilot Videotaping and Selection of Frustrational Tasks	47
	Construction of the Observational Instrument	49
	Listing Important Nonverbal Behaviors	49
	Form of the Observational Record.	51
	Quantification and Definition of the Nonverbal Behaviors	52
	Assembly of Rough Draft	53
	Additional Comments on the Instrument	53
	The Main Study.	55
	Selection of Subjects.	56
	Gathering Data	57
	Use of the Observational Instrument.	58
	Validity and Reliability	59
	Validity of the Observational Instrument -- Some Arguments	61
	Description of the Data.	63
	Data Analysis.	66
	Hypothesis One.	66
	Hypothesis Two.	67
	Hypothesis Three.	68
	Summary	69
IV	RESULTS OF THE STUDY	73
	Findings Related to the Hypotheses.	73
	Hypothesis One	73
	Reliability and Interobserver Agreement	73
	Validity of the Instrument.	74
	Summary of the Results for Hypothesis One	76
	Hypothesis Two	77
	Hypothesis Three	80
	Findings Not Directly Related to the Hypotheses	86
	Correlation Matrices	87
	Discriminant Analysis of the Data.	90

Table of Contents -- continued

<u>Chapter</u>		<u>Page</u>
	Other Nonverbal Behaviors Observed	91
	Summary of the Results.	93
V	SUMMARY, CONCLUSIONS AND RECOMMENDATIONS	95
	Summary of the Study.	95
	The Problem.	95
	Design of the Study.	96
	Results of the Pilot Phase of the Study.	98
	Results of the Main Study.	98
	Conclusions and Recommendations	99
	The Relationship of the Current Findings to	
	Other Studies.	99
	Betts' Work and Informal Diagnosis of	
	Reading	99
	The Study of Jecker, Maccoby, and	
	Brietrose and Others.	100
	Comparison of Results for the Two Tests -	
	PPVT and PIAT.	102
	Discussion of Significant Nonverbal Behaviors	
	Observed in the Study.	104
	Micromomentary Facial Expression.	105
	Mouth Twitch.	106
	Lean Back	106
	Body Movements.	106
	Manipulation.	107
	Gaze.	108
	Eyebrow Flash	109
	Eyebrow Lowering.	109
	Smiles.	110
	Mouth Dropping.	110
	Subvocalization	111
	Shoulder Shrug and Head Cocking	111
	Lip Licking	112
	Recommendations for Further Study.	112
	Replication Studies	113
	Similar Studies	113
	Instrument Modifications.	114
	Other Studies	115
	BIBLIOGRAPHY	117
	APPENDICES	
	I. Interobserver Agreement	125
	II. Notes on the Interobserver Agreement	
	Calculations	128

Table of Contents -- continued

<u>Chapter</u>	<u>Page</u>
III. Population Description.	130
IV. The Jecker-Maccoby-Brietrose Checklist.	131
V. Parental Consent Letter	132
VI. The Clock-Timer	133
VII. The Observational Instrument.	137
VIII. Relative Frequency of Behaviors Observed.	147
IX. Notes on the Clustering Program	149
X. Diagrams and Notes on the Equipment Setup During the Main Data Collection	151

LIST OF FIGURES

<u>Figure</u>		<u>Page</u>
1	Points of view in nonverbal communications literature. . . .	14
2	Communications theory frame of reference	15
3	The Shannon-Weaver model of communications	20
4	Behavioral characteristics associated with reading at three levels	38
5	Example of clustering into two groups displayed by scatter diagrams of three variables taken two at a time.	70
6	Example of clustering into two groups displayed by a profile analysis of three variables.	70
7	The process of judging the cluster analysis results.	81
8	Profile plot and verbal description.	82
9	Discriminant plot for all variables.	83
10	Discriminant plot for seven variables.	84
11	Other nonverbal behaviors observed that may be related to frustration.	92

LIST OF TABLES

<u>Table</u>		<u>Page</u>
1	Distribution of the sample population by sex and mathematics class grouping	57
2	Results for test of behavior change using T-squared test . .	77
3	Results of individual t-tests for behavior change.	79
4	PIAT data correlation matrix	88
5	PPVT data correlation matrix	89
6	Discriminant function classification accuracy results for PIAT and PPVT data	90
7	Comparison of behaviors elicited by the PIAT and PPVT. . . .	103
8	Average values for QPM and TTA for the PIAT and PPVT	103

STUDENT NONVERBAL BEHAVIOR AND FRUSTRATION IN INDIVIDUAL TEST CONDITIONS

I. INTRODUCTION

Considerable recent research has involved the systematic observation of classroom behavior (Medley and Metzger, 1963; Simon and Boyer, 1965-1970). In the majority of this research, however, the focus of the observation has been teacher behavior. This interest in teacher behavior is probably the result of concerns over teacher evaluation, and the need to evaluate changes in teaching styles as a function of curriculum changes and inservice training. Many of the instruments used in these studies were limited to the observation of verbal behaviors and were similar to, or modifications of, the Flanders System (Flanders, 1960).

There is a need for research and development of observational instruments that focus on student behaviors in the classroom. A major justification for such research is the need to carefully identify the student behaviors that can best provide feedback communications to the teacher. With such feedback information, the teacher can more effectively make both immediate and long-range changes in curriculum or methods to better meet the needs of the students.

Work on nonverbal communications has shown that observation of nonverbal behavior can give reliable and helpful information about a person's affect states and affect responses (Buck, 1975; Shapiro, 1968; Ekman, 1975; Argyle, 1975). But few researchers have systematically observed student nonverbal behaviors in the classroom for the purpose of judging student affect.

One notable exception to these studies is that of Jecker, Maccoby, and Brietrose (1964, 1965). One major goal of their study was to determine if teacher judgement of nonverbal behaviors could be improved through inservice training. Their instrument or system for observation, however, was not specifically designed for systematic observation of classroom behavior but was designed to serve as a training tool to improve teacher judgement of classroom nonverbal behaviors as

indicators of student comprehension. The results of their study indicate that (Jecker et al., 1965, p. 244):

....the ability to judge student comprehension on the basis of nonverbal cues, while largely undeveloped in the normal classroom teacher, can be improved by training in the recognition of such nonverbal cues.

It seems reasonable to generalize that teacher ability to judge non-verbal cues may be undeveloped not only in the ability to judge student comprehension but in other areas as well.

This generalization is supported by related research. Raymond (1971) was able to show that teachers were inaccurate in their judgement of student affect as defined in terms of the students' attitude towards the class and teacher. Eggen (1972) showed that student non-verbal behavior is significantly related to the students' attitude towards the class and the teacher. If these two studies are considered together with the findings of the Jecker, Maccoby and Brietrose study, it is easy to conjecture that teachers' judgements of student attitude might be more accurate with some training in the observation of non-verbal behaviors.

In the classroom, one common circumstance that may serve as a stimulus of affect is student frustration. Educators have pointed out the negative effects of frustration on student achievement and self concept (Phillips, 1978). They claim students need to be provided with classwork that is challenging but not frustrating. Jecker, Maccoby, and Brietrose (1965) point out that in the regular classroom situation the teacher must rely predominantly on nonverbal feedback. Because frustration is a stimulus for affect and affect is associated with non-verbal behavior, the teacher may be able to monitor student frustration through observation of nonverbal behaviors. Unfortunately, little research has involved observation of student nonverbal behaviors (Galloway, 1970) and no study could be found that is directed specifically at the observation and investigation of the relationship between student frustration and nonverbal behavior.

The above discussion has identified the following facts that are related to the need for the present study:

1. Little research has involved the systematic observation of student behaviors in the classroom. Even less has focused on the observation of student nonverbal behaviors.
2. Detection of student frustration in the classroom is important for effective teaching.
3. Theory suggests that student frustration may be detected through observation of student nonverbal behaviors.
4. Further research into the relationship between student frustration and nonverbal behavior may provide useful information both to the teacher concerned about detecting frustration in the classroom and to the investigator of nonverbal behavior.

In response to the above needs, the main goal of this study is the systematic observation of student nonverbal behaviors elicited in response to frustrating tasks. The first part of this study involved the construction of an instrument for such observation. Through the use of this observational instrument, the study investigated the relationship between student nonverbal behavior and task frustration.

The Problem

The central problem of the study is the systematic observation of nonverbal behaviors elicited in response to student frustration. The problem can be broken down into three subproblems:

1. Development of an operational definition of frustration;
2. Development of an observational instrument for systematic observation of nonverbal behaviors related to frustration; and
3. Application of the instrument to a sample population both to validate the instrument and to obtain information on the relationship between nonverbal behavior and frustration.

Definition of Frustration

In educational literature the term frustration is generally used in the vernacular sense of an emotion or internal state. Psychologists,

however, have usually required a more exact definition of the term frustration. In addition, because much of their work has been performed using rats and cats, they have needed a definition which does not depend on the ability of the subjects to verbally report their emotional state. Psychologists have therefore often defined frustration in behavioristic terms as the blocking of goal attainment where the goal attainment is at least initially expected (Lawson, 1966). In this way frustration is treated as a stimulus condition. Affect responses to frustration can be called frustrational behavior but are not necessarily unique to the frustration stimulus conditions. For example, anger could be an affect response to frustration, but of course not all anger is frustration caused and therefore we would not label all anger as frustrational behavior. Although this definition of frustration is more precise and may appear to be quite different from the more common usage of the educator, it does hopefully retain much of the same meaning as the vernacular from which it evolved.

Development of an Instrument

The first step in the systematic observation of student nonverbal behaviors related to frustration must be the development of a valid and reliable instrument for such observation. A central concern in such an effort is the development of a list of nonverbal behaviors that are most communicative of the frustrated affect state. Previous research provides little assistance in this effort because (1) few instruments have been developed, and (2) research using these instruments indicates that the significant nonverbal behaviors differ from affect state to affect state. The nonverbal behaviors that are listed in the Jecker, Maccoby and Breitrose (1964, 1965) checklist of student behaviors are different from the behaviors identified as important in Eggen's study. This seems to imply that the nonverbal behaviors that are related to student comprehension are somewhat different from those that are related to student attitude. This conjecture is supported by Mehrabian (1972) who concludes that gross affect, such as positive and

negative attitude, may be more related to posture and position cues but that specific affect or transitional emotional reactions such as fear, anger, or joy, is more associated with facial cues or body movements. Ekman and Friezen (1975) have developed a system based on the human face that uses nonverbal behaviors to differentiate between six emotion categories. The following hypothesis therefore focuses on this part of the problem:

Hypothesis One. A valid and reliable instrument can be constructed for the systematic observation of student nonverbal behaviors and behavior changes in response to frustration.

Application of the Instrument

The testing of the validity and reliability of the observational instrument requires the application of the instrument to a sample population. Such application will also yield information on the relationship between nonverbal behavior and frustration. The first step in the investigation of both validity and the relationship between nonverbal behavior and frustration is to confirm that there is some change in student nonverbal behavior as a result of frustration. This leads to the statement of the second hypothesis of this study:

Hypothesis Two. There is a difference in the nonverbal behaviors elicited by students performing tasks that are frustrating and those tasks that are non-frustrating.

Both theoretical and experimental work on frustration suggests that frustration acts as a stimulus for a variety of affect states or affect reactions. The affect that is induced by frustration is a function of psychological type, situation, age, past experience, and other variables. Depending on the student and the situation, frustration may cause a variety of responses. Maier and Ellen (1965, p. 102) state

the situation succinctly

Frustration theory postulates that behavior that is constructive, motivated, problem-solving, or goal oriented in nature may suddenly be replaced by behavior that is hostile (aggression), immature (regression), stubborn (fixation), apathetic (resignation), or some combination of these...

Maier and Ellen further state that such reactions suggest that a bimodal distribution of behavior data might exist in studies of frustration. This distribution, they suggest, might be labeled constructive behavior versus frustrated behavior, even though both types of behavior are frustration induced (Maier and Ellen, 1965).

Because of this variety of different affect responses, it is conjectured that some of these different modes of response would be seen in student behavior and that analysis of the data should attempt to differentiate between different modes of response. This leads to the third and last hypothesis of this study:

Hypothesis Three: There is significant clustering of measured nonverbal behaviors elicited in response to frustrating tasks.

Outline of the Study Design

The study was divided into two main parts or phases, the pilot phase and the main study. Each of these phases are described below.

The Pilot Phase

The pilot phase was principally concerned with the design of an instrument. It included a literature review to identify nonverbal behaviors other researchers have investigated and have shown to be related either directly to student frustration or were related to affect responses that are theoretically related to frustration. Videotapes were made of a variety of students in a variety of situations involving frustrating tasks. Observation of these videotapes was aimed

at identifying and defining as many as possible of the observed non-verbal behaviors. At this time the mathematics subtest of the Peabody Individual Achievement Test (PIAT) and the Peabody Picture Vocabulary Test (PPVT) were selected as the tasks to elicit frustration.

At the end of the pilot study, the observational instrument was constructed, and plans were made for the main study. Several additional videotapes were made to simulate the conditions of the main study and to work out technical problems related to the videotape recording of the individualized testing. The devised observational instrument was applied to all of the pilot videotapes to help make the definition of behaviors objective and unambiguous.

The Main Study

In the main study, 39 eighth grade students were used. Each of the students was videotaped during the administration of the two individualized tests, the PIAT mathematics subtest and the PPVT. The videotape records of this testing were systematically observed with the observational instrument. Observations were limited to a one minute segment of pre-frustrational behavior and a one minute sample of frustrational behavior for each test. Data from the main study were used to measure the interobserver agreement for the instrument. This interobserver agreement is commonly used as a measure of the reliability of an observational instrument.

The data from the pre-frustration and frustrational portions of the testing were compared to determine the significance of behavior changes and which specific nonverbal behaviors were most strongly related to frustration. Finally, the frustrational behavior data were submitted to cluster analysis in an attempt to identify specific modes or different types of affect reaction elicited by the frustrational testing.

Assumptions

The following assumptions are inherent in the study:

1. Nonverbal behavior is a significant part of a person's total behavior that communicates and augments verbal communications in a significant way.
2. There exists enough consistent nonverbal behavior that can be identified and measured so that such behavior is amenable to study.
3. Much of nonverbal behavior operates in such an unconscious or automatic way that efforts to conceal feelings and attitudes have limited effect on such behaviors.
4. Frustration is the cause or stimulus for one or more unified and dominant emotional states that are communicated through nonverbal cues.

Limitations

The following limitations apply to this study:

1. The study will be limited to the use of available and willing teachers and their students in the Corvallis, Oregon area.
2. The study is limited to the extent that the selected tasks represent tasks typical of classroom behavior.
3. The presence of a video-recorder and its operator may affect behaviors in the classroom.
4. Observations and measurements are limited by the technical parameters of the observational equipment and the limitations of observer preception.

Delimitations

The following delimitations apply to the study:

1. The study will consider only selected student nonverbal behaviors which will be selected for systematic observation by the following

criteria:

- a. They are easily observable in the classroom setting;
 - b. They lend themselves to measurement in the qualitative or quantitative sense with minimal instrumentation; and
 - c. They show promise of being somehow related to or predictive of frustrational behaviors as demonstrated in the pilot study.
2. The study will make no attempt to examine mechanisms or models for frustration or frustrational behavior. No attempt will be made to name the emotions associated with behaviors that are induced by frustration.
 3. No attempt will be made to change student behavior beyond the manipulation of the tasks.

Definition of Terms

The following definitions are relevant to the study:

1. Frustration - The term frustration is used to denote a situation in which a student is blocked from the attainment of a goal that the student is motivated to achieve and at least initially expects to achieve.
2. Task frustration - Task frustration will denote frustration that is caused when a student is motivated to complete a task but is blocked from completing due to some aspect of task difficulty.
3. Gross affect - The term gross affect is used synonymously with the term attitude.
4. Specific affect - Specific affect denotes more time dependent and transient affect states. The term is used synonymously with the term emotion.
5. Tasks - The term task will be delimited to certain specific tasks after the pilot study. During the pilot study, tasks considered will include but not be limited to:
 - a. Questions from selected individualized tests such as the Peabody Individual Achievement Test (PIAT);
 - b. Mathematical puzzles;

- c. Individualized instructional materials including laboratory instructions; and
 - d. Computerized questions using an interactive computer programming such as "Dialog".
6. Frustrational behavior - Behavior that occurs simultaneous with and just following task frustration.
 7. Nonverbal behaviors - Behaviors other than the explicit verbal communication of speech. Nonverbal behaviors include primarily body postures and movement including facial expression.
 8. Nonverbal communications - Where nonverbal behaviors are related to some message or aspect of a person's state of mind, the nonverbal behaviors are defined as communicating. Nonverbal communications assume only that nonverbal behaviors are related to some message and that therefore detection of an implicit message is possible.
 9. Anxiety - The experiencing of fear in the absence of any objective reason for fear or a reaction of fear where the source of fear is not identified or not specific.
 10. Frustrational testing - Testing where the student experiences a large degree of failure. Usually testing at the end of a "power" type of placement test, where questions are ordered from easy to hard, is frustrational, since the testing continues until a ceiling level is indicated. The ceiling level is usually defined in terms of a high student error rate in a sequence of questions. For the PIAT and PPVT, the ceiling level is defined as five wrong out of seven or six wrong out of eight, respectively. In this study the frustrated behavior sample observed was taken from this ceiling level sequence.
 11. Peabody Individual Achievement Test (PIAT) - An individualized achievement test designed to obtain quick screening and placement information on a student's achievement in reading, mathematics, spelling, and general information. In this study only the mathematics subtest of the PIAT was used in the main study. Where

the PIAT is denoted, the mathematics subtest of the PIAT is implied.

12. Peabody Picture Vocabulary Test (PPVT) - An individualized test used to assess a student's recognition vocabulary but more often to gain a quick assessment of a student's Intelligence Quotient (IQ). In this study the PPVT was used as one of the tasks of the main study testing. The testing was used only as a stimulus of frustration and the obtained IQ scores were not used in the study.
13. Pre-frustration - The state during the start of the individualized testing where the student is given easy questions and exhibits correct responses. Usually at the start of individualized "power" type placement tests, questions are easy and the testing is specifically started so that a "basal level" is indicated. This basal level is usually defined in terms of a sequence of correct answers. For the PIAT and PPVT the basal levels are defined as five and eight consecutive correct answers respectively. During this study pre-frustrational behavior observation usually included this basal level sequence.

Organization of the Remainder of the Study

The next chapter, Chapter II, is devoted to the background and related literature and is divided into three sections. The first section reviews nonverbal communications and nonverbal behavior literature with a focus on the definition on the scope and limits of nonverbal communications research. It includes a review of common communications research models, vocabulary, and theory. The second section reviews the psychological theories of frustration and research related to frustration. The third section examines the work performed in educational research related to student nonverbal communications.

Chapter III presents the design of the study. The chapter is divided into three sections - an overview of the study, the methods used in the pilot study, and finally the design of the main study. The discussion of the pilot study focuses on (1) the selection of the

pilot sample, (2) pilot videorecordings, (3) the development of the observational instrument, and (4) a description of the observational instrument. The discussion of the design of the main study includes consideration of (1) selection of the subjects, (2) data gathering, (3) selection of the behavior sample, (4) measurement of reliability, (5) validity of the observational instrument, and (6) data analysis.

Chapters IV and V present and discuss the results of the study. Chapter IV describes the analysis of the data, presents the results of the analysis and discusses the results relative to the hypotheses. Results not directly related to the hypotheses are included. Chapter V is devoted to a summary of the study, to conclusions and to recommendations for future research. It further examines the relationship of the findings of this study to other studies.

II. LITERATURE REVIEW

This chapter is divided into three major sections. The first section reviews some important research findings and issues within the field of nonverbal communications. It is important both to put the present study into context and to establish a basis for the definition of the two important terms, nonverbal communication and nonverbal behavior. The definition of these terms is especially important in laying the groundwork for arguments on the validity of the observational instrument developed in this study.

The second section reviews the literature related to frustration. Again, this review is important to put the study into context and to establish the definition of the term frustration as it is used in this study.

The last section reviews the studies that more directly relate to the present study. The goal is to provide a review of studies that are most directly related to the problem of recognition of student frustration through the use of nonverbal behavior cues. The most similar studies are presented in more detail so that the methodological and theoretical similarities and differences to the present study are apparent.

The Literature of Nonverbal Communications and Nonverbal Behavior

Nonverbal Communications as an Interdisciplinary Study

This section of the chapter reviews the literature of nonverbal communications. This review focuses on issues that help to put the present study into context and defines some terminology that is commonly used in the nonverbal communications literature. Especially important are the definitions of the terms nonverbal behavior and

nonverbal communications which are essential to the construction of the observational instrument used in this study.

One of the problems in the review of all nonverbal communications literature is the wide variety of sources of relevant literature. As Harrison states, "...there is a growing -- although far-flung -- body of research, representing several frames of reference" (Harrison, 1972).

Harrison lists some of the frames of reference and their major interest areas (Figure 1).

<u>SOURCE</u>	<u>POINT OF VIEW</u>
Natural Sciences	Pan-human
Anthropology	Cultural
Sociology	Social
Psychology	Individual

Figure 1. Points of view in nonverbal communications literature (adapted from Harrison, 1972).

To show how each of these frames of reference has contributed to nonverbal communications literature, Harrison gave examples of major contributions from each field. Darwin represented the natural sciences; Hall and Birdwhistle represented anthropology; Goffman is a sociologist; and Davitz and Ruesch are psychologists.

Another totally different frame of reference that can be used to view and organize the field of nonverbal communications is that of communications theory. The distinction between communications theory and the views of psychology, sociology, anthropology, and the natural sciences listed above is primarily the difference between a focus on process versus content in the communications. In fact, Harrison sees the communications theory frame overlapping the other fields (Figure 2).

Another important source of literature relevant to nonverbal communications theory is that of the theory of emotion and the theory of the communications of emotion. Although much of the content of nonverbal communications is beyond purely emotional communication, it is clear that the communications of emotion is one of the most important

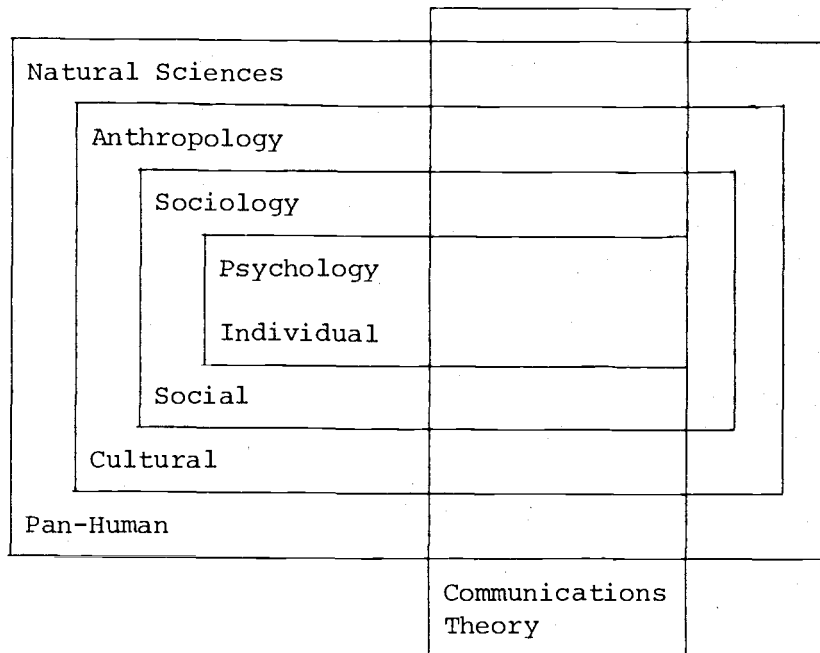


Figure 2. Communications theory frame of reference (adapted from Harrison, 1972).

functions of nonverbal communications. Some of the most important works in nonverbal communications literature are strongly interested in this area of emotional communications, such as The Communications of Emotional Meaning (Davitz, 1964), Interpersonal Messages of Emotion (Dittman, 1972), The Expression of Emotions in Man and Animals (Darwin, 1872/1965), Emotion in the Human Face (Ekman, Friesen, and Ellsworth, 1970), and The Expression of the Emotions in Man (Knapp, 1963).

Considerations in the Definition of Nonverbal Communications

Definition of the term nonverbal communications is not a simple task and there is no clear agreement among researchers. As Harper, Weins, and Matarrazo (1977, p. 2) stated,

...there have been a variety of approaches employed in the study of nonverbal communications and, as yet, there is no real consensus as to its exact definition, the domain that it encompasses, or what the best research approaches are.

To better understand and appreciate the problems of definition of nonverbal communication, several of the main issues and systems of terminology that separate the various views of nonverbal communication are given in this section of the chapter. The issues presented here not only introduce concepts and vocabulary that will be used in the definition of the present study, but lead to the definition of the terms nonverbal communications and nonverbal behavior (NVC and NVB). The definition of these terms are especially important since they are essential to validity arguments used for the observational instrument developed in this study.

Nonverbal Communication Versus Nonverbal Behavior

The definition of the terms nonverbal behavior and nonverbal communications vary among the various researchers and theorists. Some have stated that not all nonverbal communications involve behavior, and thus that nonverbal behavior is a part of nonverbal communications (Dittman, 1972). Barker and Collins state (1970, p. 344):

There has been a tendency to use the term nonverbal communication synonymously with the term nonverbal behavior. However, nonverbal communications is much broader than nonverbal behavior. A room devoid of behaving, living things communicates atmosphere and function. Static clothing communicates the personality of the wearer.

Others seem to feel that not all behavior communicates, and thus nonverbal communications is a term applied only to some nonverbal behaviors (Ruesch and Kees, 1956).

The point of view taken in this study was closer to the latter viewpoint, that not all nonverbal behavior communicate. That is, we may, for example, study a variety of nonverbal behaviors that we think are related to anxiety and find that statistical testing is only able to establish that a small number of the behaviors are actually related to anxiety. It could then be stated that these few behaviors nonverbally communicate anxiety. This point of view is dictated to some

extent by the statistical design of the study. The test of which nonverbal behaviors are associated with nonverbal communications of frustration must reflect the pragmatic consideration of how well the behaviors communicate the frustration of the student. That is, the question of which nonverbal behaviors provide nonverbal communications of frustration was the question ultimately answered and defined by statistical testing.

Definition of Nonverbal

Some researchers are very strict in the interpretation of the meaning of the term nonverbal. Mehrabian states (Mehrabian, 1972, p. 1):

In its narrow and more accurate sense, 'nonverbal behavior' refers to actions as distinct from speech....

Most researchers are not as strict in the definition as Mehrabian might like, however. The terms oral and verbal are slightly different and, strictly speaking, the term nonverbal allows some license in the definition of nonverbal behavior so as to include some aspects of oral behavior that are separable from the verbal message. Many of these aspects of nonverbal behavior are included under the term "paralanguage".

Paralanguage usually refers to the qualities of voice that communicate information beyond the verbal content. The most common research methods used in paralanguage research are the reading of content-free passages of reciting the alphabet while trying to indicate emotions of happiness, sorrow, anger, and so forth (Davitz, 1964). Other researchers have used filtering methods which make the verbal content unintelligible. Both methods have shown that some affective meaning can be conveyed through the dynamics of tone, amplitude, and timing.

Some researchers have further stretched the definition of nonverbal perhaps to its very limits to include research into speech disturbances and content analysis. Speech disturbances have been studied

by several investigators, especially in relation to stress or anxiety. George Mahl (1956) has developed a system for the classification and recording of eight kinds of speech disturbances including word repetition, word omission, and stuttering. Such speech disturbances are often included in the definition of nonverbal because like intonation, speech rate, and other dynamics, they communicate some message or augment the verbal message and provide information about the sender.

Content analysis attempts to find meaning within the content of the verbal message other than the explicit meaning. One major type of content analysis is contingency analysis. As described by Dittman, contingency analysis is a form of content analysis where (Dittman, 1972, p. 71-72):

...inferences are drawn not only from the total frequency of occurrence of the various content categories, but from the relative frequency of their occurrence in association with each other....The investigator can infer not only the fact of preoccupation with a topic but also some of the qualities of this preoccupation from the other topics with which it is associated.

Obviously including such contingency analysis into the definition of nonverbal communication may be stretching the definition to its limits.

Some researchers go even further in their unwillingness to separate verbal from nonverbal communications. Birdwhistle (1970), for example, believed that much of nonverbal communications is interrelated and tied to verbal communications so that study of nonverbal communications without considering verbal behavior is inappropriate. Knapp comments (1978, p. 3):

...It is not easy to dissect human interaction and make one diagnosis which concerns only verbal behavior and another which concerns only the nonverbal behavior.... Some of the most noteworthy scholars associated with nonverbal study refuse to segregate words from gestures and hence work under the broader terms communication or face to face interaction.

Interestingly, Mehrabian (1973) solved this problem of definition in his book Nonverbal Communications by abandoning the terms nonverbal

behavior and nonverbal communication in favor of the terms implicit communications and explicit communications.

This study used a relatively strict interpretation of nonverbal. Although it was recognized that by separating nonverbal from verbal parts of the communication most of the communication may be lost. Again this point of view was influenced by the design. The nonverbal behaviors used in the observational instrument were delimited to those behaviors that were more easily identified and measured in an objective manner without significantly affecting the subject. For this reason paralinguistic variables and content or contingency analysis were not extensively considered.

The Shannon-Weaver Model of Communications

Communications theory has adopted and adapted many models and concepts from electronic communications systems theory to describe the communications process. Although much of the electronic theory was developed for specific application to electronic communications, the success of the theory and the appeal of some of the concepts in the model has led to its great influence in the field of communications as a whole. In turn, much of the theoretical modeling of the process of nonverbal communications has been influenced by these same concepts and vocabulary. Much of this vocabulary is based on the so-called Shannon-Weaver model of the communications process.

The Shannon-Weaver model was first presented in Shannon and Weaver's, The Mathematical Theory of Communication (1949). The basic model is shown in Figure 3. In the nonverbal communications literature, the Shannon-Weaver model of communications is often assumed, and the vocabulary used in such literature includes terms from the model such as encoding, decoding, channel, source, receiver, and noise (Ekman and Friesen, 1972; Knapp, 1975; Dittman, 1972).

Dittman, in his book The Interpersonal Messages of Emotion (1972), has been the most ambitious in applying the Shannon-Weaver model to the modeling of a theory of emotion which is closely tied to the field of

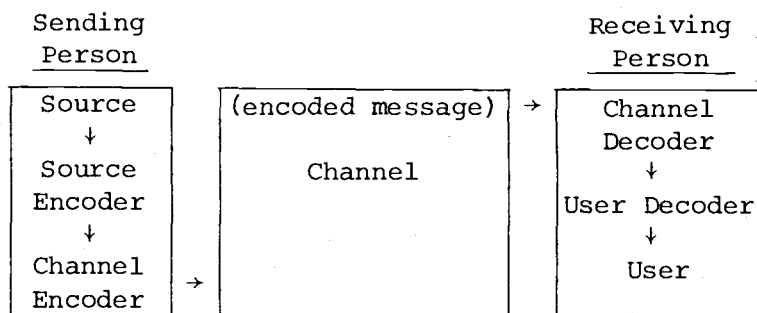


Figure 3. The Shannon-Weaver model of communications (adapted from Dittman, 1972, p. 24).

nonverbal communications. In his book, Dittman looked at the implications of concepts from the Shannon-Weaver model such as cross-talk, noise, multi-channel communications, interference, coding, and measurement of information. Although Dittman admitted that all of these concepts from the models of electronic communications did not adapt well to the modeling of the communications of emotion, many of the concepts were quite interesting and have stimulated much interest and discussion. His work has been commonly cited in current nonverbal communication literature.

To better understand how the Shannon-Weaver model is used in the nonverbal communications literature, several of the key vocabulary words arising from this model are explained below. These terms are commonly used in description of nonverbal communications processes and research.

"Encoding" and "decoding" are two terms which refer not only to the processes of encoding and decoding, but also to two types of experimental design. An encoding design is one in which the nonverbal behaviors are observed as dependent variables in the study. For example, subjects may be placed in social situations with someone they do not like and their nonverbal behaviors observed as the dependent variables. These behaviors may be contrasted with a control situation where the same people are observed in the same situation with people about whom they feel neutral or positive.

In contrast, in a decoding study, the dependent variable is some result of the decoding process, such as the receiver's inferred attitude, judgement, or feelings. For example, the experimental subjects may watch a videotape of actors who try to communicate several emotions nonverbally. The subjects are then asked to identify the emotions from their own judgement of the nonverbal behavior. Other decoding studies have used naturally occurring facial expressions or facial expressions that were the result of some controlled stimulus such as a story or film. Both motion picture film or videotape and still pictures have been used to provide the nonverbal stimulus.

Communication Specificity

Communicative specificity is a term that may have originated with Dittman (1972). Dittman referred first to a definition of the term "language" as "...a system of signs, universally applicable, whose use is agreed upon by members of a group" (Dittman, 1972, p. 78). In this definition, "system" implied the use of some definite organization, grammar or syntax in which the signs were used. Dittman then defined the term "communicative specificity" as a continuum in which language, or communications using a language base, is at one end of the continuum. Spoken language is the best example of communication high in communicative specificity. Perhaps a computer language would be one of the highest forms of communications in terms of this specificity.

Several interesting problems appear in the application of this concept of communicative specificity to nonverbal behavior. For example, if high communicative specificity is assumed then simple experimental designs are not appropriate. If communicative specificity is high, attempts should rather be made to identify the grammar and complex organizational relationships that exist. Argyle (1975, p. 15) has stated that "Investigators who believe that nonverbal communication is a structured system of communication like a language have used different, and non-statistical approaches."

Because of the simple nature of the statistical testing involved in the present study, no grammar or organization of the different non-verbal behaviors was assumed. No special sequence of behaviors was assumed or tested for. The degree of communicative specificity as defined above was therefore assumed to be low and was not tested in this study.

Awareness and Intentional Control

Dittman (1972) and Argyle (1975) specifically apply the question of awareness in the process of nonverbal communication to both the awareness of the sender and the awareness of the receiver. This question of awareness may also consider the degree of awareness and the question of awareness of the behavior versus awareness of what the behavior may communicate. The sender may or may not be aware that some behavior is significantly related to the emotion of anger for example and may or may not be aware when he or she is exhibiting that particular behavior. To the extent that people are not always aware of some behavior such as finger-nail biting, awareness can be considered to vary in degree.

It is certainly true that most adults are aware of certain non-verbal communications even if they are not always aware when they are used. Books such as Body Language by Fast (1970) and How to Read a Person Like a Book by Nierenberg and Calero (1971) have communicated some of the more interesting results of nonverbal communications research to the lay public. However, awareness of nonverbal behavior does not necessarily cause significant behavior change. It is true however that some interest in nonverbal behavior research stems from a desire to maintain more intentional control over nonverbal behavior. Awareness may be one step toward intentional control.

Both awareness and intentional control are important considerations in the theory and definition of nonverbal communication. For example, we may wish to control the communications either (1) by "acting happy" or (2) by smiling, talking more, and not frowning. In

the first case the control may be accomplished without analysis of what specific behaviors are being used. In the second case the control focuses on the control of specific behaviors.

Intentional control has been the subject of several investigations concerning how such control of nonverbal behavior can be used to deceive. Mehrabian (1972) and Ekman et al. (1970) have investigated deception. Both of these investigators were concerned with the phenomenon of "leakage", or the degree to which the subject's true feelings are expressed even when they are trying to hide their feelings. Another area of research related to intentional control is the important consideration of whether acted or posed nonverbal behaviors, which have often been used in nonverbal communications research, are significantly different from nonverbal behaviors used in everyday situations. In other words, is the use of actors valid in these experiments? This research seems to show that acted behaviors are quite similar to spontaneous behaviors but are probably significantly exaggerated and easier to judge accurately (Frijda, 1973; Allen and Atkinson, 1977).

Psychologists are especially concerned with the research on intentional control since the ideas of Freud and others indicate that we cannot conceal our feelings totally and that much of the leakage of our feelings is communicated nonverbally. Almost ironically, psychologists often attempt to consciously control their own nonverbal behaviors to manipulate the doctor-patient relationship and to facilitate the most therapeutic environment.

Intentional control and awareness are important considerations in experimental design because if nonverbal behaviors may to some degree be controlled by the subject, knowledge of the experimental situation can affect the experimental results. Questions of intentional control and awareness of NVC are delimited from the study although it was assumed that NVB of students is affected and somewhat limited by informally learned social constraints.

Psychophysiological Responses

Some or all psychophysiological responses can be considered to be within the definition of nonverbal communication. It is certainly clear that changes in heart rate, respiration, skin conduction, and other phenomena such as blushing do communicate something about the state of the individual. These responses, however, although less susceptible to intentional control, are low in communicative specificity. These same responses may occur under a variety of mental states, and therefore they may only supply one isolated bit of information that by itself communicates little and that can only be interpreted in conjunction with other information.

Perhaps the main problem with the inclusion of such psychophysiological behaviors with other nonverbal behaviors is the requirement of instrumentation for their accurate detection. Dittman states (1972, p. 115):

The cases in which these (physiological) responses can be seen and thus considered to carry interpersonal messages of emotion in free social situations are probably rare for most individuals,...For the most part, investigators have bypassed the more visible and audible psychophysiological responses... and have concentrated on those which require laboratory instrumentation to be detected.

It is especially true that for educators and others in the helping professions the use of such instrumentation may cause more problems than is justified by the additional information gained.

In spite of these problems, several examples of the use of these physiological responses as nonverbal communication exist. For example, biofeedback instruments are popular aids in the assessment of relaxation and a trained clinician may observe respiration rates unobtrusively to assess hyperventilation of some anxious patients. Instrumentation for the assessment of physiological response was not used in this study.

Major Research Studies and Strategies in Nonverbal Communications

This section of the chapter reviews some of the major research efforts in the field of nonverbal communications. An attempt is made to point out and contrast some of the most common research methods in the field, especially as they help to provide groundwork for the study presented here.

For an overall view of research in nonverbal behavior and nonverbal communications, several excellent books have appeared within the past few years. Three of the most comprehensive books are: Nonverbal Behavior: The State of the Art by Harper, Weiss, and Matarazzo (1978); Bodily Communications by Argyle (1975); and Nonverbal Communications in Human Interaction by Knapp (1978). Several other books that are listed in the bibliography include excellent literature reviews but may be limited or biased by the particular research interest of the authors (Mehrabian, 1973; Ekman, Friesen and Ellsworth, 1972; Schefflen, 1972).

Early Research

Much of the original research into nonverbal communication of emotion was related to some question about the range of human emotion and attempts to identify the number of discrete or basic emotions that are commonly expressed (Ekman et al., 1970). In addition, some of the first research asked questions about the cross cultural recognition of facial expression of emotion. Researchers also have investigated (1) how much expression of emotion is a learned cultural phenomena, (2) how much emotional expression is innate, and (3) how emotional expression has evolved and what evolutionary advantage it affords. Many of the studies were decoding studies or ethological studies where photographs of facial expressions were used as stimulus material. Subjects were often asked to judge what emotion was being expressed.

As a result of early research, several ideas about the nature of facial expression of emotion have been fairly well established. These include the following:

1. Most facial expression involves a common set of facial expressions that are cross cultural and related to a finite number of basic emotions;
2. In most cultures there are display rules which may significantly modify or change the open display of certain emotions in specific social situations; and
3. Many of the facial expressions may be fairly reliably decoded even when presented only in a single photograph.

Facial Communication of Emotion

Most of the research relating expression of emotion and nonverbal communication has concentrated on facial expression. It is fairly well agreed that the face is the primary or most important source of non-verbal behaviors that connote emotional expression. Ekman and Friesen (Ekman et al., 1970, 1971, 1975) are two principal contributors to the literature for this area of investigation, although perhaps the first major work on nonverbal communications of emotion was Darwin's Expression of the Emotion in Man and Animals (Darwin, 1872/1965). Most of the work of Ekman and Friesen has been presented within their two books - Emotion in the Human Face: Guidelines for Research and Integration of Findings (1970) and Unmasking the Face (1975).

Ekman and Friesen (1972) performed research aimed at identifying the specific aspects of facial expression that make accurate identifications of emotions expressed through facial expression possible. This research has led to the development of the Facial Affect Scoring Technique (FAST) which allows identification of emotion in facial expression from specific information about the position and movement of the mouth, eyebrows, chin, eyes, and so on (Ekman, Friesen and Tompkins, 1971).

Mehrabian (1972) performed some related work on the expression of emotion through facial expression using factor analysis. The factor analysis research was also a decoding study, and again photographs of facial expression were used as stimulus materials. However, the subject was asked to rate each of the expressions in the photographs in terms of a wide range of descriptive terms. The resulting data on ratings for a wide range of pictures were used in a factor analysis of the expressed emotions.

Mehrabian identified three major factors in facial expression. These factors were labeled as positiveness, potency or status, and responsiveness (Mehrabian, 1972). Mehrabian saw these factors as important and basic to all nonverbal behavior and not just to the expression of emotion. Further, he theorized that they emerge as the important relevant dimensions because "...nonverbal behavior is a developmentally earlier and more primitive form of communication that man shares with animals" (Mehrabian, 1972, p. 14).

Mehrabian (1972) has also performed some encoding studies where data were taken on the basis of careful observation of a variety of nonverbal behavior cues. Although his observational instrument and techniques were well documented including data on interobserver agreement and other validity information, they were influenced by his conception of the three important factors in nonverbal behavior and include relatively few measurements on facial expression. His observational instrument tended to focus on more posture and position cues of the whole body rather than specific information about facial expressions.

Studies in Psychiatry and Counseling

Much of the interest in nonverbal communication has been prompted by questions of communications in the psychotherapeutic setting. The literature contains many references to indirect communications in the therapeutic setting. But only recently has much careful experimental

research been directed at the statistical validation of some of the questions that arise from this application of nonverbal communications.

An example of such research is in the work of Mahl (1956). He has conducted several studies aimed at the validation of some ideas about the occurrence of speech nonfluencies of psychiatric patients. This research is best described as a type of encoding study where the nonfluencies of speech are the dependent variables.

Another example of research related to the counseling setting comes from the work of Matarrazo and Wiens. In one study a confederate acting as the interviewer used head-nodding during the experimental portion of a standardized interview. The effect of increased head nodding was shown to be significantly longer durations of speech by the subjects used in the interview (Matarazzo and Wiens, 1972).

Interest in nonverbal behaviors by psychologists and therapists has traditionally been limited to the manipulation of the therapist-patient relationship and to the development of increased sensitivity to implicit communications in the course of the therapeutic dialogue. In more recent research there has been some interest in measurement of the decoding and encoding of nonverbal behaviors by psychiatric patients. Observation and analysis of patient nonverbal behavior has been thought to be a help in the diagnosis of certain mental disorders, and some have hypothesized that direct work on the improvement of nonverbal communications skills would be therapeutic in many cases (Harper, Matarazzo and Wiens, 1978).

Ethological Approaches

Much of the current research in nonverbal behavior was accomplished in controlled experimental settings. However, many of the earlier research studies were based on observation of behavior outside the laboratory. The work of Darwin was typical and the methodology was traditional in anthropological study (Darwin, 1872/1965).

In nonverbal communication literature the non-experimental methodology persists and has gained somewhat in popularity through the use of

film and videotape recording. Birdwhistle and his followers have traditionally believed that the key to understanding of nonverbal communications was to be found in careful study of short samples of behavior studied again and again (Argyle, 1975; Knapp, 1978; Birdwhistle, 1972). Of such ethological methods Argyle (1975, p. 15) stated:

This has been successful in the study of animal communications, and the statistical probability that one act will lead to another can be measured. The method has also been successful in studying the sequences of nonverbal communication on that accompany speech. A number of investigators have analyzed these processes in great detail, by frame-by-frame study of short periods of film.

Birdwhistle has concentrated on trying to establish a vocabulary of nonverbal signals and the rules through which they operate. To try to understand the relationship between the different events in a sequence of nonverbal behaviors, he has studied film over and over frame-by-frame (Birdwhistle, 1972). Scheflen, a student of Birdwhistle, has used similar methodology in the study of nonverbal behavior and it's role in the regulation of social interaction (Scheflen and Sheflen, 1972).

An example of current work in the study of nonverbal behaviors and communications in animals that has used an ethological approach is the work of Jane Lawick-Goodall (1968). She has studied nonverbal communications among several categories of primates including chimps and wild dogs (Argyle, 1975).

Frustration Theory

The term frustration and the theory related to frustration have evolved over the past 50 years. In his book, Frustration: The Development of a Scientific Concept, Lawson (1973) presented a history of the development of the concept of frustration as it has been used in psychology over the past fifty or more years. Lawson contended that the use of the term frustration in the early literature of psychology was often vague but close to the every day or vernacular definition.

Since then, the definition of frustration has been made more and more precise, although the definition may have varied somewhat in different schools of psychological thinking. Although the term frustration is still sometimes used in the vernacular within psychological literature, the term often refers to some very specific theories and definitions which are different from our everyday use of the term.

In psychological literature frustration almost uniformly refers to a stimulus situation. A rough definition of this stimulus situation is the blocking of the subject from the attainment of some expected or desirable goal or the blocking of the subject from the attainment of some goal that has been reinforced in the past. With this preliminary definition of frustration in mind a brief chronological review of some of the most significant and influential research and theory of frustration is presented to put the definition of frustration used in this study in perspective. Differences in the theories and research methodologies are discussed.

The Rosenzweig Picture-Frustration Test

The Rosenzweig picture frustration test was developed by Rosenzweig in the early 1930's. This test was a projective test in the tradition of the thematic apperception test or the Rorschach ink blot test (Rosenzweig, 1945; Lawson, 1973). To Rosenzweig, frustration was caused by three types of situations - privations, deprivations, or conflicts. In his test, pictures were presented to show frustration situations. The subject was asked to state how they would respond to each situation. Rosenzweig developed a category system for the classification of each response. The classification system was based in Freudian psychology and included categories of adequate-inadequate, direct-indirect, defensive-perseverative, and specific-nonspecific. Rosenzweig used the term frustration tolerance in the assessment of testing results and hypothesized that tolerance of frustration was basic to the definition of a well adjusted person (Rosenzweig, 1945).

The picture frustration test has enjoyed some continuing popularity; Rosenzweig has continued his own research and writing, comparing cultural differences and collecting data on special populations. In the early 1950's a children's version of the picture frustration test was developed by Rosenzweig which stimulated new research on the developmental aspects of children's reactions to frustration (Angelino, 1951; Rosenzweig, Fleming and Rosenzweig, 1954). Since many psychologists feel, as Rosenzweig did, that the ability to deal with frustration is an important indicator of emotional maturity and adjustment, research into development of frustration reactions and frustration tolerance has wide interest.

Naturally, there is some difference between a projective test of frustration response and direct observation of observational response. For this reason, the results of such a projective test, although valid for some purposes, may not provide accurate data on the actual response of many subjects.

The Study of Barker, Diempo, and Lewin

In the mid 1930's an important study was conducted and reported by Barker, Diempo and Lewin (1941) in which direct observation of the reactions of children to frustration was used. Their stated objective was to attempt to observe regression induced by frustration.

In their experiment preschool age children were allowed to play with a variety of toys in a playroom. The children were observed first to determine their developmental level. Some of the toys were identified as being more desirable, based on the children's reaction and frequency of use. In their main experiment, a glass barrier was placed in the play room which blocked access to part of the room and most of the desirable toys. The behavior of the children was then observed and recorded in written notes. Much of the behavior was classified into simple categories, and although little statistical analysis was done with the data, some regressive behavior patterns were noted in the reactions of the children. A large part of the study was basically

ethological in method, and results were primarily based on the observations and subjective conclusions of the researchers.

This study has been frequently cited and certainly has been influential in the development of some of the basic ideas about frustration. Although their research inspired some follow-up study, the original study is more frequently cited today (Lawson, 1973).

The Frustration-Aggression Hypothesis

The work of Rosenzweig and the study of Barker, Diempo, and Lewin led to several hypotheses concerning frustration. One important hypothesis was the frustration-aggression hypothesis. Dollard was a major contributor to this theory (Dollard, Douth, Miller, Mowrer and Sears, 1939; Miller, 1941). The idea that frustration sometimes led to negative behavior has been almost an universal assumption. The frustration-aggression hypothesis basically stated that although behavior caused by frustration is not always aggressive, aggressive behavior is caused by frustration. This notion that all aggressive behavior is caused by frustration is not easy to dispute if the notion of cause and effect or stimulus and response is allowed to include more than immediate temporal relationships. In the limit, frustration in childhood could cause aggressive behavior as an adult.

The frustration-aggression hypothesis has been popular in psychology and is still widely accepted today. Although current research acknowledges the role of hormones and instinctual behavior in looking for causes of aggression, most psychologists still look for frustration as the specific stimulus for episodes of aggressive behavior (Morris, 1977).

The Frustration-Fixation Hypothesis

The Rosenzweig picture frustration test includes in the classification of reactions to frustration the defensive-perseverative

dimension. If the defensive mode is aggressive, the perseverative mode suggests fixation.

In 1949 Maier published a book on his experimental work that had as its focus the presentation of the frustration-fixation hypothesis (Maier, 1949). Maier's operational definition of a frustrational situation was one in which no adaptive, goal-oriented behavior can develop in which the subject must continue to respond. His major experimental technique was a forced-choice situation in which a rat was forced to jump from a pedestal to one of two doors. One of the doors was locked; when the rat hit that door, it would fall four feet to a net below. The other door was open. Maier trained his rats to respond to a variety of clues as to which door was open. However, in his experiment he tricked the trained rats so that the clue was no longer valid; the previously correct response met with a locked door. Maier found that instead of learning the new situation the rats usually became fixated on the old correct response. In some cases they became so fixated they repeatedly jumped toward one door without regard to the clues. Although there has been criticism of both the experimental methods and the conclusions of this experiment, the notion that fixation is frequently related to frustration has been widely accepted.

Frustration and Behaviorism

Rosenzweig (1945), Barker et al. (1941) and Maier (1949) treated frustration as a unique phenomenon. Maier especially stressed the uniqueness of the frustration-fixation phenomena. However, many of the critics of Maier argued that frustration was not a unique phenomenon, and in the spirit of the growing fascination with behaviorism, frustration experiments were reviewed to try to explain the phenomena through several simpler mechanisms.

The work of Skinner was significant to the behaviorism movement in American psychology. In its purest form behaviorism stressed viewing all human behavior as learned. He treated the human subject in experiments as a black box and stressed the use of tightly controlled

experimentation in the laboratory where relationships between variables could be studied one at a time. As experimental results started to appear from this laboratory work, behavioristic psychologists often ignored more complex human behavior on the grounds that experimental work should explore the most simple behavior mechanisms first. They attempted to explain complex behavior in terms of a string of more simple stimulus-response (S-R) relationships.

In the research of the past 20 years, the study of frustration as a unique phenomena has been largely ignored because of the influence of behaviorism. When the term frustration has appeared in the recent psychology, it has been used primarily in one of two ways. First, where the term appeared in the less behavioristic literature such as social psychology or psychotherapy, it was frequently used in the more vernacular sense or in much the same way as more classical psychology used the term frustration. The current interest in the picture frustration test reflects this view. Second, in current literature the term frustration has taken on special meanings, different from the vernacular, especially within the literature of learning theory and other more behavioristic areas of psychology. The most common of these special usages are (1) the "frustration effect", and (2) frustration as a mediating variable.

The frustration effect is related to a phenomenon seen in a maze experiment with rats. In the experiment, the maze is a long runway that leads to a chamber at the midpoint where a food reward is given and finally to another chamber at the end of the runway where more reward food is given. After a rat has learned the maze and becomes acclimated to the reward situations, the frustration effect refers to the phenomenon that occurs when reward usually given at the central chamber is stopped. The rat runs, on the average, faster in the second leg of the runway toward the final chamber. The frustration experiment has been replicated a number of times with several variations and is still the focus of current research studies.

Lawson (1973) pointed out that behaviorism has had trouble using the simple S-R model to explain this and several other related

phenomena and that is why the frustration effect experiment has demanded so much attention. The related phenomena Lawson listed are the failure to explain why resistance to extinction of behavior is greater (1) when intermittent reinforcement is used, (2) when the response requires greater rather than less effort, and (3) when delayed rather than immediate reinforcement is used. "These facts are true in spite of the fact that experimental subjects can be shown to prefer continuous, immediate reinforcement of a response requiring little effort" (Lawson, 1973; p. 36).

In some of the behavioristic theory frustration is seen as an intervening variable. In models of complex behavior frustration can be hypothesized to be an internal state or response which causes certain response behavior. Brown and Farber (1951) have used such an approach in extending Hullian models of behavior. In their work, frustration was described as "...the consequence of either (1) the simultaneous activation of two competing excitatory tendencies, or (2) the presence of a single excitatory tendency and an opposing inhibitory tendency" (Brown and Farber, 1951, p. 153).

Common Frustrational Tasks

In research on frustration, different theoretical definitions of the term frustration have been used. In experimental research these definitions must be operationalized through the construction of a frustrating situation. The structure of these frustrating situations ultimately defines frustration in the empirical literature. Over the years several common experimental situations have been used to frustrate the organism or subject being studied. Lawson (1965) provided the following list of commonly used situations in frustration research:

1. Non-reinforcement after a history of reinforcement;
2. Prevention of completion of a reinforced response sequence;
3. Prevention of a response aroused by goal stimuli;
4. Delayed reinforcement;
5. A change in incentive conditions;

6. Failure; and
7. Hypothetical frustrating situations.

Pareek (1964) listed several specific tasks that have been used in psychological research to investigate frustration in humans. These tasks included the following:

1. Stopping a movie at the climatic point;
2. Blocking a fire exit during a fire alarm;
3. Use of a pegboard game where half of the subjects were provided with pegboards that could not be completed;
4. Asking a subject to recall a sequence of numbers where the digit span was too difficult; and
5. Falsifying test scores.

Student Nonverbal Behavior and Frustration

The two previous sections of this chapter present the background and theory in the two topics of nonverbal communications and frustration. This section reviews major studies of student nonverbal behavior and student frustration. The aim is to present briefly some of the studies and literature that are most directly related to the present study.

Betts and Frustration in Reading Education

In reading education the term frustration sometimes is used in the everyday or vernacular sense of the word but is sometimes used with a special meaning. One special usage of the term frustration is in the definition of the frustrational reading level.

Betts (1957) is usually given credit for the definition of the frustrational reading level. His popular remedial reading textbook presents a testing procedure for obtaining a quick and informal evaluation of reading ability called the Informal Reading Inventory (IRI). On the basis of the IRI test results the student is given a

grade level score for his independent reading level, his instructional reading level, his frustrational reading level, and his hearing capacity. The student is first presented with a short reading selection to read silently, then with a smaller sample of the passage to reread out loud, and finally with some comprehension questions over the material in the reading passage. This procedure is repeated with harder and harder reading selections until the oral reading mistakes and comprehension scores indicate the student has reached the frustrational reading level. The frustrational reading level is defined as that selection where correct oral word recognition is 90 percent or less and comprehension is 50 percent or less. The next two or three more difficult passages are read to the student by the examiner, and the comprehension questions are asked to see if the student understands the more difficult vocabulary and content when given orally. This determines the hearing capacity. A large difference between hearing capacity and frustrational reading level often indicates a high potential to gain quickly from remedial reading instruction.

The term frustration, as used in this reading level definition, is similar to psychological definitions of frustration. The student is blocked from the attainment of a desired goal. In the case of the reading frustration, the student is blocked by lack of decoding skills and vocabulary from reading and comprehending the selected passage.

Of special interest to the present study is Betts' listing of related behavioral characteristics of each level of reading. In his text Betts listed student behaviors that he associated with each level of reading difficulty such as independent, instructional, and frustrational. These characteristics are not formally used in the scoring of the IRI but do act as a further check on the students' ease of reading in each of the passages selected. The student behaviors that Betts identified are listed in Figure 4. The list of behavioral characteristics suggests that Betts recognized the significance of nonverbal behavior as a fairly reliable indicator of student frustration. Interestingly, the student nonverbal behaviors have not been emphasized

Independent Reading Level (Basal Reading Level)

Criteria: Word recognition - 99%; Comprehension - 90%

Behavioral characteristics:

Rhythmical and expressive oral reading

Accurate observation of punctuation

Acceptable reading posture

Silent reading more rapid than oral

No evidence of: Lip movement; finger pointing; head movement;
vocalization; sub-vocalization; anxiety about
performance

Instructional Reading Level

Criteria: Word recognition - 95%; Comprehension - 70%

Behavioral characteristics:

Same as independent reading level

Frustrational Reading Level

Criteria: Word recognition - 90% or less; Comprehension - 50% or less

Behavioral characteristics:

(May show one or more of the following behaviors)

Abnormally loud or soft voice

Increased tendency to stutter

Inaccurate observation of punctuation

Frequent requests for help

High pitched voice

Lack of expression in oral reading

Arrhythmic or word-by-word reading

Lip movement or sub-vocalization

Substitutions, repetitions, omissions, or word reversals

Tension as seen by frowning, blinking, excessive and erratic body
movements, "nervousness", and/or faulty breath control

Withdrawal as seen by non-interest in selection, unwillingness to
continue, refusal to continue, and/or attempts to distract the
examiner's attention.

Figure 4. Behavioral characteristics associated with reading at three levels (adapted from Betts, 1957; and Johnson and Kress, 1965).

as part of the criteria in the determination of the reading levels or for the recognition of frustration within the remedial reading setting.

The list of behaviors associated with the different reading abilities has been either overlooked or substantially reduced by many authors in reading education (Johnson and Kress, 1965; Guillard, 1974). This may be due to the fact that only the quantitative scoring of comprehension and word recognition are used in the actual assessment of reading abilities.

One interesting study related to the IRI compared the accuracy of the determination of frustration level by two alternative methods for scoring the accuracy of the oral-reading word recognition. The issue was whether word repetitions should or should not be counted as mistakes in the word recognition scoring. A trained polygraph operator was used to independently judge the point in the testing when the student experienced frustration. The frustrational level was determined by the two alternative scoring methods and by the polygraph operator independently. This study and several other studies have found significant agreement between the IRI and polygraph determined frustration (Rugel, 1971; Ekwall, 1974a, 1974b). Further, this type of study validated the notion that there is some significant change in the internal state and/or affect of the student as reading frustration occurs.

The Study of Jecker, Maccoby, and Breitrose

The major goal of the Jecker, Maccoby, and Breitrose (1965) study was to test whether or not inservice training in the recognition of nonverbal cues could improve the decoding ability of teachers. The specific decoding task was the judgement of student facial expression to determine level of student comprehension.

In their study, students' facial expressions were videotaped in a classroom setting while short lessons were presented to the students. The lessons were either easy or hard. At the end of each short lesson, comprehension questions were asked to check student

comprehension. Short segments of videotaped behavior of the students were then spliced together and shown to teachers who were asked to judge whether the student looked as though he/she was listening to the easy or hard lesson. The accuracy of the teachers' judgements was defined as the proportion of correct judgements.

Inservice training was then given to the experimental group of teachers. The training mainly involved the use of a checklist of nonverbal behaviors while viewing a number of similar practice videotapes. The accuracy of judgement of the teachers was assessed after the inservice training for both the experimental and control groups. Analysis of the data indicated that a small but statistically significant improvement in the accuracy of the teachers' judgement or decoding skill was attained as a result of inservice training. Another important outcome of this study was the data on the decoding accuracy of teachers. This data implied that the students' nonverbal behaviors did provide significant communications of student comprehension in the classroom. In the report of their study, the authors did not speculate on the link between student comprehension and nonverbal behavior. In the present study the assumption is made that part of the link between student comprehension and nonverbal behavior is frustration. Because of the use of very easy and very hard stimulus situations to elicit the student reactions, the judgement of comprehension may be equivalent to the judgement of student frustration. Further discussion of the similarities between the Jecker, Maccoby and Brietrose study and the present study is presented in Chapter V where some results are compared.

The nonverbal behavior checklist used in the inservice training of the teachers was far from being an objective observation instrument. However, it was constructed by trained psychologists, and because of the significant outcome of the study, this list provided important clues for the development of a classroom observation instrument in this study. The Jecker, Maccoby and Brietrose checklist is included as Appendix IV.

The Studies of Allen and Others on Decoding Ability

A series of studies conducted by Vernon Allen and his associates have adapted the basic design of the Jecker, Maccoby and Brietrose study to investigate the decoding abilities of various groups of subjects under a variety of conditions (Allen and Atkinson, 1970, 1977; Allen and Feldman, 1975; Allen and Brideau, 1977). For example, one study compared the ability of third grade students, fifth grade students, and adults in the decoding of the nonverbal behavior of third grade students. The third grade students used in the decoding task had been videotaped while viewing both easy and hard lessons (Allen and Feldman, 1975). Another study compared the skill of parents in decoding the nonverbal behavior of their own and other parents' children (Allen and Brideau, 1977).

Some of the work of this research group has been aimed at observing differences in encoding and decoding of the nonverbal behavior when the students are in the experimental situation versus when they are asked to role play or act as they would in the easy-hard lesson situation (Allen and Atkinson, 1977). This study found that the student actors tended to overplay emotion and were easier to judge correctly. This is an important issue since many of the early research studies used actors to portray emotions nonverbally, and the use of actors and role play is still common in nonverbal behavior research.

Eggen's Study

Eggen (1972) investigated nonverbal behaviors and student affect. The aim of his study was to systematically measure nonverbal behaviors elicited by students who expressed positive affect toward the science teacher and science class. The same observations were made on students who expressed negative affect toward their teacher and class. Students with mixed or neutral attitudes were not used in the study. The behavior of the positive and negative students was videotaped over a

period of time and the behavior was then analyzed using Eggen's observational instrument.

Statistical analysis of the data from the observations in Eggen's study showed a significant relationship between selected nonverbal behaviors and student affect. In this study the focus was on overall attitude towards the class or gross affect, rather than the more commonly studied transient emotional reactions or specific affect. The videotapes were made to reflect a sample of behavior taken over a period of several days.

Mehrabian has pointed out that the expression of more transient emotions is usually accomplished through facial expression, whereas the communications of attitude or gross affect is more closely associated with posture and position cues (Mehrabian, 1972). It is sometimes argued that the facial musculature is better suited to the transient expression of emotional reactions than are body postures. Some of Ekman's research has indicated "...that stationary facial expressions and posture are more likely to communicate gross affect (for example, liking), whereas facial and bodily movements are more likely to communicate specific emotions" (Mehrabian, 1972).

Eggen's study seems to validate these assumptions about the differences in nonverbal communications of gross and specific affect. These findings also indicate that the direct application of Eggen's observational instrument and behavior sampling techniques to the present study - involving more specific affect - is probably inappropriate.

Dittman and the Development of Nonverbal Behavior Skills

Dittman's (1972) research on nonverbal behavior abilities of children suggested that nonverbal communications skills develop during childhood. In one experiment, children were asked to give a series of clues to an adult experimental confederate in a question and answer session. The adults acted out nonverbal behaviors trying to

nonverbally communicate the need for more information, i.e., further clues. Testing with third grade and fifth grade students indicated there was a significant difference in the ability to correctly interpret or decode these nonverbal behaviors (Dittman, 1972).

Developmental change in the decoding skill of young children has also been found by Peterson, Danner and Flavell (1972). This research indicated the need for further investigation of the development of the encoding and decoding abilities of children. The study also suggested that studies of nonverbal communications in children may be limited in external validity because of the differences in the nonverbal skills of different age groups.

Summary of the Related Literature

This literature review found no study that involved the observation of encoding of nonverbal behaviors of students elicited by frustrational tasks. There are, however, several related studies that suggest important behaviors and research methods relevant to this study. Several of the most important points found in the review of the literature are summarized below.

1. Betts' definition of frustration used in reading corresponds well with most psychological theory of frustration. In Betts' work on the development of the IRI testing procedure, he identified a number of behavioral changes that were associated with student frustration in the reading situation. Betts indicated that these behavioral changes could reliably be used as part of the process of identification of student frustration in reading.
2. Task failure has previously been used to stimulate student frustration in the experimental setting.
3. Studies of judgement of student comprehension appear to involve judgement of affect that are similar or the same as judgement of student frustration. These studies indicate that student nonverbal behaviors should be reliable indicators of student comprehension or frustration.

III. DESIGN OF THE STUDY

This chapter presents the methods and design of this study. The study is divided into two main phases - the pilot phase and the main study. Because the design and methods of the main study depend on the results of the pilot study, the results of the pilot study are presented in this chapter. The chapter is divided into three major sections. The first section presents an overview of the study design. The second section discusses the pilot study. The final section presents the design of the main study.

Overview of the Study Design

The central purposes of the study are to construct an instrument to systematically observe student nonverbal behaviors resulting from frustrating classroom tasks and to use the instrument to examine the relationship between nonverbal behavior and frustration. The specific focus of the study is reflected in the following hypotheses:

- H1: A valid and reliable instrument can be constructed for the systematic observation of nonverbal behavior related to student affect states and affect changes induced by frustration.
- H2: There is a difference in the nonverbal behavior elicited by students performing tasks that are frustrating and those tasks that are non-frustrating.
- H3: There is a significant clustering of measured student nonverbal behavior elicited in frustrating tasks.

The study consisted of two parts or phases. The first phase or the study, the pilot phase, had as the major outcome the writing of an observational instrument for the collection of data on the nonverbal behaviors elicited while students attempt some frustrating and

non-frustrating tasks. This first phase of the study had as a secondary goal the planning of the second phase or main study where a specific frustrating task was investigated with the observational instrument. In the second or main phase, application of the observational instrument assisted in the validation of the instrument, identified areas where the instrument might be further modified, and provided information on the nonverbal behavior of students elicited by frustrational tasks.

The following plan, written in terms of goals and objectives, presents an overview of the study methodology and the temporal sequence of study tasks.

Phase I - The Pilot Study Plan

Goal: The major goal of the pilot study was the construction of an instrument for the observation of frustration induced behaviors.

Objectives:

1. Identify two or three tasks that lend themselves to the study of task frustration by the following criteria:
 - a. conditions of the administration of the task can be easily standardized and controlled;
 - b. tasks are similar to tasks used in the classroom; and
 - c. tasks can be controlled so as to be frustrating or non-frustrating.
2. Videotape students attempting frustrating and non-frustrating tasks.
3. Review literature on observation of nonverbal behaviors to identify other behaviors that have been found to be related to expression of specific affect or frustration.
4. Identify behaviors to be measured and quantified in the observational instrument.
5. Define each nonverbal behavior selected for study and identify observational techniques to be used to objectify and quantify each behavior.

6. Design rough draft and overall observational instrument, including record sheets, recording methods, and tabulation methods.
7. Apply the rough draft of the observational instrument to the pilot videotapes.
8. Apply statistical analysis to the data obtained with the draft observational instrument to help identify which of the nonverbal behaviors appear to be most important and which need redefinition.
9. Write the final observational instrument.

Phase II - The Main Study Plan

Goals: The major goals in this phase of the study were the validation of the observational instrument and the use of the instrument to investigate the relationship between nonverbal behavior and frustration. Secondary goals included the determination of which specific nonverbal behaviors were most related to student frustration and the determination of whether clustering of nonverbal behaviors occurred.

Objectives:

1. Administer frustrating and non-frustrating tasks to a sample of students. Videotape the students' behaviors for further analysis with the observational instrument.
2. Apply the observational instrument to the videotape records of student behavior.
3. Determine interobserver agreement through the application of the observational instrument by two independent observers.
4. Analyze the data from the application of the observational instrument statistically to test the major hypotheses. The analysis includes:
 - a. Comparison of frustrated and non-frustrated behavior data using Hotelling's T-squared test and Student's t-test;
 - b. Identification of clustering in the behavior data using cluster analyses; and

- c. Determination of how effectively nonverbal behavior data could be used to identify or judge frustrated from non-frustrated students through the use of discriminant analysis.

The Pilot Study

The major goal of the pilot phase of the study was the construction of an instrument for the observation of nonverbal behaviors that are related to frustration in students.

Definition of Frustration

As was discussed in Chapter II, there are a variety of ways that frustration has been defined in the literature of psychology and education. In this research, frustration is defined as a behavioral precondition or stimulus. Frustration involves a situation in which the student is blocked from the achievement of a goal. In this study, frustration is defined as a situation where a student is faced with repeated task failure. The specific circumstances of task failure used to provide frustration are further discussed later in this chapter.

Pilot Videotaping and Selection of Frustrational Tasks

Through a series of videotapes, an evaluation was made of possible tasks that could be used to induce frustration in a controlled way. Criteria for the selection of the tasks included the conditions that the tasks could be easily controlled and that the tasks were similar to some common classroom tasks.

The use of individualized tests for the task was appealing from the start. The main selection process involved investigation of a variety of individualized tests and testing procedures. One major advantage here was that the student responses, correct or incorrect answers, could be used to identify student failure and frustration.

Pilot videotapes were made both in the field and in the Oregon State University microteaching facility. Field videotapes were made of David Hawley, a psychologist with the Corvallis School District, administering individualized tests to a variety of children ranging in ages from eight to sixteen who were being tested as part of screening processes for special problems. The individual tests observed included:

1. The Peabody Picture Vocabulary Test;
2. Raven's Colored Progressive Matrices;
3. The Wepman Auditory Discrimination Test;
4. The Woodcock Johnson Achievement Tests;
5. The Parquetry Blocks;
6. The Peabody Individual Achievement Test;
7. The Bender Visual Motor Gestalt Test; and
8. The Key-Math Test.

The outcome of this investigation was the selection of the Peabody Picture Vocabulary Test (PPVT) and the mathematic subtest from the Peabody Individual Achievement Test (PIAT) for use in the rest of the pilot study and in the main study. The PIAT and PPVT tests were chosen for use in the main study because they are commonly used, standardized, and individualized tests that confront the student with a series of questions ranging in difficulty from easy to very difficult. The tests are designed so that the examiner need have little if any special training, and the recommended physical set-up for testing is quite compatible with the use of videotape recording.

In the second part of the pilot study, videotapes were made in the Oregon State University microteaching laboratory with the PIAT and PPVT tests using sons and daughters of several fellow graduate students. Much of this testing was done to work out details of the procedures to be used in the main study data collection.

In this study, frustration was defined in terms of student performance on the individualized tests. The test questions were arranged in order from easy to hard. Basically, during the first part of the testing pre-frustration behavior was assumed to exist. Towards the end

of the testing where the item difficulty was beyond student ability, frustrational behavior was assumed to exist. In these individual tests, instructions require the examiner to continue testing until a sequence of wrong answers is noted. Once this sequence was encountered the student was defined as being frustrated.

Construction of the Observational Instrument

Listing Important Nonverbal Behaviors

The first step in the writing of the observational instrument was to construct a list of behaviors to be included for observation. This list construction proceeded in three stages - an inductive observation of the initial pilot tapes, a review of the literature and a consideration of practical limitations.

The initial tapes were viewed many times to try to list and define as many nonverbal behaviors as possible. This process was undoubtedly influenced by the familiarity of the researcher with the literature, but an attempt was made to openmindedly observe any behaviors, including brow and forehead movements, eye movements, mouth movements, body shifts, paralanguage cues, and so on.

After an initial listing of behaviors was made from the tapes, a systematic review of the literature identified other behaviors to add to the list that had been shown to be related to frustration or were related to emotional reactions that might be induced by frustration. This process added few new behaviors to the list that had been constructed.

The initial list of behaviors to be considered for inclusion in the observational instrument contained a number of behaviors that had to be dropped from consideration for practical reasons. Common among these reasons were the limitations of camera angle and clarity of the videotape record. A complete listing of each behavior considered and reasons for elimination of the behavior seems unwarranted. However, four examples are listed below to explain some of the process.

First, several researchers, including Mehrabian (1972), have studied the communications of attitude through nonverbal behaviors and found that there are some interesting correlations between posture cues and attitude. Other researchers such as Ekman and Friesen (1975) have explored the communications of emotion through facial expression. However, with the use of videotape in a situation involving a fixed and unobtrusive camera location and angle, it was necessary to choose whether to zoom-in and obtain a more detailed record of the facial expressions or to pan-out and observe posture cues including leg and foot movements with a loss of detail in the record of facial expression.

As has been discussed in Chapter II, there is fairly good agreement among researchers that the more transient expressions of specific affect, such as frustration, are more related to facial expression, whereas gross affect and attitude are related to posture and position cues. For this study the decision was therefore made to focus more on facial expression, and thus most posture cues, leg positions, and foot movements were delimited from consideration in the construction of the observational instrument.

A second example of practical limitations to the nonverbal observations is the use of eyeblinking. Some research has indicated some relation of eyeblink activity to anxiety (Harper, Wiens, and Matarazzo, 1978). However, where such research has been performed, eyeblinking is often recorded by monitoring electrical activity of eye muscle to detect movement. Towards the end of the pilot study it became obvious that there were too many conditions that could cause the counting of eyeblinks from videotapes to be unreliable and the eyeblink record was eliminated. Some of these problems included poor lighting or focus, subjects who wear glasses, and subjects where hair was hanging at least partially over the eyes.

Third, studies of paralanguage, intonation, phrasing, and speech rate seemed unpractical for they either require judgement or extensive and expensive analysis of the audio record. Further, in the testing situation oral response was limited and there was some desire to avoid

using paralanguage cues in detection of frustration since oral behavior is often limited in the classroom setting.

Fourth, some researchers consider physiological response as part of nonverbal behavior. Instrumentation using wires and probes, however, could easily affect student behavior. Further, research on such behavior would have limited its usefulness since the physiological information is not easily accessible in the classroom setting.

Form of the Observational Record

At this point some consideration had to be made as to the form the observational instrument would take. Medley and Metzger (1963) review several methods commonly used to systematically record classroom behavior. Two types of observational instruments that they discuss are category systems and sign systems (Medley and Metzger, (1963)). However, many of their arguments for and against each system are somewhat irrelevant to a study such as the present one where a videotape record of the behavior allows for repeated viewing of the same sequence of events over and over to record complex data on the behaviors observed. Many of their considerations have to do with observation records made in the real time setting of classroom observation without benefit of video records.

The observational instrument constructed in this study is a sign system type of instrument. A sign system is one in which a list is made of "...incidents of behavior which may or may not occur during a period of observation. The record will show which of these incidents occurred during a period of observation and in some cases, how frequently each occurred" (Medley and Metzger, 1963, p. 299). In the instrument a one minute sequence or sample of behavior was divided into five-second intervals. For most of the behaviors, if the behavior occurred during a five-second interval, a "1" was recorded in the cell representing that behavior. The "score" for a behavior was the number of five-second intervals in the sampled minute in which a "1" was recorded. This five-second interval is quite small compared to many

other observational instruments reviewed and would not be possible without videotape recording of the student behavior. Also the one minute sample of behavior is comparatively short.

There were several important considerations in the choice of these short time periods. First, the use of videotape allows for rather intensive analysis of small sequences of behavior, and some researchers, such as Birdwhistle (1970) have claimed that a great deal can be gained from this intensive observation. Second, Ekman and Fiesen (1975) and others have studied the speed of important facial nonverbal behaviors and have studied the speed of important facial nonverbal behaviors and have identified what they term micro expressions and macro expressions. These two types of expressions have display times of two to four seconds and four to five or more seconds respectively. Micro-momentary facial expressions were identified by Haggard and Isaacs (1966) to last from one-eighth to one-fifth of a second. Therefore, it was decided that five-second intervals were sufficiently small to record such expressions in sign system fashion without significant loss of information. Third, because one of the goals of the research was to help teachers or researchers identify student frustration, it seemed reasonable to demand that this identification would require only limited observation. Hopefully, a teacher could recognize frustration in one minute or less of observation; therefore important nonverbal behaviors should be identified in less than one minute of intense observation. Finally, it was assumed that the type of emotional reactions that would elicit these nonverbal behaviors may last only a short amount of time.

Quantification and Definition of the Nonverbal Behaviors

The next step in the pilot study was to define each of the nonverbal behaviors so that quantification of the videotape record could be as objective as possible. This process of defining the behaviors required several cycles where the behaviors were defined, then the definition was used in conjunction with careful observation of pilot

tapes. Each application of the instrument suggested revisions and clarifications to improve the objectivity and precision of the instrument. The definitions of each of the behaviors selected are given in Appendix VII.

Assembly of Rough Draft

The list of behaviors and their definitions were assembled into a rough draft of the observational instrument. At this time a record sheet was also designed for the scoring of the observational data. In this record sheet a grid was made to score and tabulate observations. Each row of the grid was used for the recording of each behavior for each of the 24 five second intervals - 12 for pre-frustration behaviors and 12 for frustrational behaviors. In each case the nonverbal behaviors were recorded as a one or zero (1 or 0) in each box depending on whether the behavior was seen to occur in the five second interval. The final form of the instrument and record sheet are presented as Appendix VII.

As a last check in preparation for the main study, the observational instrument was applied to the last 10 pilot video tapes. These tapes were made in the Oregon State University micro-teaching center and were arranged to simulate and prepare for the main study. The data from this use of the instrument was analyzed in several ways. A significant difference in overall nonverbal behaviors could be seen between the pre-frustration and frustrational behavior. Several clustering schemes were tried to identify modes of frustrational reaction but with only 10 data points no significant clustering could be claimed.

Additional Comments on the Instrument

Although each of the categories used in the observational instrument is defined in Appendix VII, some further comments are appropriate

about the instrument and some of the categories to help explain the genesis of the design.

Sequence. The purpose of the sequence track is basically to aid in the location of the correct place in the video record from which to score the behaviors. A "clock" was designed and constructed that could be used in the field to mix a time track with the audio record as the field recordings were made. The time track consisted of "clicks" mark each second and a short "beep" to mark the five-second intervals. The clock also counted the five-second intervals and displayed the interval count on a LED digit display that could be viewed on the video portion of the videotape recording. A block diagram and description of the clock is included in Appendix VI. The placement of the clock in the experimental set-up is shown in Appendix X.

Movement. Comments about movement or restlessness are common in descriptions of student behaviors and are sometimes associated with stress or frustration. It was therefore thought to be desirable to make some attempt to measure the amount of movement taking place during the observation even though the camera setting was made to optimize recording of facial behaviors.

Although activity level, including hyperactivity, has been an area of interest to researchers for several years, no agreement has been reached on an operational definition of hyperactivity nor has there been much agreement on methodology of measurement of body motion (Werry, 1968; Bryan and Bryan, 1975). In the studies that have been done elaborate electrical instrumentation including special chairs, electrodes, and accelerometers are dominate; such methods seemed undesirable for this study.

In this study, the main means of recording body motion was to try to record the coordinate (X,Y) position of the nose of the subject under observation as it was viewed in the videotape monitor.

Coordinates were recorded at the start of each five second interval during the one minute sequence. Although coordinate positions on the TV screen actually correspond to angular displacements and are a function of monitor screen size and camera lens setting, the displacements can be converted into actual approximate displacements for each particular experimental set-up. Because of the camera angle in the experimental set-up, the coordinate positions recorded reflect a mixture of both front to back and side to side motion of the subject. In the main study set-up the position coordinates were recorded to the nearest whole coordinate and the distance between coordinates corresponded to about 2.5 centimeters or approximately one inch.

The coordinate positions were not used as data in the comparison of pre-frustration and frustrational behavior by themselves. The coordinate position data were used to calculate a score for each of two categories which reflect aspects of motion called "change" and "magnitude". "Change" notes the number of intervals in which the coordinate position shows a change. "Magnitude" is basically a calculation of the average distance moved from coordinate to coordinate during the observation sequence. Both change and magnitude are explained more fully in Appendix VII.

The Main Study

In the main study the major goals were the validation of the observational instrument constructed in the pilot study and the examination of the relationship between nonverbal behavior and frustration. This section of the chapter describes the methods and procedures for the examination of the relationship between nonverbal behavior and frustration for both the gathering of data and the data analysis. Included is some discussion of how the data analysis provides a basis for checking the validity of the observational instrument. The discussion provides enough detail about the method so that the study could be replicated.

The topics included in this section of the chapter include (1) selection of the subjects, (2) gathering data, (3) measurement of validity and reliability, (4) overall validity of the observational

instrument, (5) description of the data, and (6) description of the data analysis.

Selection of Subjects

The Lebanon School District granted permission to conduct the main study using students in its schools. School district restrictions required that parental permission for testing be obtained and that the testing be voluntary. The consent form letter is shown in Appendix V. Although the requirement that parental permission be given for the participation in the study limits the random nature of the sample population, it was felt that this could not be avoided in most studies conducted in the public schools today.

Unfortunately, the fact that this self-selection introduced bias in the sample was quite obvious as the parental permission forms returned. In the Lebanon Middle School, the mathematics classes for the eighth grade students are grouped or tracked into high, middle, and low ability classifications. Response was much greater from students in the advanced classes and from girls so that more girls than boys and more advanced than average or low ability math students were included. Because only 40 positive permission slips were returned out of a potential of slightly over 100, an attempt was made to use all 40 students rather than try to even the distribution of the sample by some random selection process since that would significantly reduce the sample size. Thirty-eight of the forty students were used in the main study. One of the 40 students was unavailable during the testing due to a class schedule change. One other student was tested but not used in the main study because she was a recently immigrated Cambodian and had significant difficulty with the English language. The sample distribution with respect to sex and math ability grouping is given in Table I.

Table 1. Distribution of the sample population by sex and mathematics class grouping.

Sex	Math Class Grouping			Total
	Low	Average	High	
Male	3	2	9	14
Female	<u>9</u>	<u>5</u>	<u>10</u>	<u>24</u>
Total	12	7	19	38

Gathering Data

The students participating in the study were tested during the months of March and April, 1979. Students were taken one at a time in random order from their mathematics class to the conference room of the middle school counseling center. Testing of each student required from 20 to 30 minutes and therefore only one or two students could be taken from any particular mathematics class during the same day.

The two tests, the PIAT and the PPVT, were used and the order of testing was mixed so that half of the students took the PIAT first followed by the PPVT and half were tested in the reverse order. Further, two alternate but equivalent forms of the PPVT were used; students were instructed not to discuss the testing with each other until the testing was complete.

The two tests were administered largely in accordance with the instructions provided by the test publisher. These instructions include both the physical set-up of the testing and the actual testing procedure. The physical set-up as shown in Appendix X is essentially as suggested by the instructors manual.

The publisher's instructions on the testing procedure specify some informality that is contrary to what might be suggested for repeatable and standardized testing procedures. For example, the instructions make suggestions concerning the use of a little "small-talk" to relax the subject and make him or her more receptive to the testing. In addition the use of some verbal praise and positive feedback is suggested especially at the start of a sequence of testing.

Since the publisher's reliability and validity data on the two tests must reflect this degree of informality these testing guidelines were used as much as possible.

Both the PPVT and PIAT tests are wide range informal tests using a so-called power test format: that is, the questions are arranged in order from easy to hard. The starting point in the test for each student is determined by age or grade information. However, for the PPVT, if the student doesn't achieve eight consecutive correct responses, the "basal level", then the examiner is instructed to ask additional questions working backwards until the basal requirement is fulfilled. The basal level requirement for the PIAT is five consecutive responses.

After the basal level is established, the test proceeds in order towards the harder questions. The testing continues until a "ceiling level" is established. The ceiling level is defined by a sequence of incorrect responses. For the PPVT the ceiling level is six incorrect in a sequence of eight consecutive questions, and for the PIAT the ceiling is five incorrect in seven consecutive questions. Note that both tests consist of four part multiple choice questions and the ceiling level must consider correct guessing.

During usual use of these tests a record is made by the examiner on a score sheet for each test. Although all of the testing was recorded on video tape the written record was also made both to help record the test performance and to help simulate the usual testing conditions, especially since it was thought that the record keeping must affect the examiner's own nonverbal behavior such as posture, eye contact, and so forth.

Use of the Observational Instrument

After the videotape data were collected the observational instrument was used to record the nonverbal behaviors. First it was necessary to identify a good sample of pre-frustrational behavior and a good sample of frustrational behavior. To identify the pre-frustrational

behavior each of the test record score sheets were examined. The basic idea here was to identify a one minute sequence of questions where the student was clearly confronted with easy questions. In operational terms this meant that a sequence was identified where the student answered all questions correctly. If this was not possible, a sequence was selected where only one or at the most two answers were incorrect. This one minute sequence usually corresponded to eight to ten questions. In most cases the identification of this pre-frustrational behavior sequence was not difficult.

A similar method was used to identify the frustrational behavior sequence, although because of the possible correct guessing it was usually not possible to identify a sequence where all answers were incorrect. Where possible, the frustrational behavior sequence was chosen to include much of the so-called "ceiling level" sequence. The necessary one minute sequence of behavior usually included fewer questions because the pace of the exam was usually much slower. In the PIAT mathematics exam this reflected longer questions but in both the PIAT and PPVT tests the students' time to answer did increase.

The use of the test instrument to record behaviors has already been described as part of the description of the pilot study; complete instructions for the use of the instrument are included in Appendix VII.

Validity and Reliability

The reliability of any evaluation instrument is a necessary but not a sufficient condition for the validity of the instrument. In the case of an observational instrument such as the one developed for this study, reliability is measured by interobserver agreement. That is, reliability, which is basically a measurement of the consistency of the instrument, depends on the amount of agreement that can be attained by two or more observers while using the same instrument to independently view the same behavioral record.

As has been discussed the greatest possible objectivity was a major goal in the construction of the instrument. Observer ratings and judgements were avoided where possible and the use of a videotape recording system and precise definition of the observed behaviors were also planned to increase interobserver agreement.

To measure the actual amount of interobserver agreement the statistics called Scott's Pi was calculated. Scott's Pi was developed for the measurement of interobserver agreement and has been used for over ten years in such research (Scott, 1968; Medley and Metzger, 1963; Light, 1973; Frick and Summell, 1978). Scott's Pi is basically an estimate of the amount of agreement observed beyond that agreement due to random chance. P_o is the agreement observed; and P_e is the agreement due to chance for each case. Scott's Pi is then given by:

$$\pi = \frac{P_o - P_e}{1.0 - P_e} ,$$

To gather data from which Scott's Pi could be calculated, a fellow graduate student was enlisted as the second observer. This student had only general knowledge of the aims of the research and at that point was not aware of the observational instrument. First two tapes were selected at random from the main study videotapes for use in observer training. The training consisted of about one hour of informal discussion of each of the categories and instruction on the use of the videotape recorder. The graduate student used one more hour for practice in the use of the videotape recorder and the observational instrument.

In the second part of this interobserver agreement measurement five videotapes were selected at random and scored independently by the two observers. Scott's Pi was calculated for each of the behavior categories in the observational instrument. Although the use of only five videotapes seemed a minimal number for the assessment of interobserver agreement, it was necessary to minimize the burden on the second observer. The time required for each observer to view the five

student videotapes (10 test records) was approximately seven to ten hours.

The results of the interobserver agreement measurements are included as Appendix I and were overall very acceptable, being similar to interobserver agreement figures found in similar studies of non-verbal behavior (Eggen, 1972; Mehiabian, 1972). Further discussion of the results is presented in the next chapter.

After the interobserver agreement study was complete the two observers discussed the sources of the disagreements that were found in the two independent records of the student behavior. Some comments and conclusions about the interobserver agreement measurements are included in Appendix II.

Validity of the Observational Instrument -- Some Arguments

The proof of validity of any evaluation instrument is not a simple process. Validity arguments vary considerably depending on the evaluation instrument under consideration. Some discussion of basic and relevant validity arguments is presented below to show how the proof of validity of the observational instrument is attempted.

Validity is sometimes defined as the degree to which an instrument measures what it claims to measure. Validity is not assigned specifically to an evaluation instrument but is sometimes argued for an evaluation process. That is, the validity of a particular evaluative procedure depends on the sample population characteristics, the purpose of the evaluation, and other factors. One procedure may be valid for one purpose with a particular population but not for another. Often reference is made to the validity of a particular test instrument or evaluation procedure. In these cases validity seems to assume that the instrument is being used for the purpose for which it was intended and is being given to a "normal" population.

The first step in the determination of the validity of the observational instrument developed in this study is to identify clearly

the purpose of the instrument. The three purposes of the observational instrument were considered. First, the instrument can be seen as an instrument to objectively observe certain student nonverbal behaviors, especially facial nonverbal behaviors seen in the individual testing setting. Second, the instrument can be viewed as an instrument to detect emotional reaction or specific affect that is stimulated in the testing situation. Finally, the instrument can be seen as an instrument to detect student frustration.

The proof of validity of the instrument seems the most straightforward in the first case where the instrument simply measures student nonverbal behavior. It is necessary to define nonverbal behavior and be sure that the behaviors measured concur with the definition. Then, measurement of interobserver agreement is important. Finally arguments from the data in the pilot and main studies can be used to show that most of the interesting or important nonverbal behaviors are used in the instrument.

Proof of the instrument as a valid measure of emotional reaction or as a tool to detect student frustration is much more difficult. The outcome of the use of the instrument is a series of scores related to the nonverbal behaviors observed. Therefore, the instrument certainly could not be claimed as a tool to detect frustration or emotional reaction by itself. That is, application of this instrument to a particular student does not give information on the emotionality or frustration of the student observed. This instrument might later provide the basics for such an instrument, but the instrument was not constructed with that goal in mind.

However, to determine if the instrument is sensitive enough to detect nonverbal behavioral changes that correlate with emotional reaction or frustration is a different argument. There might be some need to validate the use of this particular observational instrument, as a research tool to study relationships between nonverbal behavior and emotional reaction or frustration. Here arguments commonly called construct validity can be used.

Construct validity can be defined as the degree to which the instrument fits or meets the requirements of the theory from which the psychological construct comes. Gronlund (1976) describes several types of construct validity tests and arguments, but basically they all involve the use of psychological or educational theory to provide a hypothesis about the characteristic being evaluated. To be valid, the evaluation instrument must provide data that fit the hypothesis. For example, if theory states that a certain population should be more creative than another, then a test of creativity can be in part validated by providing data that confirms this hypothesis. Naturally, such a method of validation tests or validates the theory as much as the evaluation instrument.

In this study, the testing of hypotheses two and three can be considered directly related to the construct validity of the observational instrument. There is strong evidence suggesting that nonverbal behavior of students in the individual testing situation should change from the pre-frustrational testing to the frustrational testing. Using the data from the observational instrument to measure nonverbal behavior in this individual testing situation and then testing the above hypothesis not only validates the theory but validates the observational instrument. Similarly, the testing of hypothesis three can help in the validation of the instrument as well as provide information about the reactions to frustration that occur in the individual testing situation.

Description of the Data

Application of the instrument results in a series of scores, most of which represent the number of occurrences of a particular behavior during the one minute sequence (or sample) of behavior. The data from the main study thus yielded two sets of scores - pre-frustration and frustrational behavior - for each student for each of the two tests. Each set of scores was considered as a data point in a multivariate space where each dimension corresponds to one of the behavior scores.

Further, the pairs of data points corresponding to the pre-frustrational and frustrational behavior observation of each student were considered paired data since they corresponded to the behaviors and behavioral changes noted for that student.

During the application of the observational instrument, extensive use was made of the "other comments and observations" space on the behavioral record in an attempt to record every behavior that was seen. However, after the tabulation of the behavioral records, some of the behavioral categories were eliminated from further consideration because they were so infrequently observed. The criterion that was used in the elimination of such behaviors was that if the behavior was not observed at least one time for more than 25 percent of the students that were studied, it was eliminated from the data analysis. Actually, this only eliminated six behaviors from further analysis (nose movement, lip bite, question repeat, question answers, sigh, and don't know). By the same criterion, none of the behaviors noted in the "other comments and observations" were of sufficient frequency to be included in the data analysis.

This is not meant to imply that the infrequently observed behaviors were not important. Some of these behaviors seemed quite communicative, but being less frequent they are less amenable to statistical analysis with a limited sample of behavior and small population. A few behaviors seemed ideosyncratic and limited to one individual; these were largely ignored. Descriptions of these more infrequent nonverbal behaviors are included in Appendix VII, and further discussion is included in Chapter V.

The 17 behaviors or behavior scores that were taken from the application of the observational instrument and used in statistical analysis are listed below. The three and four letter abbreviations are those that were used in much of the computer analysis of the data. More complete definitions are available in Appendix VII.

- 1 TTA - Time to Answer - the average time from the asking of the question to the student response.

- 2 AMTM - Movement Frequency - reflects the number of position changes seen between successive position observations, at the beginning of each five-second interval.
- 3 DSTM - Average Distance Moved - the average distance of movement seen between successive position observations. Position observations are made at the start of each five-second interval.
- 4 PSFT - Posture Shift - the number of shifts of seating in the chair.
- 5 LEAN - Lean Back - number of fairly rapid leans back and away from the test.
- 6 HSHK - Negative Head Shake - number of occurrences of negative shakes of the head (side to side).
- 7 MANP - Self Manipulation - number of active uses of the hands to manipulate any part of the face or clothing.
- 8 HTH - Hands to Head - number of touches of one or both hands to the head.
- 9 GAZE - Gaze - number of looks up towards the examiner.
- 10 EBFL - Eye Brow Flash - number of rapid raises of the eye brows.
- 11 EBLO - Eye Brow Lower - number of lowerings of the eye brows.
- 12 MTWT - Mouth Twitch - number of movements of the sides or corners of the mouth, transient or sustained.
- 13 SMIL - Smile - number of smiles.
- 14 LICK - Lip Lick - number of times the tongue is seen licking the lips.
- 15 MDRP - Mouth Drop - number of times the mouth drops open.
- 16 FLPA - Filled Pause - number of occurrences of non-intelligible oral filler. For example, "aa-h" or "um --".
- 17 SUBV - Subvocalization - number of occurrences of lip movements that are associated with subvocalization, with or without slight vocalization.

Data Analysis

The discussion of the data analysis that follows is focused on the testing of the three hypotheses to be tested in this study. Included are descriptions of statistical tests and arguments that provide the basis for the acceptance or rejection of each hypothesis. Although, the three hypotheses are interrelated, the discussion of each hypothesis is presented as separately as possible.

Hypothesis One

The first hypothesis, that a valid and reliable instrument can be constructed for the systematic observation of important aspects of nonverbal behavior related to student affect states and affect changes induced by frustration, cannot be directly tested by statistical means. The test of the first hypothesis rests mainly on the tests of the validity and reliability of the observational instrument for the purpose stated. Since the validity of an instrument is a matter of degree, it may be argued that this is not simply an accept or reject hypothesis. Certainly the evidence and arguments for or against validity are diverse and not usually quantitative.

An appeal to construct validity arguments provides most of the basis for the validity of the observational instrument. First, as already stated "construct" does not deal with the construction of the instrument but refers to the psychological construct or constructs related to the theory from which the evaluation instrument is designed.

One type of construct validity involves the actual construction of the instrument in accordance with the theory. For the observational instrument designed in this study, three key constructs are nonverbal behavior and nonverbal communications theory, theory of emotional response, and theory of frustration. All of these have been reviewed in relation to this study in Chapter II. The instrument has been constructed to include most of what the current literature identifies as important nonverbal behaviors. These behaviors were chosen on the

basis of what previous theory and research has identified as most related to specific affect response or on the observation of students' nonverbal behavior in the classroom setting. Careful consideration of the literature and theories of nonverbal behavior, emotional response, and frustration in the construction of the observational instrument help validate the instrument.

Further validation of the observational instrument is also provided by other construct validity arguments. As has been already discussed, validity can be further established for the instrument if it provides data to support hypotheses related to the theory of the psychological constructs related to the instrument. In this way, the testing of hypotheses two and three can help to validate the observational instrument.

As additional evidence to test hypothesis one, the reliability of the observational instrument was tested through measurement of interobserver agreement. Although reliability is a necessary but not a sufficient condition for the validity of an evaluation instrument, it is one place where some quantitative testing may be applied. The measurement of interobserver agreement has already been discussed in this chapter and the results of that measurement are presented in the next chapter.

Hypothesis Two

Hypothesis two states that there is a difference in the nonverbal behaviors elicited by students performing tasks that are frustrating and tasks that are non-frustrating. Statistical testing of hypothesis two was performed for each of the two testing situations, the PPVT and the PIAT. In each case a series of scores on the measured nonverbal behaviors were taken for each subject at the start of the test and again during the frustrating part of the test. Each series of scores provides one multivariate data point; the application of the instrument to both the pre-frustration and frustrational testing yields a pair of multivariate data points. Thus for the student sample a number

of paired data points were obtained for each of the two testing situations, the PIAT and the PPVT.

Testing univariate paired data is often accomplished through the use of Students t-test. For the case of multivariate data, Hotelling's T-squared test is similarly used, which is a multivariate extension of Students t-test (Morrison, 1976; Harris, 1975). Hotelling's T-squared test was therefore applied to the data from each of the two testing situations, the PIAT and the PPVT, to test the overall hypothesis of whether any change in nonverbal behavior occurred. Where Hotelling's T-squared test indicated a significant difference in the multivariate data, the next step in analysis is often to apply univariate tests such as Students t-test to try to identify which variables contribute the most to the observed difference. Such testing was performed for each behavior category for each test to try to identify which of the non-verbal behaviors contribute most to the observed overall change.

Although the T-squared test is a multivariate test that makes some assumptions about the normality of the distributions, the use of this statistic is justified in two ways. First, the T-squared test, like the individual Student's t statistic has been shown to be quite robust; especially where the sample size is large ($N = 30$ or more) the variable distribution does not significantly affect the use of the statistic. Second, the data is multivariate in nature and the field of non-parametric multivariate analysis is just now emerging so that few statistical tools are available.

Hypothesis Three

Hypothesis three, that there is a significant clustering of measured student nonverbal behavior elicited in frustrating tasks, was investigated using statistical clustering techniques (Anderberg, 1973; Harris, 1975). Clustering is a matter of degree, however, and clustering statistics are essentially descriptive statistics. Therefore the acceptance or rejection of the null hypothesis is a matter of judgement.

Clustering is used to separate groups of data points that are "close" together and therefore represent similar behavioral responses. To explain the idea of clustering consider the following example. Suppose in a remedial reading classroom data were taken on students' IQ, reading ability, and frequency of disruptive behavior. One hypothesis might be that two distinct types of students were commonly found - low IQ students who have low reading ability and are seldom disruptive and high IQ students who also have low reading ability but who are commonly disruptive. Hypothetical data for this example, showing clustering into two groups, are shown in Figures 5 and 6.

Statistical clustering analysis of the hypothetical data should be able to identify the two clusters. Although a number of clustering algorithms have been developed for a variety of individual applications and preferences, in this study a clustering technique called "block clustering" has been used (Hartigan, 1975). Some further description of the clustering algorithm used is contained in Appendix IX.

As can be seen in the hypothetical clustering sample, the clustering of the data points usually involves only some of the variables and for this and other reasons can be seen best when the multivariate data is projected in certain directions or along certain axes called canonical coordinates. For the clustering algorithm used in this study, standard discriminant analysis can be used to find the best two dimensional projection of the data in which the clustering can be seen (Morrison, 1976; Harris, 1975). Such discriminant analysis was used to help visualize the amount of clustering that exists in the data.

Summary

The goal of the study is the systematic observation of student nonverbal behavior to investigate the relationship between nonverbal behavior and student frustration. This investigation was divided into two phases, the pilot phase and the main study. The goal of the pilot phase was the construction of an observational instrument for the systematic observation of nonverbal behaviors that are related to frustration. The second phase, or main study, involved the application

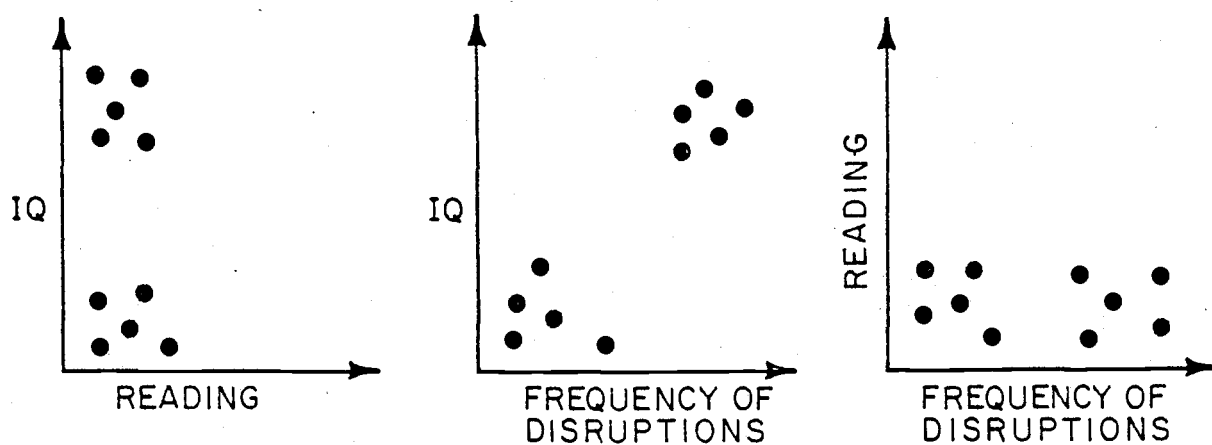


Figure 5. Example of clustering into two groups displayed by scatter diagrams of three variables taken two at a time.

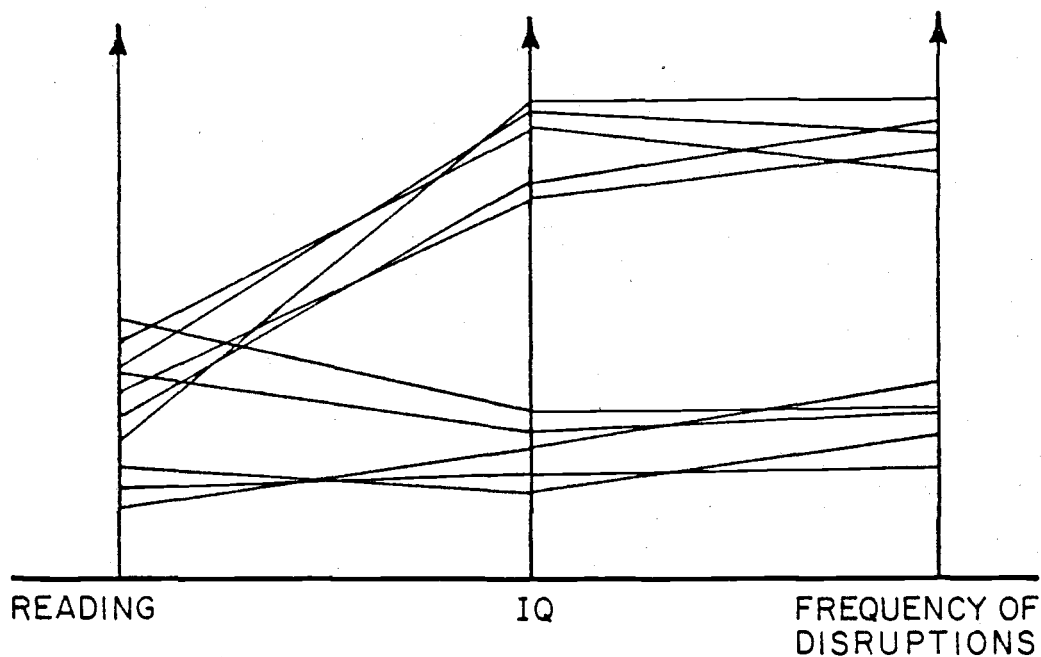


Figure 6. Example of clustering into two groups displayed by a profile analysis of three variables.

of the observational instrument to a sample of students. The main study was aimed both at the validation of the observational instrument and at the observation and analysis of the relationship between frustration and nonverbal behavior.

The design of the pilot study began with pilot videotape recording of students that were attempting both frustrating and non-frustrating tasks. The analysis of the videotapes and a literature review were used to list and define nonverbal behaviors to be included in the observational instrument. After the list of behaviors was defined, the form of the observational instrument, the record sheet, and the tabulation techniques were drafted. During this stage, the decision was made to use two individualized tests in the main study, the PIAT and the mathematics subtest of the PPVT. Several final videotapes were made to help in preparation for the main study.

In the main study, a sample of students was videotaped during individual testing with both the PIAT and PPVT tests. Selection of the frustrational and pre-frustrational behavior samples to be systematically observed was made on the basis of the students' responses. During the first part of each test, the questions were selected to be easy and were considered to elicit pre-frustrational behavior. During the last portion of each test there was a sequence of difficult questions that were assumed to elicit the students' frustrational behavior. The observational instrument was designed to record nonverbal behaviors over a one minute sample of student behavior and the application of the instrument yields a series of scores that generally represent the frequency of occurrence of each of the behaviors.

The data were analyzed to test the study's three hypotheses:

- H1: A valid and reliable instrument can be constructed for the systematic observation of nonverbal behaviors and behavior changes in response to frustration.
- H2: There is a difference in the nonverbal behavior elicited by students performing tasks that are frustrating and those tasks that are non-frustrating.

H3: There is a significant clustering of measured student nonverbal behavior elicited in frustrating tasks.

The testing of hypothesis one is based on the process used in the construction of the instrument, the interobserver agreement, and to some extent on the outcome of hypotheses two and three. Hypothesis two is tested by comparing the pre-frustrational and frustrational behavior using Hotelling's T test and individual t-testing. Hypothesis three is more subjective but is based on the cluster analysis of the data from the frustrational behavior observed. A variety of techniques are used to examine the results of the clustering analysis but the eventual acceptance or rejection of the hypothesis is based on judgement.

IV. RESULTS OF THE STUDY

This chapter presents the findings of the study. The first section focuses on the major findings related to the three major hypotheses. The second section presents some additional findings that were not directly related to the hypotheses.

Findings Related to the Hypotheses

Hypothesis One

The first hypothesis states that a valid and reliable instrument can be constructed for the systematic observation of student nonverbal behavior elicited by frustration. This hypothesis was not amenable to simple statistical testing. The foundations for the arguments used in the testing of this hypothesis have been laid in the previous two chapters. Chapter II discussed the psychological constructs of nonverbal behavior, nonverbal communication, emotional expression, and frustration. Chapter III outlined a proof of construct validity. In the following, a summary of these arguments is presented along with a discussion of how the observational instrument created out of this study met each of the conditions for validity and reliability.

Reliability and Interobserver Agreement

Reliability of the instrument was evaluated by the measurement of interobserver agreement. Because reliability is a necessary but not a sufficient condition for validity of the instrument, it is important to establish a high degree of reliability. Unfortunately, few theorists or statisticians have discussed what constitutes an acceptable level of reliability, at least in terms of the interobserver agreement measured for observational instruments (Scott, 1968). In other testing situations where reliability is measured through test-retest and similar

methods, some have stated that a reliability of 0.65 is acceptable (Cronbach, 1970).

The results of the interobserver agreement study are given in Appendix I. The individual interobserver agreement estimates range from 0.54 to 1.00. The average value for the interobserver agreement estimates was found to be 0.88. These values seem adequate compared to other studies involving systematic observation of nonverbal behavior.

A further argument for reliability can be made from the procedures used to construct the instrument and record the data. Considerable effort was made to objectify the data taking process. The goal was to make the definitions of the behaviors observed to be as precise and unambiguous as possible. Considering the short training period for the second observer, the acceptable interobserver agreement appeared to testify to the objectivity of the instrument.

More discussion of the interobserver agreement measurement procedure is included in Chapter III. Some technical notes related to the interobserver agreement procedure and calculations are included in Appendix II.

Validity of the Instrument

Validity is established by determining that the instrument measures what it claims to measure. Validity is a matter of degree and depends on the purpose of the instrument or evaluation technique.

In the case of the observational instrument, the focus was thus on whether the instrument measures (1) nonverbal behavior, (2) important nonverbal behaviors, and (3) important nonverbal behaviors elicited by frustration.

The literature review process and the use of pilot videotapes were designed to help insure that the instrument considered behaviors defined as nonverbal behavior. Nonverbal behaviors that were identified by previous study to be related to the kind of specific affect responses expected in the frustrational situation were selected for consideration. There was however a simultaneous filtering process that delimited certain

behaviors as a result of technical considerations, trade-offs, and the desire to keep the instrument objective and, where possible, quantitative. Most of this process was described in Chapter III. In summary, the construction of the instrument insured that it measures nonverbal behaviors. This construction process, together with the interobserver agreement results, argue for the validity of the instrument as a tool to measure certain nonverbal behaviors.

One of the best tests of the validity of the instrument rests on the construct validity arguments that can be made in relation to the testing of hypotheses two and three, especially hypothesis two. In theory, frustration will cause some affective changes, and these affect changes will cause certain changes in nonverbal behavior that will be related to that affect change and indirectly to the frustrating situation. If the observational instrument is a useful and valid instrument for observation of nonverbal behavior in this situation, then the instrument should detect changes in behavior. Hypothesis two tests not only a portion of the theories of frustration and nonverbal behavior but also the validity of the observational instrument.

Similarly, the testing of the third hypothesis could help to validate the instrument. Theory suggested that several reactions to frustration should be noted, including fixation, aggression, regression, and motivation. If several of these or other modes of reaction were seen in the frustrational setting of the individual testing, clustering analysis of the data from the observation of students should reflect these modes of response. Detection of modes of response to frustration through observation of nonverbal behaviors would certainly add to our knowledge of frustrational reactions in the classroom setting and would support the validity of the observational instrument as a useful tool for the observation of nonverbal behavior in this setting.

Cluster analysis of the data from this study did not show significant evidence for the hypothesized modes of response. However, it was also concluded that the data sample may have been too small for such cluster analysis and so the results are inconclusive.

Naturally, the failure to reject the null hypothesis for either or both of the hypotheses does not invalidate the theory or the observational instrument. Validation of an evaluation process or tool often is only gained through the repeated application of the instrument in a great variety of circumstances.

Summary of the Results for Hypothesis One

The focus of the first hypothesis is the establishment of the reliability and validity of the observational instrument. Reliability as measured by interobserver agreement was found to be quite adequate.

The establishment of the validity is not as straightforward, but the results are acceptable. First, construction methods for the instrument were chosen to insure the objectivity of the instrument. In addition an attempt was made to identify all nonverbal behaviors that would be important in the observation of students during frustrational tasks. The use of a videotape record allowed rather detailed and comprehensive analysis of the behaviors observed. There were some technical and practical considerations that restricted the use of a few potentially desirable nonverbal behaviors in the observational instrument. However, these restrictions were not great.

The other major argument for the validity of the observational instrument rested on the results of the testing of the other two hypotheses. The validity of the instrument as an instrument for the observation of frustrational behavior would be supported if that instrument could be used to detect significant behavioral changes and/or if it could differentiate and identify clustering in the data. In the testing of hypotheses, significant changes were detected through the use of the observational instrument, providing evidence for validity. Hypothesis three was not supported by the data and thus cannot be claimed as evidence to support the validity of the instrument. More discussion of the testing of hypotheses two and three is given below.

Hypothesis Two

Hypothesis two states that there is a difference in the nonverbal behavior elicited by students' performance of frustrating and non-frustrating tasks. As has already been explained, showing that definite differences in nonverbal behavior in the two situations can be detected through the use of the observational instrument establishes three main points. First, it supports the hypothesis as stated above. Second, it argues for the validity of the observational instrument. And finally, it further validates the theory and constructs on which the hypothesis is based.

The test of this hypothesis was performed by statistically testing data from the main study. The data were paired data with each pair of data points representing the pre-frustrational and the frustrational behavior of one student. Two sets of paired data resulted from the main study, one for the PPVT testing and one for the PIAT testing. The test statistic used was Hotelling's T-squared test. This calculation tested all variables simultaneously to determine if any overall change in the variables occurred.

The results of this testing are shown in Table 2.

Table 2. Results for test of behavior change using T-squared test.

	PIAT test	PPVT test
<u>All 17 variables</u>		
T-squared	152.28	332.75
Degrees of Freedom	17,15	17,21
F	4.33**	11.11**
<u>17 variables less TTA</u>		
T-squared	72.85	134.90
Degrees of Freedom	16,16	16,22
F	2.35*	5.01**

*Significant, $p < 0.05$

**Significant, $p < 0.01$

The results are presented both for all 17 variables and for all 17 variables except time to answer (TTA). Because one of the strongest and most obvious changes in student behavior for the frustrational and non-frustrational testing situations was the time to answer, the T-squared test was applied to the 16 variables without TTA to determine if there was significant change in other behaviors when TTA was not considered. The results showed significant behavioral change indicated with or without TTA.

These results clearly indicate that significant behavioral changes were found in the data taken with the observational instrument. The fact that the results indicate a significant difference even when TTA was omitted indicates the observational instrument could be applied to situations and tasks other than individual testing where the TTA variable would not be applicable.

When Hotelling T-squared test indicates significant change, it is common to apply other statistical analysis to help identify which of the observed behaviors (variables) contributed the most to the overall change noted. It is common practice to use individual tests such as Student's t-test to test each variable individually. Because of the multiple use of the test, the simple application of p-values or α levels may be misleading, but these calculations can certainly show which of the variables were most significant in the detected behavioral change.

The results of the individual t-test calculations are shown in Table 3. The calculation of the simple t-values were used only to give some indication of the magnitudes of the behavior changes and was done a posteriori to determine which of the behaviors contributed the most to the overall behavior change. The term significant is used in a loose way to provide a standard to identify the strongest behavior change.

The results of the individual t-tests indicate which of the variables were most important to the overall behavior change that was noted by the T-squared test. The largest change was indicated by the time to answer (TTA). However, there were several very strong changes

Table 3. Results of individual t-tests for behavior change.

Variable		PIAT (df=31)	PPVT (df=37)
TTA	Time to Answer	7.52***	11.06***
AMTM	Amount of Movement	-.62	0.33
DSTM	Distance or movement	0.42	0.67
PSFT	Posture Shift	-0.52	-1.13
LEAN	Lean Back	1.43	2.92**
HSHK	Head Shake	0.49	1.55
MANP	Self Manipulation	2.55*	2.40*
HTH	Hands to Head	0.48	1.08
GAZE	Gaze	-4.85***	0.92
EBFL	Eyebrow Flash	2.06*	3.61***
EBLO	Eyebrow Lowering	2.17*	2.13*
MTWT	Mouth Twitch	3.66***	3.64***
SMIL	Smile	-2.87**	-0.55
LICK	Lip Licking	-0.33	0.90
MDRP	Mouth Drop	2.46*	2.96**
FLPA	Filled Pause	0.22	0.27
SUBV	Subvocalization	4.16***	4.37***

*p < 0.05

**p < 0.01

***p < 0.001

indicated for other behaviors, such as mouth twitching (MTWT) and subvocalization (SUBV). Because the data for the PIAT and the PPVT testing were independent, it is especially interesting to note the similarities in the t-test results for the two tests. Significant behavior changes were noted for both tests for time to answer (TTA), self manipulation (MANP), eyebrow flashing (EBFL), eyebrow lowering (EBLO), mouth twitching (MTWT), mouth dropping (MDRP), and subvocalization (SUBV). Other behaviors that showed significant change for one but not both of the tests include gaze (GAZE), smiling (SMIL) and the

lean back (LEAN). Further discussion of these similarities and differences is included in Chapter V.

In summary, the test results of the main study indicate that there was a significant change in nonverbal behavior between the non-frustrating and the frustrating portions of the two individual tests. This finding allows for the rejection of the null form of hypothesis two, i.e., there was a significant change in nonverbal behavior for students performing frustrating and non-frustrating tasks. The findings also contribute to the validity of the observational instrument and the acceptance of hypothesis one. The results of the individual t-testing, although not directly related to the test of hypothesis two, indicate that certain variables were most significant in the detected behavior change.

Hypothesis Three

Hypothesis three states that there is clustering of measured non-verbal behavior elicited in response to frustrating tasks. Although this hypothesis was investigated through the use of computer aided statistical analysis, cluster analysis is a descriptive rather than an inferential technique and results cannot be directly translated to an acceptance or rejection criterion. The following discussion examines the process used to test the hypothesis. The cluster analysis and the judgement of the results involved several steps. The process is diagrammed in Figure 7. Each of the steps of the process is numbered, and a brief description of each numbered process is given below.

1. The data used were the observations of behavior from the frustrational portions of the testing only. Both the PIAT and PPVT were used.
2. Since there was no a priori knowledge of which variables would be most important in the clustering, the first attempt at clustering used all 17 variables. However, a concern about the size of the data samples relative to the number of variables resulted in a decision to attempt clustering with some reduced number of

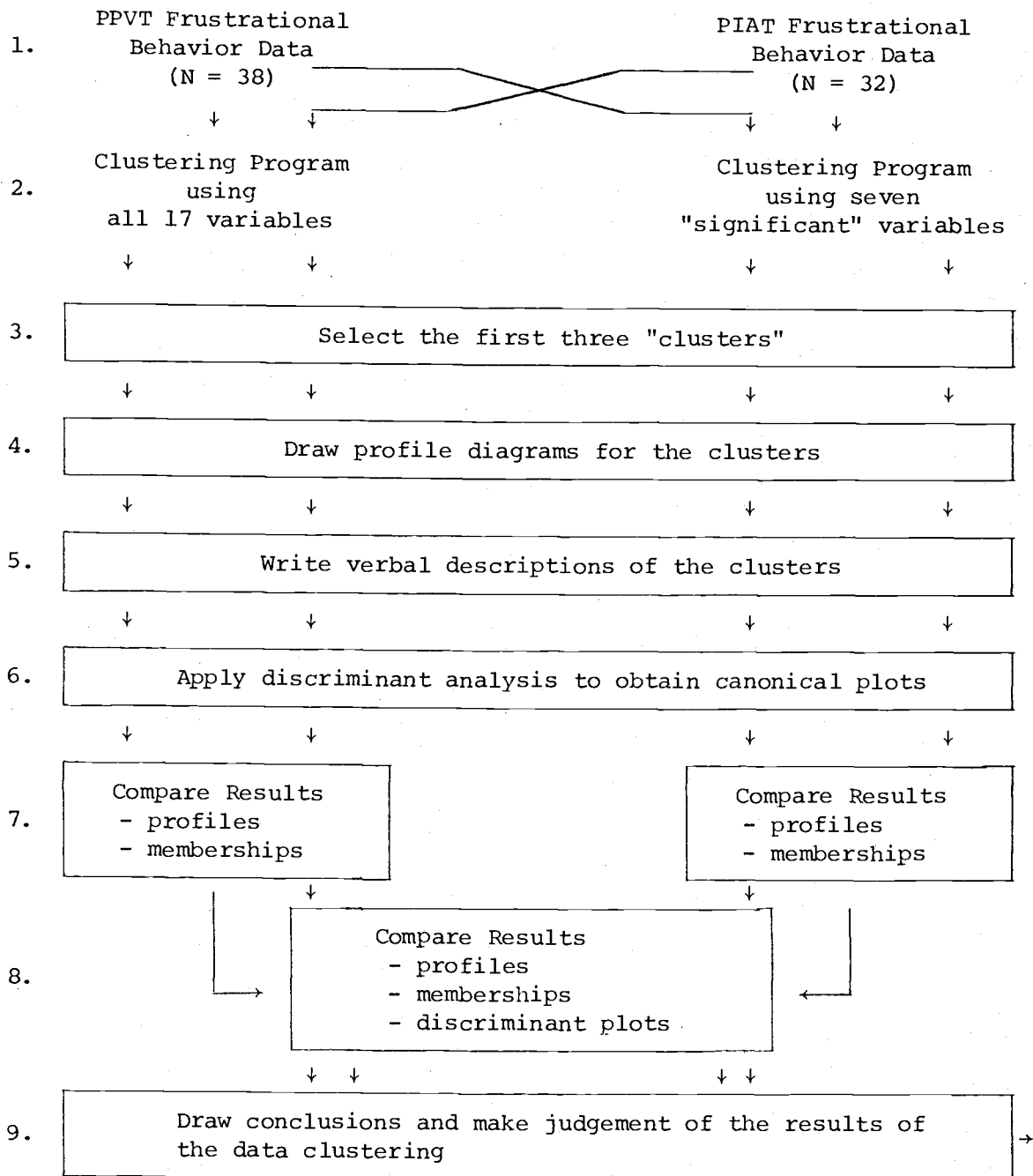


Figure 7. The process of judging the cluster analysis results.

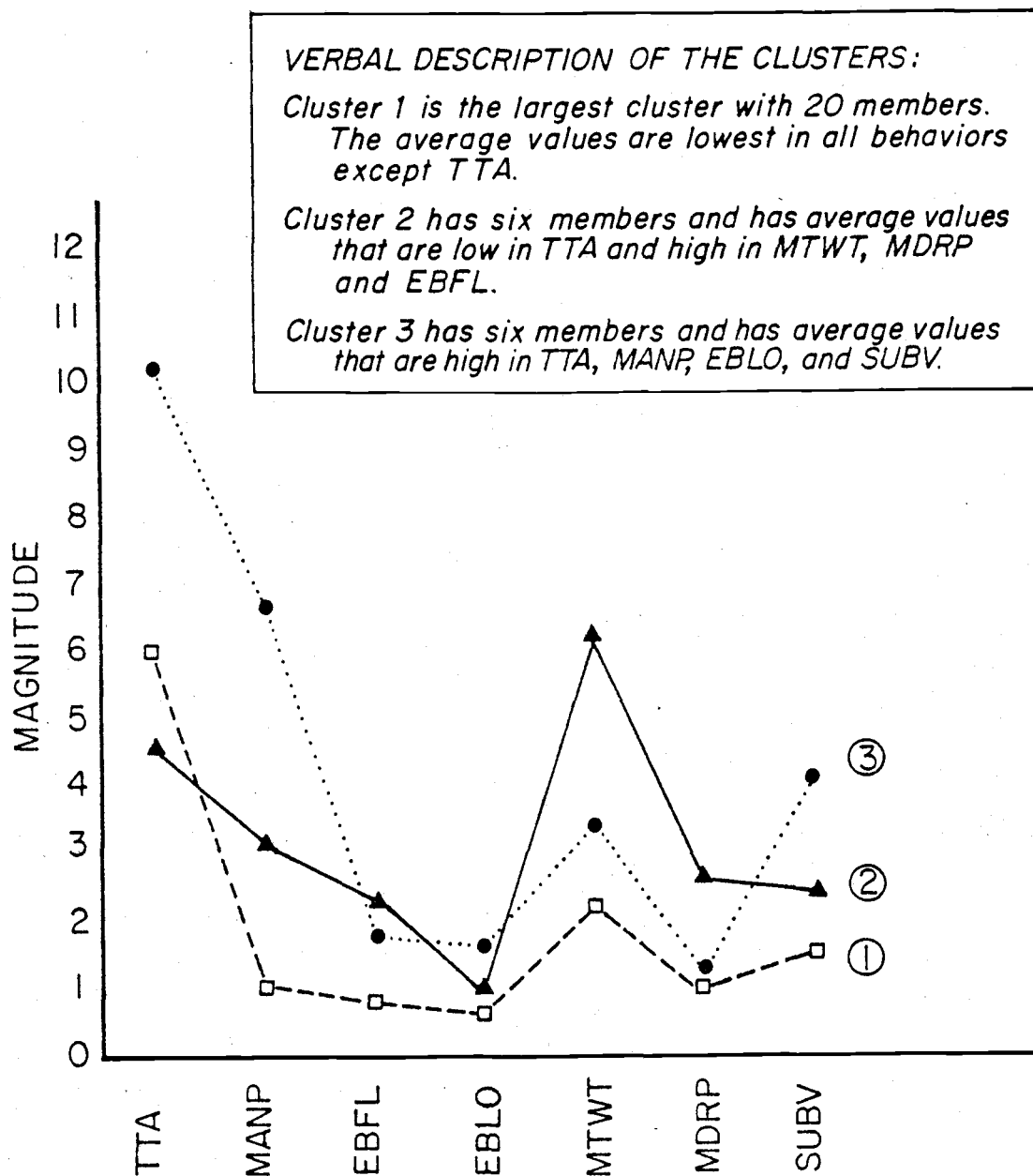


Figure 8. Profile plot and verbal description. An example - PIAT data, seven significant variables.

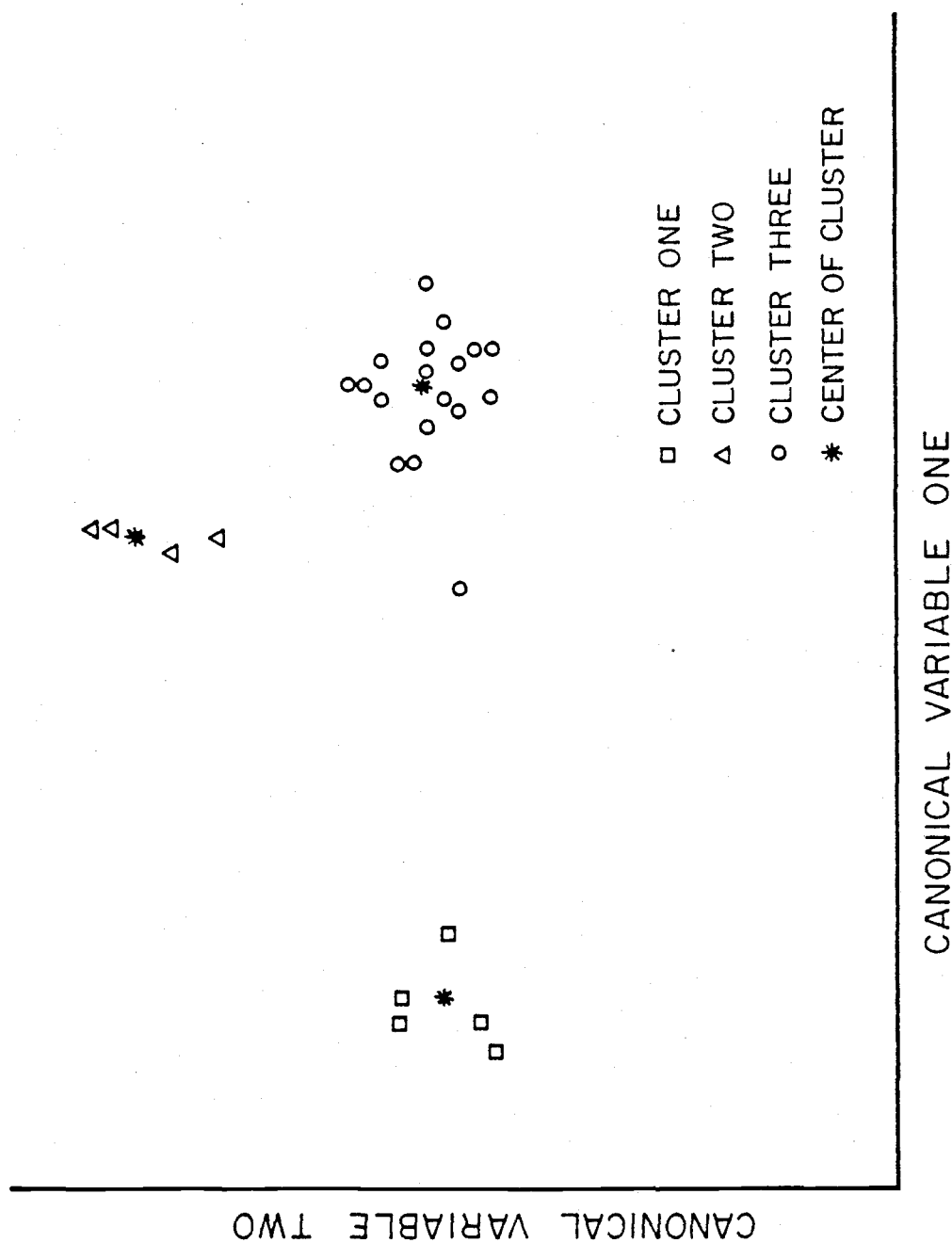


Figure 9. Discriminant plot of three clusters. PIAT behavior data - all variables.

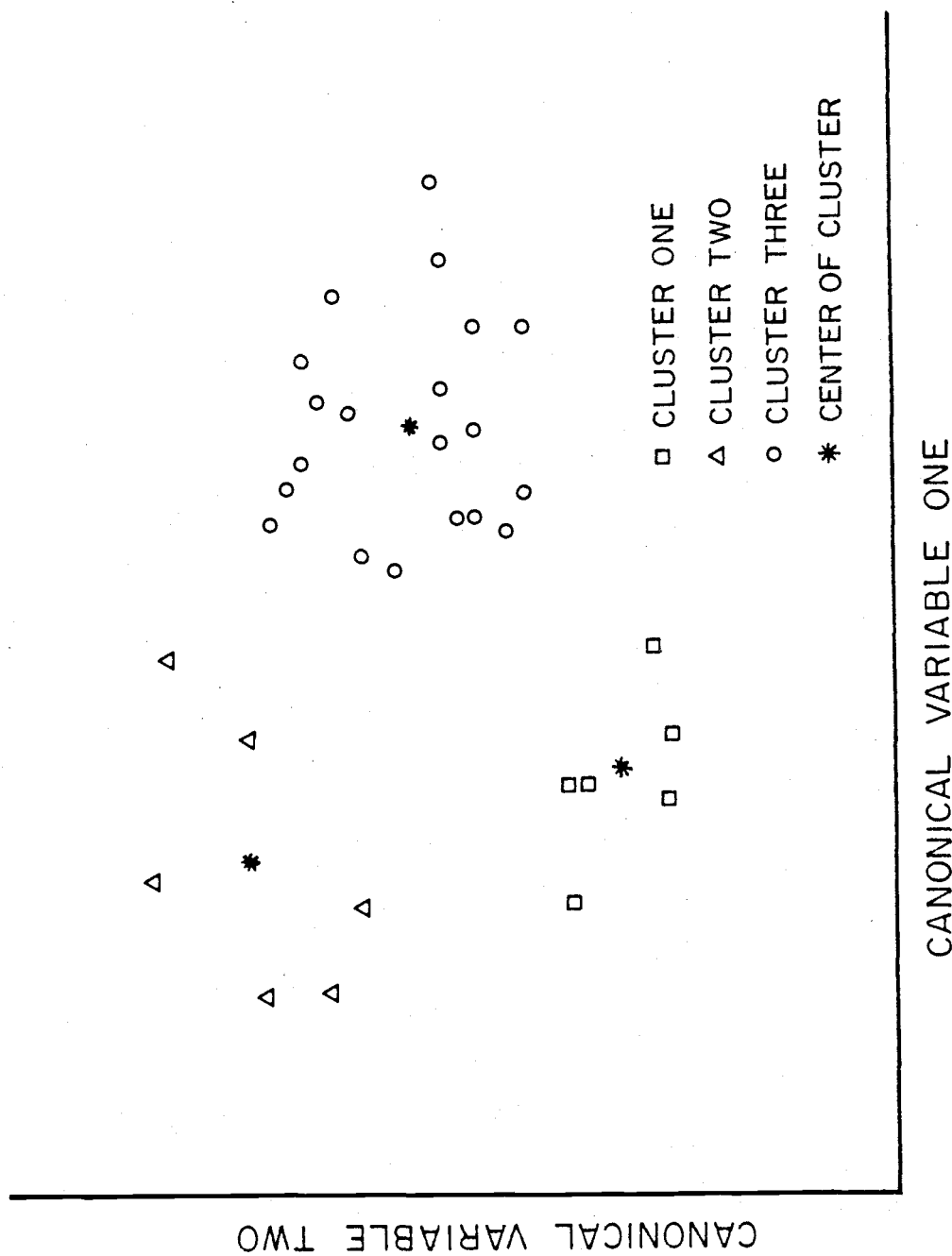


Figure 10. Discriminant plot of three clusters. PIAT behavior data - seven significant variables.

variables. The criterion used in the reduction of the variables referred to the results of the Student's t-testing of individual behaviors presented in the section on hypothesis two. The seven variables that were significant at the 0.05 level for both the PIAT and the PPVT were used for this second clustering attempt. For this list of seven significant variables, the term significant is used only for convenience to identify this reduced set of variables. These seven variables include TTA, MTWT, MANP, EBFL, EBLO, MDRP, and SUBV.

3. Clustering of the data was accomplished through the use of the clustering program described in Appendix IX. The program starts by considering all of the students or data points as one cluster and then divides the data into two clusters, three clusters, and so on. The first three clusters identified by the computer were selected for further study, subject only to the constraint that clusters with only one or two members were considered to be outlier data and ignored. Thus the result of the clustering was to separate the students into three groups, or clusters, on the basis of similarities in their behavior data. This clustering was performed for both the PPVT and the PIAT data using both all 17 variables and the reduced seven significant variables, so that four sets of clustered data were obtained.
4. The cluster program output included average values for each of the variables within each of the clusters. Graphing of these average values in the form of a profile plot allowed a visual comparison of the clusters. An example of such a profile plot is given in Figure 8.
5. From the profile plots verbal descriptions of each cluster were written based on the profile plots. An example of such a description is provided as part of Figure 8.
6. In order to graphically display the clustering in variable space, the clusters were subjected to discriminant analysis. As part of this discriminate analysis, a plot was constructed that showed the data points projected on two dimensions where the two

dimensions represented the first two canonical variables, or linear combinations of the original variables. Basically, this plot uses the two dimensions in which the most separation between the data points can be seen (Dixon and Brown, 1979). Examples of such plots are given in Figures 9 and 10.

7. Comparisons of the clustering results helped to determine the consistency of the results. It was conjectured that if there were significant clustering into modes of behavior, then the clustering of the PIAT and PPVT data should both show similar clustering. Comparisons to identify this consistency included comparison of the profile plots and the verbal descriptions for the two tests. Little similarity was found. It was further conjectured that there should be enough similarity in the behavior modes that the membership of clusters should be similar for the two tests. Again little similarity was found in the membership of the clusters between the PIAT and the PPVT.
8. In an effort to identify the effects of the reduction of the number of variables from seventeen to seven, the results under the two separate conditions were compared. Again both profiles membership lists, and discriminant plots, were inspected. The most interesting result of this comparison was the significant decrease in the apparent clarity of the discriminant plots when the number of variables was decreased (Figure 10). It was concluded that the better looking plots, using all 17 variables, were an aberration caused by the relatively small sample size in comparison to the number of variables.
9. Final consideration of all of the above results lead to the conclusion that these data do not provide enough evidence to justify the rejection of the null hypothesis that no clustering occurred.

Findings Not Directly Related to the Hypotheses

This section of the chapter presents some results not directly related to one of the three hypotheses. The results presented include

correlation matrices for the variables used in the observational instrument, results of discriminant analysis used to determine how well the observational instrument data can discriminate between frustrated and non-frustrated students, and, finally, a listing of nonverbal behaviors that were observed infrequently during the main study and may be related to student communication of frustration.

Correlation Matrices

The main purpose for the determination of the correlations between variables is to investigate the degree of independence of each of the variables used in the observational instrument from other variables. A secondary purpose results from the fact that when strong correlations exist between variables, some multivariate statistical techniques are affected. For example, the coefficients and therefore the form of discriminant functions or linear regressions are sometimes misleading and hard to interpret when correlations are very large. High correlations could also effect the results of cluster analysis. Although the Hotelling's T-tests would be unaffected, the results of individual t-tests may be hard to interpret when large correlations exist.

Fortunately, the results of the correlation studies showed small correlations between variables. Results of the correlation studies are shown in Tables 4 and 5. These results reflect correlations between variables during the frustrational part of the testing for both the PIAT and PPVT testing.

The correlations shown in Tables 4 and 5 show that the strongest correlations exist within the movement variables MAMT, MDST, LEAN and PSFT. This finding is not surprising when the definitions of each of these movement variables is considered. There are some obvious relationships. For most of the other variables used in the observational instrument the correlations are small.

Table 4. PIAT data correlation matrix.

88

CORRELATION MATRIX

		TTA 1	MAINT 2	MOIST 3	PSHFT 4	LEAN 5	HSNK 6	MANIP 7	HTH 8	GAZE 9	EBFLSH 10
TTA	1	1.0000									
MAINT	2	-.1759	1.0000								
MOIST	3	-.0380	.7755	1.0000							
PSHFT	4	.0069	.1159	.1061	1.0000						
LEAN	5	.0847	.2651	.1439	.1025	1.0000					
HSNK	6	-.0101	.0203	-.0762	.0231	.3784	1.0000				
MANIP	7	.3383	-.1128	-.0724	.0966	-.1280	-.0377	1.0000			
HTH	8	.2269	-.0595	.0335	.1090	-.1451	-.0222	.6370	1.0000		
GAZE	9	-.0320	-.3253	-.2540	.0199	-.1619	-.2012	.1175	.0515	1.0000	
EBFLSH	10	.3204	-.0285	.3118	-.0663	-.0346	.0955	.1454	-.2370	-.0856	1.0000
EBLOW	11	-.2476	.0158	-.0080	-.1441	-.0817	.2943	-.1710	-.0515	-.1728	.2772
MTWICH	12	.1221	-.1025	-.1842	.0480	-.1167	-.0729	-.0507	-.2130	.0505	-.0245
SMILE	13	.1728	.1175	.1625	.2048	.0090	-.0166	.2824	-.3629	-.0705	.4001
LICK	14	-.1713	.0445	-.0342	-.1371	.0740	.0344	-.2643	-.2242	.0804	-.1895
MORP	15	-.1139	-.1490	-.2168	.1501	-.0485	.1142	-.1338	-.0451	.2796	-.0295
FILOPSE	16	-.1158	-.0054	-.0843	-.0364	-.1415	-.1468	-.1811	-.1538	.2905	-.0799
SURVOC	17	.0654	.1891	.2140	.3681	.2088	.2942	-.1418	-.0885	.1382	.0577
		EBLOW 11	MTWICH 12	SMILE 13	LICK 14	MORP 15	FILOPSE 16	SURVOC 17			
EBLOW	11	1.0000									
MTWICH	12	-.0220	1.0000								
SMILE	13	.0318	-.1280	1.0000							
LICK	14	.0198	.0875	.3316	1.0000						
MORP	15	-.0833	.0572	-.0415	-.1590	1.0000					
FILOPSE	16	-.1057	.1737	.0845	-.0124	.4405	1.0000				
SURVOC	17	-.0870	-.1271	.0207	-.0704	.3109	-.1926	1.0000			

Table 5. PPVT data correlation matrix.

Discriminant Analysis of the Data

Discriminant analysis is used to determine the separability of individuals into two or more groups based on a series of measurements on each individual (Harris, 1975). In this study, the aim of the discriminant analysis was to determine how well students can be classified as frustrated or not frustrated on the basis of the nonverbal behavior data.

The discriminant analysis involves the construction of a mathematical function which optimizes the separation of the groups using all of the data and provides a criterion for judging which group a data point belongs to. To test the quality or accuracy of this mathematical function, data on known frustrated and non-frustrated students are placed into either the frustrated or non-frustrated category by the discriminant analysis criterion. The results of such a classification of known behavior data can be used to check the accuracy of the classification criterion and discriminant function. In this study, data from the frustrated and non-frustrated behavior observed in the main study were used to construct the discriminant function. Then, using a jackknifing procedure, the same data were used to check the accuracy of the discriminant function classifications.

The results of this discriminant function analysis are summarized in classification tables for both the PIAT and PPVT test results (Table 6). The analysis was accomplished using data from all 17 variables

Table 6. Discriminant function classification accuracy results for PIAT and PPVT data.

	Accuracy of Classification	
	PIAT	PPVT
All 17 variables	84.4%	85.5%
All variables less TTA	76.6%	71.7%

and from all variables except time to answer (TTA). Where all of the variables were used in the discriminant function analysis, the classification accuracy was 84.4% and 85.5% for the PIAT and PPVT tests, respectively. Where TTA was not used, the classification accuracy was 77% and 71% for the PIAT and PPVT test data, respectively.

Some interesting comparisons can be made of the discriminant function classification accuracies and the judgement accuracies found in the study of Jecker, Maccoby and Brietrose (1965). This comparison and a discussion of these findings is included in Chapter V.

Other Nonverbal Behaviors Observed

Figure 11 lists the other nonverbal behaviors observed during the analysis of the main study videotapes. As already indicated, the statistical analysis of nonverbal behavior data was limited to only those behaviors that were observed at least one time in more than 25 percent of the students studied. This limitation excluded only six variables that had been selected for the observational instrument from the pilot study results. However, during the main study, a number of other nonverbal behaviors were observed in the behavior of more than one student, and although they occurred too infrequently to be included in the statistical analysis, they may still be communicative of the frustrational situation.

These behaviors were noted under the "other comments" section of the observational instrument record sheet. An attempt was made to record and describe essentially all observed behaviors. As the study proceeded, most of these other behaviors were named and described in a manner similar to the behavior descriptions of the observational instrument. This list and description of behaviors was helpful in making the notation of these behaviors more quick and concise. The description of each of these other behaviors is given in Appendix VII. The list of behaviors in Figure 11 is given in the approximate order of frequency, with the most frequent listed first.

-
1. Nose Movement - any movement or wrinkling of the nose.
 2. Lip Biting - biting of the upper or lower lip.
 3. Question Repeat - part or all of the question is repeated by the subject.
 4. Questioning Answer - inflection of the voice indicates that the answer is given in a questioning way.
 5. Sigh - an audible exhaling of breath.
 6. Don't Know - the subject states that he/she doesn't know an answer.
 7. Rocking Foot - a rapid rocking of the foot.
 8. Shoulder Shrug - a rapid raising of the shoulders.
 9. Head Cocked - a tilt of the head sideways.
 10. Head Wag - movement of the head side to side.
 11. Head Toss (Hair Toss) - throwing the head back with a quick motion.
 12. Wide Eyes - a rapid widening of the eyes.
 13. Closed Eyes - closing the eyes longer than a blink.
 14. Tongue in Cheek - the tongue pushing out the cheek.
 15. Tight Smile - the mouth pulled back on both sides.
 16. Swallowing - swallowing seen in observation of the throat.
 17. Laugh - any laughing.
 18. Quiet answer - barely audible student response.
 19. Drumming the Table - a type of manipulation where the table is tapped with one or several fingers.
-

Figure 11. Other nonverbal behaviors observed.

Summary of the Results

The findings presented in this chapter include data from the main study and the results of analysis of that data. Most of the findings are related to the study's three hypotheses, but several findings that are not directly related to the hypothesis testing are also included.

The findings related to the first hypothesis confirm the observational instrument developed in this study as a valid and reliable instrument. The main evidence supporting this finding includes (1) the methods used in the design of the instrument; (2) the reliability measured by interobserver agreement; and (3) the detection of significant differences in nonverbal behaviors between frustrational and non-frustrational testing. The findings of the discriminant analysis also support the validity of the observational instrument.

The findings of the main study support the second hypothesis that significant differences in nonverbal behavior are elicited in response to the frustrational testing situation. The data from both the PIAT and PPVT testing showed significant changes between the pre-frustrational and the frustrational portions of the testing.

The findings related to the third hypothesis do not support the idea of clustering in nonverbal behavior data. Although some clustering could be detected in the data, these clusters lacked consistency between the two testing situations, the PPVT and the PIAT. One problem that may inhibit the clear detection of clustering is the small number of data points relative to the number of variables.

Findings not related to the hypothesis testing include correlation analysis of the nonverbal behavior data, discriminant analysis of the pre-frustration and frustrational data, and a listing of other behaviors observed during the main study. The correlation analysis showed few high correlations among the behavior data. This finding is important for the interpretation and use of some multivariate analytical techniques. The discriminant function analysis is related to the findings of the second hypothesis in that it not only shows that a significant difference in behavior exists, but that the behaviors can

be used to correctly classify a high portion of the students as frustrated or not frustrated on the basis of the observational data alone. The listing of the "other behaviors" is included as an important outcome of the main study, since many of these behaviors, although infrequent, may be quite communicative of the frustrational state of the student.

V. SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

This final chapter presents a summary of the study plus overall conclusions and recommendations. The chapter is divided into two main sections. The first section summarizes the entire study along with the major findings. The second section presents conclusions and recommendations and relates the findings of this study to other research and theory.

Summary of the Study

The Problem

There is a lack of data on student nonverbal behavior in the classroom. Little research has been directed at the process of student encoding of nonverbal behavior. Of the work that has been done on encoding, most has been directed at detection of nonverbal communications of attitude or gross affect.

The ability to decode nonverbal behavior could be important in the educational setting. Some studies have been done that suggest the ability to correctly judge or decode nonverbal behavior is a function of familiarity with the student and that this ability can be improved with training. Specialists in evaluation have suggested that important information can be learned from the nonverbal behaviors students express during the individual testing process.

In this study the overall goal was to study the nonverbal behaviors and behavior changes that are elicited when students face frustrational tasks. More specifically, the goal was to develop an observational instrument for the observation of nonverbal behaviors elicited by frustrational tasks within the individual testing setting and to use this instrument to investigate the relationship between frustration and nonverbal behavior.

The three major hypotheses to be tested by the study follow:

1. A valid and reliable instrument can be constructed for the systematic observation of nonverbal behaviors elicited by frustrating tasks in the classroom;
2. There is a difference in the nonverbal behaviors of students performing frustrating tasks and nonfrustrating tasks.
3. There is a significant clustering of student nonverbal behavior elicited in response to frustrating tasks.

Design of the Study

The study was broken into two phases, the pilot phase and the main study. In the pilot phase the major objectives included the selection of the specific tasks to be used to stimulate frustration, the construction of the observational instrument, the gathering of some pilot videotape data, and the trial application of the observational instrument.

Some of the pilot videotapes viewed individual testing administered by staff psychologists in the Corvallis school district. After this observation, the decision was made to use two individualized tests to stimulate frustration. The tests chosen were the Peabody Picture Vocabulary Test (PPVT) and the mathematics sub-test from the Peabody Individual Achievement Test (PIAT). Later in the pilot phase more videotapes were taken using only the PIAT and PPVT tests as part of the preparation for the main study.

The observation of the pilot videotapes and a literature review assisted in the construction of the observational instrument. The observational instrument that was developed consisted of a listing and definition of a variety of behaviors, a record sheet, and rules for the determination and scoring of nonverbal behaviors during a one minute sample of behavior. The instrument described and defined the behavior categories and how they are recorded over the one minute sample of behavior.

After the instrument was constructed, it was applied to the pilot videotapes in order to refine the procedures and definitions of each category. The data from pilot tapes were limited, but some analysis of these data helped in the planning of the statistical analysis and in the identification of behavior categories that might be significant in the main study.

The data for the main study was collected during the months of March and April of 1979. A sample of 39 eighth grade students was used. Each student was videotaped while the two tests, the PIAT and the PPVT, were administered. From the videotape records two one minute samples of behavior, one frustrational and one non-frustrational, were identified for each student for each test and the observational instrument was applied to each. In each case the non-frustrational behavior samples came from the start of the testing where the test material was easiest for the student; the frustrational behavior sample came from the end of the testing where the material was hardest. The application of the observational instrument to each behavior sample yielded a series of scores, usually a number of behaviors per minute.

Hypothesis one was tested by establishing the reliability and validity of the instrument. Two independent observers applied the observational instrument to the same videotapes in order to measure interobserver agreement; interobserver agreement has been used as the measure of reliability of an observational instrument. Validity was established through arguments of construct validity based on the process used to design the instrument and on the results of the tests of hypotheses two and three.

Hypothesis two was tested through a statistical comparison of the data from the pre-frustrational and the frustrational behavior. Both Hotelling's T-squared test for all variables simultaneously and Students' t-test for each variable individually were used. Hypothesis three was tested through the application of computer aided cluster analysis aimed at the identification of clustering within the behavior data from the frustrational portion of the testing. Other data analysis included procedures and tests not directly related to the specific

hypothesis testing. This included correlation studies and discriminant analysis.

Results of the Pilot Phase of the Study

Some of the results of the pilot phase have already been identified above. The two primary outcomes of the pilot phase of the study were (1) the selection of the two tests (PIAT and PPVT) to be used as stimuli for frustration in the main study and (2) the design and construction of the observational instrument.

The two tests that were chosen for use in the main study had the advantage of being standardized, individually administered tests in which the prescribed administration procedures require that the student be faced with a series of questions that range from very easy to very difficult. The tests were designed so that the examiner need have little special training, and the physical set-up recommended for test administration was compatible with the use of videotape recording.

The observational instrument developed during the pilot phase of the study has already been briefly discussed. Basically, the instrument was designed to objectify and tabulate as much nonverbal behavior as possible during a one minute sample of behavior. The use of the videotape records allowed for intensive observation and review of the behavior record. Although the instrument record sheet was designed to record behaviors in the sequence in which they occur, the scores taken from the record were primarily a tabulation of the total number of behaviors per minute for each of the behavior categories.

Results of the Main Study

The reliability was confirmed by measurement of interobserver agreement. An overall interobserver agreement of 0.67 was found; the results were quite consistent with similar studies using systematic observation of nonverbal behavior. The validity of the instrument was

demonstrated by both the procedures and methods used in the design of the observational instrument and by the results of the test of hypothesis two which showed that the instrument was sufficient to detect behavioral changes induced in the frustrational situations. Further validation of the observational instrument came from a discriminant analysis of the data that showed that the instrument could discriminate between frustrated and non-frustrated students with an accuracy similar to that found in judgement studies using teachers or parents to judge comprehension of students (Jecker et al., 1965; Allen and Atkinson, 1977).

The test of hypothesis two showed that the behavior observed does show significant change between the frustrated and non-frustrated testing conditions. Beyond this, individual t-tests identified which of the nonverbal behaviors contributed most to the change. This list of most significant behaviors was almost the same for both of the tests.

The findings related to hypothesis three were mixed. Overall they did not establish conclusively the existence of clustering in the behavioral data. It was not clear whether or not a larger sample size would help identify the hypothesized clustering and behavioral modes.

Conclusions and Recommendations

The Relationship of the Current Findings to Other Studies

Betts' Work and Informal Diagnosis of Reading

One of the important results of this study is the objective demonstration of behavioral changes shown by students in the frustrational testing situation. These findings confirm the observations of Betts (1957) and others who quite perceptively indicated that such

nonverbal behavior changes occur and considered observation of such nonverbal behaviors an integral part of the diagnostic process. Recent trends in testing appear to be away from more informal and judgemental evaluation practices and toward the more quantifiable scores such as percent correct or number of words per minute. This study indicates that significant information is communicated nonverbally and supports the notion that nonverbal behavior observation can be a valid part of some school diagnosis.

The Study of Jecker, Maccoby, and
Brietrose and Others

The findings of the studies of Jecker et al. (1964, 1965) showed that judgement of student nonverbal behavior could be used to assess student comprehension and that improvement in judgement skill could be obtained through the use of some simple inservice training. This training consisted largely of practice in the use of an observational checklist. The checklist developed by Jecker et al. was informal and qualitative in style and designed primarily as a training aid.

The methods of Jecker et al. (1964, 1965) were used and expanded in the studies of Allen and his colleagues (1975, 1977, 1978). Allen and Atkinson (1977) compared the abilities of teachers and parents to correctly judge student comprehension through observation of nonverbal behavior.

There is a relationship between these comprehension studies and this study if the assumption is made that comprehension or lack of comprehension reflect student frustrational state. In this study, difficult stimulus material, which was identified by the number of correct or incorrect responses was defined to be frustrating. The assumption seems well founded here since only very easy and very difficult stimulus material was used in both this study and in the Jecker et al. and Allen et al. studies.

The results of this study largely confirm and validate the behaviors listed in the checklist used by Jecker et al. and Allen et al. in the communication of frustration. This fact also tends to

confirm the notion that the studies of judgement of student comprehension could be viewed as studies of student frustration and visa versa.

The relationship between task difficulty and frustration in the everyday sense of the word is an open area of research. The present study suggested that for such research frustration might actually be defined in terms of nonverbal behavior changes as indicators of affect change. Perhaps affect changes could then be more continuously monitored while task difficulty was varied. There certainly needs to be more research that could clarify the relationships between task difficulty, comprehension, frustration, affect, and nonverbal behavior.

It is interesting to compare the judgement accuracy of pre-service teachers in the Jecker, Maccoby and Brietrose study (1965) to the accuracy of the discriminant analyses results presented in Chapter IV. In the Jecker study, the overall accuracy of judgement of pre-service teachers in identifying easy versus hard lesson material from the student nonverbal behaviors was 56.7 percent correct. In this study, the overall accuracy of identification using the discriminant function was about 85 percent when the time to answer variable (TTA) was included and about 74 percent when the TTA variable was not included (Table 6).

Comparison of the two judgement situations may provide clues as to the differences in judgement accuracy. In the Jecker study the stimulus videotapes showed about 45 seconds of student behavior which was taken while the students were listening to either an easy lesson or a hard lesson. In this study, behavior data was gathered during a 60 second sample of student behavior from the pre-frustrational or frustrational portion of the individual testing. In the Jecker study, some comprehension questions were asked at the end of each short lesson, but videotape of the question-answer process was not used in the judgement of the students. One major difference that must have contributed to the differences in judgement accuracy was that in the present study the individual test situation involves a two-way communications process. Verbal communication was taking place in the question-answer sequence, and much of the non-verbal behavior that was seen may have been

elicited as a byproduct of the verbal communications. Another hypothesis is that frustration does not really take place, even without comprehension, until the student is questioned and thus forced to respond. Then the frustration caused an affect response that could be seen in nonverbal behaviors. This definition of frustration corresponds more to Maier's ideas (Maier, 1949).

Comparison of Results for the Two Tests - PPVT and PIAT

Because two different individualized tests, the PIAT and PPVT, were used for the stimulus of frustration in the main study, comparison of the similarities and differences in the results can provide some data on the external validity of the study. Comparison of results also suggests interesting insights into hypotheses about the differences in observed nonverbal behavior.

The most significant finding about the two tests was the fact that both of the test situations elicited significant behavior changes. This finding alone supports the external validity of the study; it suggests that similar changes in nonverbal behavior may be noted in other similar testing situations.

In the statistical analysis that followed the finding of significant overall behavior changes, individual t-tests identified individual behaviors which contributed most strongly to the observed behavior change. The results of the t-tests applied to the two test behaviors show strong similarities and some interesting differences (Table 7).

The difference in the speed or pace of the two tests might help to explain some of the differences noted in the elicited nonverbal behaviors. For the PPVT testing the number of questions per minute was significantly greater and the time to answer was smaller (Table 8). Differences in the pace of the testing were related to the nature of the testing material itself. In the PPVT each question was a single word; in the PIAT questions often required a sentence or two of explanation. Similarly, the PPVT tasks require only vocabulary

Table 7. Comparison of behaviors elicited by the PIAT and PPVT.

PPVT t-test Results	PIAT t-test Results	
	Significant*	Not Significant
Significant*	TTA	EBLO
	MIWT	MDRP
	MANP	SUBV
	EBFL	
Not Significant	GAZE	AMTM
	SMIL	HTH
		DSTM
		LICK
		PSFT
		FLPA
		HSHK

*Significant at the 0.05 level

Table 8. Average values for QPM and TTA for the PIAT and PPVT.

Variable	Pre-frustration	Frustration
Questions per minute (QPM)		
PPVT	9.02	6.74
PIAT	5.91	2.88
Time to answer (TTA)		
PPVT	2.13	4.39
PIAT	2.06	10.93

recognition, but in the PIAT math subtest students sometimes had to do some mental calculation in order to answer the question.

Another possible factor in the differences could be related to the idea of synchrony in verbal dialogue. Research has shown that the length of the response in a dialogue depends on the length of the communication of the other party (Harper et al., 1978, p. 47-51). If one person responds in shorter statements the other person will usually shorten his/her response. Similarly, longer pause times by one party will tend to induce longer pause times in the other party. It is possible that these phenomena contribute to the quicker pace and shorter time to answer noted since the shorter questions of the PPVT may induce quicker responses by students.

These differences in the format of the testing may provide some clues as to the nature of or reasons for the similarities and differences in the nonverbal behaviors noted. For example, in the case of Gaze and Smiles, which showed significant change in the PIAT but not the PPVT, it appeared that the slower pace at the start of the PIAT testing allowed the student some chance to look up and smile occasionally between questions. Later in the testing there was less looking-up and the behavior change was noted. For the PPVT there was little time to look up during any part of the test because of the faster pace.

As another example, there seemed in general to be more nonverbal behavior elicited during the PIAT testing. It appeared that nonverbal behaviors often occurred during the time while the student was thinking about possible answers, i.e. between the asking of the question and the student's answer. The longer time to answer (TTA) appeared to allow for more nonverbal expression.

Discussion of Significant Nonverbal Behaviors Observed in the Study

This section reviews many of the nonverbal behaviors identified in this study and relates them to other research findings. The nonverbal behaviors that were more frequently observed or found to be most

strongly related to the frustrational testing are of primary interest. However, some nonverbal behaviors have been more researched than others; therefore, discussion is often limited by the degree to which other researchers have investigated these particular behaviors.

Micromomentary Facial Expression

Haggard and Isaacs (1966) have used the term micromomentary facial expression to describe facial expressions that last less than one-third of a second. They noticed these facial expressions first in the analysis of videotapes of patients in the psychotherapeutic setting. Haggard and Isaacs then investigated the ability of untrained observers to detect these expressions and concluded that although micromomentary expressions often went unnoticed, but sensitivity could be increased with training. Haggard and Isaacs assumed that these facial expressions were important indicators of emotional reaction, such as reactions to sensitive topics in a patient's life history. The short duration of the facial expression was thought to be the result of the subject's containment of emotional expression due to social or psychological repression mechanisms.

Haggard and Isaacs used judgement to identify the expressed affect of each expression. Videotapes were slowed or stopped to more clearly judge the expression. Their work could therefore be considered more of a judgement or decoding study; therefore little information was available as to the specific facial movements that commonly constituted such micromomentary facial expressions.

In the present study several of the encoded nonverbal behaviors could be considered micromomentary expressions. Mouth twitching for example was often very short in duration and was sometimes concurrent with one or more other behaviors such as an eyebrow flash. The finding that untrained observers have difficulty in detection of such facial expressions is significant. In fact, many of the nonverbal behaviors used in this study were difficult to observe and record reliably without checking and rechecking the videotape records. Ease of use of the

observational instrument did improve with time indicating improved sensitivity with training.

Mouth Twitch

In this study mouth twitch was significantly related to frustration. Little has been written specifically about the mouth twitch behavior except as in relation to micromomentary facial expression. Quite a variety of mouth twitches were observed in this study. With more sophisticated videotape equipment the mouth twitch category might be broken into subcategories such as smile, frown, and neutral. Obviously this would involve more observer judgement and subjectivity.

Lean Back

The lean back was the only category of body movement or position that showed significant relation to the frustrational testing. Eggen (1972) found some relationship between body lean and gross affect. However, in his experimental set-up the subject was in a classroom lecture, and he was investigating gross affect not specific affect. In another study, Mehrabian (1972) showed a link between body lean and affiliation or liking, but his stimulus situation was usually a loose social setting such as a waiting room. It is interesting to conjecture that lean backs observed during frustration might arise from similar affect states in all three cases, that of dislike.

Body Movements

In this study there were several complications associated with the experimental setting that might influence body movement, posture, and lean. First, the subject had to stay close enough to the testing material so that he or she could see the test materials clearly. Second, movement of the body or head also influenced the ability to see the test material clearly. Third, students appeared to maintain fairly

motionless posture while listening to the questions and then while looking for the answer. Finally, the pace of the questions slowed during the frustrational portion of the testing, and time to answer increased allowing more time for movement.

Based on these factors and the observation of the tapes, I have conjectured that two opposing factors were at work in the testing situation. The tension and frustration seemed to make the student generally more agitated and fidgety, but the greater concentration required of the harder questions tended to cause more motionless listening and thinking activities. Thus, little change in overall movement was observed.

Manipulation

This study found manipulations to be significantly related to frustration. Manipulation or self-manipulation has been used as a behavioral category in the studies of several investigators (Mehrabian, 1972; Eggen, 1972; Schelfin and Schelfin, 1972). However, none of these studies seemed to directly investigate the relationship of such manipulations to frustration. Morris (1977) considers manipulation such as head scratching to be a "displacement activity", i.e. one that allows some vent for suppressed affect. Schelfin (1972) has identified certain behaviors such as preening the hair as a quasi courtship behavior and wiping the finger across the upper lip as an attempt to hide deceit or disagreement. Schelfin's studies used more of the ethological observation method; however, his conclusions have not been statistically tested. Preening of the hair and the finger wipe would have been simply identified as manipulations in this study, but there may be some justification in expanding the category into sub-categories to gain more information.

Gaze

In the present study gaze was simply recorded as being directed towards the testing material or not directed towards the testing materials. Although not thoroughly analyzed, most of the gaze that was not directed at the testing materials was directed towards the examiner. This looking at the examiner occurred more frequently during the asking of the question and during the pause between the subject's answer and the next question.

Gaze toward the examiner was found to be significantly less frequent during the frustrational portion of the PIAT testing but did not significantly change for the PPVT testing. There are several factors that may help to explain these results. The effect of the difference in the pace of the PIAT and PPVT testing situations has been discussed above and probably helps to explain why change was noticed in one situation but not the other. The decrease in gaze noted for the PIAT testing may be because of the fact that a failing student may be a little embarrassed and tend to avoid eye contact; this decrease may also be at least partially due to the increase in eye contact with the test materials while the student is thinking about the question and trying to identify the correct answer.

Although gaze has been one of the most researched areas of non-verbal behavior, few if any of the studies have been directed at a task oriented situation such as the present study. In one study, however, Argyle and Graham (cited in Argyle, 1975) showed that when a simple prop was introduced in the experimental setting, almost all eye contact was directed towards the prop. In Argyle's study the prop was a map, and the conversation was about summer vacation plans. This effect must also be noted in the testing situation.

Many studies of gaze including those reported in Argyle's book Gaze and Mutual Gaze (Argyle and Cook, 1975) have focused on the effects of culture, status, sex, age, and other variables on gaze. Gaze has been identified as having a role as a regulator of verbal interaction in an unstructured setting.

In an individual testing setting, such as in the present research, where the verbal interaction is structured and the test materials serve as a prop, gaze probably does not have much importance in regulating the interaction and may be limited as a nonverbally communicative behavior.

Eyebrow Flash

In this study the term eyebrow flash was used to denote a raising of the eyebrows, because the act appeared to be identical with the non-verbal behavior commonly called a flash in the literature. The eyebrow flash usually refers in the literature to the gesture that denotes recognition of a friend (Argyle, 1975). Although no specific references were found that related the eyebrow flash or eyebrow raising to such a task oriented or frustrating situation, in this study the eyebrow flash was found to increase in the frustrational setting. The gesture seemed to communicate mock surprise right before or right after the answer was given. The implicit message seemed to be "I don't know" or "I'm guessing".

Eyebrow Lowering

Lowering of the eyebrows may be commonly associated with worry or sorrow, but there are few studies that identify this gesture specifically. In the studies of Ekman and Friesen (1975), eyebrow lowering was associated with the emotions of anger and sadness. However, Ekman's work was primarily a study of pure emotion and may be hard to relate to the more complex and transient emotions that the frustrational testing situation elicit. Eyebrow lowering increased significantly with frustration. The message communicated by the brow lowering seemed to be "I'm thinking" or "That's a hard question".

Smiles

The message commonly associated with smiling is most commonly a positive affect of liking or happiness. Perhaps because of the almost obvious messages associated with smiles less research has been done on the smile as a nonverbal cue. Eggen (1972) showed that smiling was associated with liking of the class and the teacher. Other studies have related smiles to an innate and cross-cultural expression of happiness.

In this study, the number of smiles decreased significantly in the frustrational portion of the PIAT testing but not in the PPVT. This result was similar to the decrease in Gaze that was noted in the PIAT but not the PPVT. As discussed above, both GAZE and SMIL were thought to be partially the result of the differences in the pace of the testing as discussed above. Further, smiles seemed to occur mostly when the student looked up at the examiner. If this relationship existed, then most of the discussion of the results for Gaze apply here. That is, failing students may avoid gaze and smiles partially because of embarrassment or anger, and gaze may be focused more on the testing materials as the students try to identify the correct answer.

Mouth Dropping

Mouth dropping as a behavior was not specifically noted in the literature review. Ekman and Friesen, however, have associated an open mouth with expression of surprise (Ekman and Friesen, 1975). Comedians and pantomimists use an open mouthed expression both for surprise or to act dumb or dumbfounded (Darwin, 1872/1965). The increase in mouth dropping observed in this study during frustration may be a communication of surprise or mock surprise at the hardness of the question.

Another hypothesis is that the behavior is more common in children and is repressed through social conditioning (Knapp, 1978, p. 65-76). The open mouth may be seen less in adults because of greater self

control and awareness of their facial expressions. During cognitive or emotional stress, however, these behaviors may be noted in adults.

Subvocalization

Subvocalization increased significantly during frustration for both the PIAT and PPVT. Subvocalization has been often noted in the reading behavior of children. In the case of this study, in the PIAT testing the students may have been in some cases rereading the questions, because the question was usually printed on the stimulus card. This does not explain the significant relationship between subvocalization and frustrational testing for the PPVT test. In the PPVT, the stimulus cards were totally free of writing. In the case of the PPVT, careful observation of the videotapes revealed that the subjects were usually repeating the question, i.e. the stimulus word. Some educators believe subvocalization may help the student attend to a difficult task and block out distractive stimuli. This could be the reason for the increased subvocalization observed during frustration. In a few cases the subjects were subvocalizing some negative comment such as "oh no" or "that's not right".

Shoulder Shrug and Head Cocking

Among the "other behaviors" noted in the study, there were several that appeared to be quite communicative even though they were less frequently noted. Two of the most striking of these behaviors were the shoulder shrug and the head cocking behaviors. Darwin noted that, "When a man wishes to show that he cannot do something or prevent something being done, he often raises with a quick movement both shoulders.... The head is often thrown a little to one side..." (Darwin, 1872/1965, p. 263-264). Both of the behaviors seemed to clearly communicate "I don't know", and in some cases almost an "I don't care". At this point in the testing the probability was high that the student was guessing. However, perhaps because of the almost obvious nature

of the nonverbal behavior, no research studies were found that dealt with this specific behavior.

Lip Licking

It may be significant that there was no difference noted in the amount of lip licking during the frustrational testing. Lip licking is assumed to be associated with dryness of the mouth due to a reduction of saliva flow caused by stress or tension (Morris, 1977, p. 166-167). Lip licking was frequently noted in both pre-frustrational and frustrational behavior of some students, and the lack of change in the behavior is interesting. One conjecture is that lip licking, associated with some dryness of the mouth, was more a function of test anxiety than a specific affective response to frustration. This hypothesis warrants further investigation.

Swallowing, like lip licking, might be a result of dryness of the mouth and a function of anxiety more than specific frustration. Unfortunately, swallowing could not be easily and reliably observed in all subjects and therefore was not systematically studied here.

Morris (1977) also noted that besides lip licking, the tongue may also protrude slightly during periods of concentration. He called this nonverbal behavior the concentration tongue. He claimed that nursery school children use this behavior as a sign that they do not want to be disturbed, and that the behavior may be related to another rejection signal of sticking out the tongue.

Recommendations for Further Study

Several suggestions for further research arise out of the discussion of the results of this study. These recommendations are presented below. First, the study should be replicated, perhaps with some modifications. Second, there are some similar studies that would extend the findings of this study. Third, some suggestions are made for instrument modification, especially where the observation is to be done in

more dissimilar circumstances. Finally, there are some suggested research studies that are related to the findings of this study but that require quite different experimental designs.

Replication Studies

Replication of any reasonable study needs no specific justification other than the desire to confirm the findings independently. However, replication studies usually incorporate modifications of design and method that serve not only to verify findings but to improve the study.

There are several modifications that might be incorporated. First, because the instrument has been developed, it would be advisable to spend more energy on an increased sample size. This might allow statistical analysis of the less frequently observed behaviors. An increase in sample size might also make it possible to observe the hypothesized clustering of the frustrational behavior.

Another modification would be to analyze the data in terms of the timing of nonverbal behaviors. Which nonverbal behaviors often occur concurrently? This type of analysis would investigate the "micro-momentary facial expression" phenomena. Further, when during the testing do the nonverbal behaviors occur? This would require investigation of which behaviors occur during questioning, during the time to answer, during or near the answer, and during the pause to the next question.

Similar Studies

Other individualized testing situations that are similar to the PPVT and the PIAT might be used for a replication study. Although there was much more similarity than difference in the nonverbal behaviors seen in the PIAT and the PPVT testing, the investigations of other testing situations would extend arguments for the external validity of this study.

One of the most common individualized testing situations has been the individualized testing of reading, especially the IRI that was designed by Betts. Betts' IRI (Informal Reading Inventory) provided some of the initial justification for this study. Thus the IRI testing would be a logical candidate for a study of nonverbal behaviors.

A different type of study might use some non-testing task to provide the frustrational situation. The observational instrument developed in this study might require modification but could easily be adapted to observation of other classroom tasks. One key element in the selection of frustrational tasks to study is the control of the task difficulty or frustrational stimulus. One type of classroom task that seems quite amenable to this type of study would be the use of individualized instructional materials. Such materials, often in the form of self instructional packets are usually designed so that the student can achieve success without aid from the teacher. The difficulty is controlled by design and could be made harder or easier for an experimental situation.

Videotapes of lessons or slide-tape presentations could be used as in the studies of Allen and Atkinson (1977). Certainly the use of computer assisted instruction, computer simulations, or computer games could also be considered for study of the nonverbal behaviors and frustration. Pareek (1964) has listed other task situations that have been used in the study of frustration. They include peg board games, problems given with too little time to solve them, falsified test scores, and several difficult motor coordination tasks. This research is certainly needed since it could be important for the teacher to be able to identify students who are experiencing frustration.

Instrument Modifications

Further research may require modification of the instrument. When the situation to be observed is different than the individualized testing used in this study, it would be advisable to do some pilot videotaping in order to be sure that all commonly seen behaviors are

included. As research progresses it might also be wise to ignore certain nonverbal behaviors that are not found to be related to the frustrational stimulus. Depending on the researcher and the needs of the study, the nonverbal observation record sheet might include more or fewer behaviors. For studies such as the study of the timing of nonverbal behaviors previously suggested, it would be interesting to consider reducing the five second intervals to two or one second intervals. Such a study would require more sophisticated videotape equipment. Also, where the sample size is large, it would probably be important to consider inclusion of some of the behaviors observed infrequently.

Other Studies

The results of this research suggest some interesting hypotheses for further investigation. Two such studies are suggested in the following paragraphs.

The first investigation would be the investigation of the connection between specific nonverbal behaviors elicited by the students in the frustrational situation and the ability of teachers to decode these behaviors accurately. Correlation of specific nonverbal behaviors with the judgement accuracy scores of a panel of teachers could provide some insight. We might hypothesize that certain nonverbal behaviors such as mouth twitching or a shoulder shrug more clearly communicate the students' frustration. A related hypothesis would be that students that are higher in total overall expression or frequency of nonverbal behaviors would be easier for an observer to judge accurately. And finally, the hypothesis that certain personality types more openly express their frustration could be tested if the study included personality testing for the student subjects.

The second type of related research suggested is a study of the relationship between the nonverbal behaviors and test anxiety. As was noted above, lip licking was a reasonably common behavior but was not found to be related to the frustrational stimulus specifically. Test

anxiety has been researched extensively. The largest body of research is included in the book Anxiety in Elementary School Children (Sarason, Davidson, Lighthall, Waite, and Ruebush, 1960). Most of this research uses a student questionnaire that asks questions such as, "Do you worry a lot before you take a test"? Much of the data that has been collected on classroom anxiety and test anxiety compares anxiety levels of different populations such as minority groups, age groups and so on.

In psychology, anxiety is defined quite differently than frustration, and although there may be some connections, the interesting results for lip licking suggest that the differences between anxiety and frustration may be observable in terms of nonverbal behaviors. Certainly research here could enhance our understanding of frustration, anxiety, emotional communications, and nonverbal behaviors.

BIBLIOGRAPHY

- Allen, V. L. & Atkinson, R. S. "Encoding of Nonverbal Behavior by High and Low Achieving Children." *Journal of Ed Psych*, 1978, 70 (3), 298-305.
- Allen, V. L. & Atkinson, R. S. The Generality-Specificity and Decoding with Spontaneous and Deliberate Nonverbal Behavior (Tech Report No. 443). Madison Wisconsin: Research and Development Center for Individualized Schooling, University of Wisconsin, 1977. (ERIC Document Reproduction Service No. ED 148 189).
- Allen, V. L. & Budeau, L. B. Ability of Parents to Decode Nonverbal of Their Own and Other Children (Tech Report No. 444). Madison Wisconsin: Research and Development Center for Individualized Schooling, University of Wisconsin, 1977. (ERIC Document Reproduction Service No. ED 153 186).
- Allen, V. L. and Feldman, R. S. Decoding of Children's Nonverbal Responses (Tech Report No. 365). Madison Wisconsin: Research and Development Center for Individualized Schooling, University of Wisconsin, 1975. (ERIC Document Reproduction Service No. ED 117 044).
- Anderberg, M. R. Cluster Analysis for Applications. New York: Academic Press, 1973.
- Angelino, H. R. The Validity of the Rosenzweig Picture - Frustration Study (Children's Form). Unpublished doctoral dissertation, University of Nebraska, 1951.
- Argyle, M. Bodily Communication. New York: International Universities Press, 1975.
- Argyle, M. & Cook, M. Gaze and Mutual Gaze. Cambridge: Cambridge University Press, 1976.
- Barker, L. L. & Collins, N. B. Nonverbal and kinesic Research, In P. Emmert and W. D. Brooks (Eds.), Methods of Research in Communications. Boston: Houghton-Mifflin, 1970.
- Barker, R., Dembo, T., & Lewin, K. Frustration and Aggression: An Experiment with Young Children. University of Iowa Study of Child Welfare, 18 (1), 1941.
- Bettelheim, Bruno. Love is Not Enough. New York: MacMillian Co., 1965.

- Betts, E. A. Foundations of Reading Instruction. New York: American Book Co., 1957.
- Birdwhistell, Ray L. Kinesics and Context. Pennsylvania: University of Pennsylvania, 1970.
- Boehm, A. E. The Classroom Observer. New York: Teachers College Press, 1977.
- Boucher, J. D. & Ekeman, P. Facial Areas and Emotional Information. Journal of Communications, Spring 1975, 25 (2), 21-29.
- Breed, G. and Colaivta, V. "Looking, Blinking and Sitting: Nonverbal Dynamics in the Classroom." Journal of Communications, 1947, 24 (2), 75-81.
- Bryan, T. H. and Bryan J. H. Understanding Learning Disabilities. New York: Alfred Publ. Co., 1975.
- Buck, R. "Nonverbal Communication of Affect in Children." Journal of Personality & Social Psychology, 1975, 31, 644-653.
- Buck, R., Miller, R. E. & Caul, W. F. "Sex, Personality, and Physiological Variables in the Communication of Affect via Facial Expression." Journal of Personality and Social Psychology, 1974, 30 (4), 587-596.
- Buehler, R. E. & Richmond, J. F. "Interpersonal Communication Analysis: A Research Method." Journal of Communication, 1963, 8 (3), 146-155.
- Cronbach, L. J. Essentials of Psychological Testing (3rd ed.). New York: Harper and Row, 1970.
- Dittman, A. T. "Developmental Factors in Conversational Behavior." The Journal of Communication, December 1972, 22, 404-423.
- Dittman, A. T. Interpersonal Messages of Emotion. New York: Springer Publ. Co., 1972.
- Dittman, A. T. "The Relationship between Body Movements and Moods in Interviews." Journal of Consulting Psychology, 1962, 26 (5), 480.
- Darwin, C. The Expression of Emotions in Man and Animals. Chicago: University of Chicago Press, 1965. (Originally published 1872).
- Davitz, J. R. (Ed.). The Communications of Emotional Meaning. New York: McGraw-Hill, 1964.

- Davitz, J. R. & Davitz, L. J. "Correlates of Accuracy in the Communication of Feelings." The Journal of Communications, 1959, 9 (3), 110-117.
- Davitz, J. R. and Davitz, L. J. "The Communication of Feelings by Content Free Speech." Journal of Communication, 1959, 9 (1), 6-13.
- Dibner, A. S. "Cue-counting: A Measure of Anxiety in Interviews." Journal of Consulting Psychology, 1956, 20 (6), 475-478.
- Dixon, W. J. and Brown, M. B. BMDP-79: Biomedical Computer Programs P-Series. Berkeley: University of California Press, 1979.
- Dollard, J., Doob, L. W., Miller, N. E., Mowrer, O. H., & Sears, R. R. Frustration and Aggression. New Haven: Yale University Press, 1939.
- Dominick, B. "Nonverbal Communication." Journal of Communication, 1963, 8 (3), 146-155.
- Eggen, P. D. A Comparison of Student Affect and Kinesic Behavior. Unpublished doctoral dissertation, Oregon State University, 1972.
- Ekman, P. "Movements with Precise Meanings." Journal of Communication, Summer 1976, 26 (3), 14-26.
- Ekman, P. & Friesen, W. V. Unmasking the Face. New Jersey: Prentice-Hall, 1975.
- Ekman, P., Friesen, W. F., & Ellsworth, P. Emotion in the Human Face: Guidelines for Research and an Integration of Findings. New York: Pergamon Press Inc., 1972.
- Ekman, P., Friesen, W. V. "Hand Movements." The Journal of Communication, Dec. 1972, 22, 353-374.
- Ekman, P., Friesen, W. V. and Tompkins, S. S. "Facial Affect Scoring Techniques: A First Validity Study." Semiotica, 1971, 3, 37-58.
- Ekwall, E. E. "Informal Reading Inventories: The Instructional Level." The Reading Teacher, 1974b, 29, 662-665.
- Ekwall, E. E. "Should Repetitions be Counted as Errors?" The Reading Teacher, 1974a, 27, 365-367.
- Ellsworth, P. C. & Ludwig, L. M. "Visual Behavior in Social Interaction." Journal of Communication, Dec. 1972, 22, 375-403.
- Fast, J. Body Language. New York: Pocket Books, 1971.

- Flanders, N. A. Interaction Analysis in the Classroom: A Manual for Observers (Rev. ed.). Ann Arbor, Mich: University of Michigan, 1964.
- Frick, T. & Semmel, M. I. "Observer Agreement and Reliabilities of Classroom Observation. Review of Educational Research, W 1978, 48 (1), 157-185.
- Frijda, N. H. The relation between emotion and expression. In M. Von Cranach and I. Vine (eds.), Social Communication and Movement. New York: Academic Press, 1973.
- Galloway, C. M. Teaching is Communicating: Nonverbal Language in the Classroom. Washington, D. C.: National Education Association, 1970.
- Galloway, C. M. An Analysis of Theories and Research in Nonverbal Communication. Washington, D. C.: National Center for Educational Communications, 1972. (ERIC Document Reproduction Service No. ED 059 988).
- Gronlund, Norman E. Measurement and Evaluation in Teaching (3rd Ed.). New York: MacMillian Publ. Co., 1976.
- Gilliland, H. A Practical Guide to Remedial Reading. Columbus, Ohio: Charles E. Merrill, 1974.
- Haggard, E. A. & Isaacs, K. S. Micromomentary facial expressions as indicators of ego mechanisms in psychotherapy. In L. A. Gottschalk and A. H. Auerbach (Eds.), Methods of Research in Psychotherapy. New York: Appleton-Century-Crofts, 1966.
- Hall, E. T. The Silent Language. New York: Doubleday, 1959.
- Hopson, J. L. "Scent and Human Behavior: Olifaction or Fiction?" Science News, April 1979, 115 (17), 282-283.
- Harper, R. G., Wiens, A. N. & Matazazzo, J. D. Nonverbal Communication: The State of the Art. New York: John Wiley & Sons, 1978.
- Harris, Richard J. A Primer of Multivariate Statistics. New York: Academic Press, 1975.
- Harrison, R. P. The human face. In T. M. Steinfatt (Ed.), Readings in Human Communication. Indianapolis: Bobbs-Merrill, 1977.
- Harrison, Randall R. P. Beyond Words: An Introduction to Nonverbal Communication. Englewood Cliffs, N. J.: Prentice Hall, 1974.
- Harrison, R. P., and Knapp M. L. "Toward an Understanding of Nonverbal Communication Systems." The Journal of Communication, Dec. 1972, 22, 339-352.

- Hartigan, J. Clustering Algorithms. New York: Wiley, 1975.
- James, W. "A Study of the Expression of Bodily Posture." Journal of General Psychology, 1932, 7, 405-437.
- Jecker, J. D., Maccoby, N. & Brietrose, H. S. "Improving Accuracy in Interpreting Nonverbal Cues of Comprehension." Psychology in the Schools, 1965, 2, 239-244.
- Jecker, J. D., Maccoby, N., Brietrose, H. S. & Rose, E. D. "Teacher Accuracy in Assessing Cognitive Feedback from Students." Journal of Applied Psychology, 1964, 48 (6), 393-397.
- Johnson, M. S. & Kress, R. A. Informal Reading Inventories. Newark, Delaware: International Reading Assn., 1965.
- Knapp, M. L. Nonverbal Communication in Human Interaction (2nd ed.). New York: Holt, Rinehart, and Winston, 1978.
- Knapp, M. L. The role of nonverbal communications in the classroom. In J. M. Civikly (Ed.), Messages: A Reader in Human Communication (2nd ed.). New York: Random House, 1977.
- Knapp, P. H. (Ed.). The Expression of the Emotions in Man. New York: International Universities Press, 1963.
- Krout, M. H. "An Experimental Attempt to Produce Unconscious Manual Symbolic Movements." Journal of General Psychology, 1954, 51, 93-120.
- Krout, M. H. "An Experimental Attempt to Determine the Significance of Unconscious Manual Symbolic Movements." Journal of General Psychology, 1954, 51, 121-152.
- Lawick-Goodall, J. V. In the Shadow of Man. Boston: Houghton Mifflin, 1971.
- Lawick-Goodall, J. V. The behavior of free-living chimpanzees in The Gombe Stream Reserve. Animal Behavior Monographs, 1968, 1 (3), 161-311.
- Light, R. J. Issues in the analysis of qualitative data. In R. M. W. Travers (Ed.), Second Handbook of Research on Teaching. Chicago: Rand McNally, 1973.
- Lowen, A. The Language of the Body. New York: Collier Books, 1958.
- Mahl, G. J. "Disturbances and Silences in the Patients' Speech in Psychotherapy." Journal of Abnormal Social Psychology, 1959, 53, 1-15.

- Maier, N. R. F. & Ellen, P. "Integrative Value of Concepts in Frustration Theory." Journal of Consulting Psychology, 1959, 23, 195-206.
- Maier, N. R. F. Frustration: The Study of Behavior Without a Goal.
- Matarazzo, J. D. & Wiens, A. N. The Interview: Research on its Anatomy and Structure. Chicago: Aldine-Atherton, 1972.
- Medley, D. M. & Mitzel, H. E. Measuring classroom behavior by systematic observation. In N. L. Gage (Ed.) Handbook of Research on Teaching. New York: Rand McNally, 1963.
- Mehrabian, A. Nonverbal Communication. Chicago: Aldine-Atherton, 1972.
- Mehrabian, A. A. "A Semantic Space for Nonverbal Behavior." Journal of Consulting and Clinical Psychology, 1970, 35 (2), 248-257.
- Miller, N. E. "The Frustration-Aggression Hypothesis." The Psychological Review, 1941, 48, 337-342.
- Morris, D. Man Watching: A Field Guide to Human Behavior. New York: Harry N. Abrams, 1977.
- Morrison, D. F. Multivariate Statistical Methods. New York: McGraw Hill, 1976.
- Pareek, U. N. Developmental Patterns in Reactions to Frustration. New York: Asia Publ. Co., 1964.
- Peterson, C. L., Danner, F. W., and Flavell, J. H. Developmental Changes in Children's Response to Three Indications of Communicative Failure. Child Development, 1972, 43, 1463-1468.
- Phillips, B. School Stress and Anxiety. New York: Human Sciences Press, 1978.
- Raskin, A. "Observable Signs of Anxiety or Distress During Psychotherapy." Journal of Consulting Psychology, 1962, 26 (4), 389.
- Raymond, A. D. The Acquisition of Nonverbal Behaviors by Preservice Science Teachers and Their Application During Student Teaching. Unpublished doctoral dissertation, Oregon State University, 1972.
- Reich, W. Character Analysis (3rd ed.). New York: Simon and Schuster, 1972.
- Robbins, O., Shannon, D., & Wiener, M. "Social Patterns of Turn Taking: Nonverbal Regulators." Journal of Communication, Summer 1978, 28 (3), 38-46.

- Rosenzweig, S. Psychodiagnosis: An Introduction to Tests in the Clinical Practice of Psychodynamics. New York, Grune and Stratton, 1949.
- Rosenzweig, S. "The Picture-Association Method and Its Application in a Study of Reactions to Frustration." Journal of Personality, 1945, 14, 3-23.
- Rosenzweig, S. "An Experimental Study of 'Repression' with Special Reference to Need-Persistive and Ego-Defensive Reactions to Frustration." Journal of Experimental Psychology, 1943, 32, 64-74.
- Rosenzweig, S., Fleming, E. E., and Rosenzweig, L. "The Children's Form of the Rosenzweig Picture-Frustration Study." Journal of Psychology, 1954, 24, 171.
- Rugel, R. P. "Arousal and Levels of Reading Difficulty." Reading Teacher, 1971, 24, 458-460.
- Ruesch, J. and Kees, W. Nonverbal Communications. Berkeley: Univ. of California Press, 1956.
- Saltz, E. The Cognitive Basis of Human Learning. Homewood, Ill., Dorsey Books, 1971.
- Sandoval, J. The Measurement of the Hyperactive Syndrome in Children. Review of Educational Research, Winter 1977, 47 (1), 293-318.
- Sarason, S. B., Davidson, K. S., Lighthall, F. F., Waite, R. R. & Ruebush, B. K. Anxiety in Elementary School Children. New York: John Wiley, 1960.
- Saslow, G., Matrazzo, J. D., and Guze, S. B. "The Stability of Interaction Chronograph Patterns in Psychiatric Interviews." Journal of Consulting Psychology, 1955, 19 (6), 417-430.
- Schefflen, A. E. and Schefflen, A. Body Language and Social Order. Englewood Cliffs, New Jersey: Prentice-Hall, 1972.
- Scott, W. Attitude measurement. In G. Lindley & E. Aronson (Eds.), The Handbook of Social Psychology (Vol. 2). Reading, Mass: Addison-Wesley, 1968.
- Scott, W. A. "Reliability of Content Analysis: The Case of Nominal Scale Coding." Public Opinion Quarterly. 1955, 19, 321-325.
- Shannon, C. E. and Weaver, W. The Mathematical Theory of Communication. Urbana: University of Ill. Press, 1949.

- Sharipo, J. G. "Responsivity to Facial and Linguistic Cues." Journal of Communication, Mar. 1968, 18, 11-17.
- Sheflen, A. E. & Sheflen, A. Body Language and Social Order. Englewood Cliffs, N. J.: Prentice-Hall Inc., 1972.
- Simon, A. & Boyer, E. G. Mirrors for Behavior: An Anthology of Observation Instruments (4 volumes). Phil., Pa.: Research for Better Schools Inc., 1965-1970.
- Strongman, K. T. The Psychology of Emotion (2nd ed.). New York: John Wiley and Sons, 1978.
- Toft, R. "The Ability to Judge People." Psych. Bulletin, 1955, 52 (1), 1-23.
- Weitz, S. (Ed.). Nonverbal Communication: Readings with Commentary. New York: Oxford University Press, 1974.
- Werry, John S. The diagnosis, etiology, and treatment of hyperactivity. In J. Hellmich (Ed.), Learning Disorders (Vol. 3). Seattle, Wash: Special Child Publication, 1968.
- Wiermann, J. M. & Knapp, M. L. "Turn-Taking in Conversations." Journal of Communication, Spring 1975, 25 (2), 75-92.
- Williams, F. & Tolch, J. "Communication by Facial Expression." Journal of Communication, 1965, 15 (1), 17-27.
- Yates, A. J. Frustration and Conflict. New York, Methuen, 1962.

APPENDICES

APPENDIX I
INTEROBSERVER AGREEMENT

The interobserver agreement was calculated using Scott's Pi (Scott, 1955; Frick and Semmel, 1978):

$$\pi = \frac{p_o - p_e}{1.0 - p_e}$$

where: p_o = observed agreement = $\frac{1}{N} \sum_i n_{ii}$

n_{ii} = number of cases where two observers agree on the behavior

p_e = expected agreement due to chance $\sum [p(i)]^2$

$p(i)$ = the average proportion of tallies made by all observers for the i th category.

A sample calculation of interobserver agreement is provided below. This example uses data from the observation of eyebrow lowering (EBLO). For the five tapes used for the interobserver agreement calculations, there were five prefrustrational segments and five frustrational segments, and thus there were a total of 10 one-minute segments of behavior observed. In this 10 minutes of observation there were a total of 120 five second intervals. In each 120 five second interval each observer recorded a one (1) or a zero (0) depending on whether the behavior was observed. Table I-1 summarizes the comparison of the observational record for the two observers for EBLO.

Table I-1. Interobserver agreement data for EBLO.

Observer 2	Observer 1		Total
	Behavior Occurred (1)	Behavior Did Not Occur (0)	
Behavior Occurred (1)	6	2	8
Behavior Did Not Occur (0)	2	110	112
Total	8	112	120

Appendix I (continued)

From the data in the above table, interobserver agreement for EBLO was calculated as follows:

$$p(1) = \frac{8 + 8}{240} = 0.0667$$

$$p(0) = \frac{112 + 112}{240} = 0.9333$$

$$p_e = [p(1)]^2 + [p(0)]^2 = 0.8756$$

$$p_o = \frac{6 + 110}{120} = 0.9667$$

$$\pi = \frac{p_o - p_e}{1.0 - p_e} = 0.7323$$

Table I-2 lists the results of the interobserver agreement calculations for other behaviors.

Appendix I (continued)

Table I-2. Interobserver agreement.

Variable	Agreement
PSFT	1.00*
LEAN	1.00*
HSHK	0.80
MANP	1.00
HTH	1.00
GAZE	1.00
EBFL	0.74
EBLO	0.73
MTWT	0.81
SMIL	0.85
LICK	0.82
MDRP	-0.01**
FLPA	1.00
SUBV	0.54
ALL BEHAVIORS	0.88

*These two behaviors were not observed in the tapes selected for the interobserver agreement measurement. 1.00 is indicated since there was agreement that these behaviors were not observed.

**This behavior was observed only one time in the videotapes selected for the interobserver agreement measurement. There was a disagreement as to the correct placement of the behavior in time but the total behaviors for both observers was the same.

See the other interobserver agreement measurement notes in Appendix II.

APPENDIX II

NOTES ON THE INTEROBSERVER AGREEMENT CALCULATIONS

At the end of the interobserver agreement study the two observers compared behavior records and discussed the disagreements in the two records. In many cases the videotape record was then viewed again to determine who was wrong or what the source of the disagreement was. The following notes primarily relate the conclusions of those discussions.

1. Several researchers suggest that interobserver agreement should be calculated for the case where two or more observers view specially edited videotapes that include examples of each behavior (Frick and Semmel, 1978). Often these tapes are prepared so that the tapes include only pure and nonambiguous examples of behavior. In other cases the tapes are developed to provide a check of observer training in the use of an instrument. In these cases the interobserver agreement can be calculated between the trainee's record and a record prepared by the judgement of experts and thus provides a check on the ability of the observer.

The preparation of a special videotape was rejected for this study in favor of the use of randomly chosen tapes because it was thought to be the best check of the real reliability of the instrument and because the time and effort involved did not seem to be justified by any particular improvement in accuracy of the measurement of interobserver agreement.

One drawback in the rejection of this special tape method was that for certain behavior listed in the observational instrument the frequency of occurrence was quite low for the random sample of videotapes that was used. In those cases, just one or two disagreements in the few behaviors viewed resulted in very low calculated values of Scott's Pi.

It is therefore felt that the most reliable or least uncertain of the Scott's Pi measurements is found where the frequency of occurrence is higher. Note that the frequency of behavior can be roughly taken

Appendix II (continued)

from the behavioral record and is noted on the listing of the results of the Scott's Pi calculations in Appendix I.

2. Some of the disagreements in the behavioral record occurred only in which of the five second intervals the behavior was seen to occur in. Because the interobserver agreement was calculated on the basis of looking at the detail of the record, that is placement in each five second interval, some of the cases of noted disagreement resulted in the same total score for the subject behavior. This method of calculation although statistically much easier was therefore somewhat conservative in estimating interobserver agreement.

It was found that the disagreements in determination of the correct five second frame in which to score the behavior could be greatly reduced or eliminated with a different video system. This discovery was made because the videotape record was made with two different videotape systems - port-a-pack system and a studio camera and cassette recorder. The studio system seemed to have better overall video quality; tapes for the interobserver agreement study were all taken from the cassette videotapes. However, it was found that because of the more automatic controls on the cassette videotape and the lack of a slow motion feature, it was sometimes difficult to determine exactly in which time period a particular behavior occurred.

On the other hand, the port-a-pack is a simple reel to reel videotape machine; the operator can advance the tape by hand, allowing frame by frame viewing of the videotape record. This frame by frame viewing allowed the operator to much more easily determine in which five-second interval a behavior occurred. In retrospect, it was felt that the smaller port-a-pack machine could have eliminated some of the disagreements seen in the interobserver agreement study and therefore the reliability of the instrument could be increased above the level indicated by the calculated values of Scott's Pi through the use of the port-a-pack system.

APPENDIX III POPULATION DESCRIPTION

The population was taken from the Lebanon Middle School in Lebanon, Oregon. Students in five eighth grade classes were asked to volunteer for the study. The students in these classes were told very little about the research except that it was aimed at the improvement of testing materials and techniques. Each of the students in these classes was given a letter of parental consent to take home. Of the approximately 125 students who were asked to participate in the study, only 40 returned the completed consent form indicating approval for the testing.

Lebanon, Oregon is a small town in the Willamette Valley area with a population of approximately 9,000 people. The economy is primarily based on lumber and agriculture. The local school district supports one middle school for grades five through eight and one high school for grades nine through twelve.

In the eighth grade math classes, the students are placed into low, average, and high groups. In this study it appeared that (1) more of the higher ability students volunteered for the testing and (2) more girls than boys volunteered. The breakdown of the student population is given in Table III-1.

Table III-1. Breakdown of study population by sex and ability.

Ability	Sex		Total
	Male	Female	
High	9	10	19
Average	2	5	7
Low	3	9	12
Total	14	24	38

APPENDIX IV
THE JECKER-MACCOBY-BRIETROSE CHECKLIST
(Jecker et al., 1965)

- | | |
|---|---|
| <p>1-Clip Number</p> <p>2-Amount of time looked at source</p> <p> 1-Little or none</p> <p> 2-Somewhat</p> <p> 3-A great deal</p> <p> 4-Virtually all the time</p> <p>3-Number of times looked away from source</p> <p> 1-Many times (more than 4)</p> <p> 2-Some (3 or 4)</p> <p> 3-Little (1 or 2)</p> <p> 4-None</p> <p>4-Speed of eye movements away from and returning to source</p> <p> 1-Slow</p> <p> 2-Fast</p> <p> 3-None away from source</p> <p>5-Amount of blinking</p> <p> 1-More than normal</p> <p> 2-Normal</p> <p> 3-Less than normal</p> <p>6-Grouping of blinking</p> <p> 1-Spaced</p> <p> 2-Clustered</p> <p> 3-Virtually no blinking</p> <p>7-Duration of lowering eyebrows</p> <p> 1-A lot</p> <p> 2-A little</p> <p> 3-None</p> <p>8-Strength of lowering eyebrows</p> <p> 1-Strong</p> <p> 2-Weak</p> <p> 3-No eyebrow lowering</p> <p>9-Duration of raising eyebrows</p> <p> 1-A lot</p> <p> 2-A little</p> <p> 3-None</p> <p>10-Strength of raising eyebrows</p> <p> 1-Strong</p> <p> 2-Weak</p> <p> 3-No eyebrow raising</p> | <p>11-Movement of hands on face</p> <p> 1-A lot</p> <p> 2-Some</p> <p> 3-A little</p> <p> 4-None</p> <p>12-Frequency of general body movement</p> <p> 1-A lot</p> <p> 2-Some</p> <p> 3-A little or none</p> <p>13-Extent of general body movement</p> <p> 1-A lot</p> <p> 2-Some</p> <p> 3-A little or none</p> <p>14-Amount of mouth movement</p> <p> 1-A lot</p> <p> 2-Some</p> <p> 3-A little or none</p> <p>15-Chewing gum or candy</p> <p> 1-Yes</p> <p> 2-No</p> |
|---|---|
- 16-My guess of student's answer

 1-Wrong

 2-Right

17-Student's actual answer

 1-Wrong

 2-Right

APPENDIX V

PARENTAL CONSENT LETTER

Parental Approval for Participation in Educational Research

TO: Prospective participants and parents

March 6, 1979

During the months of March and April, some testing will be conducted at Lebanon Middle School by a researcher from Oregon State University. The testing is part of research aimed at better understanding of tests and testing procedures.

The testing will consist of two or three brief tests given individually to each participating student. The total testing will last about 15 minutes or less and will be recorded on videotape. The videotapes are necessary to allow careful examination of the testing procedure and results. The test results will be strictly confidential and the tapes will be erased after their analysis. The test results will not be used for any grading or evaluation and will not be placed in the students' records.

Participation in the testing is strictly voluntary and requires parental permission.

_____ Permission is given to conduct the testing described above.

_____ Permission is denied.

Parent's Signature _____ Date _____

Student's Name _____

Any questions about this research program should be addressed to:

Bob Roberts or Dr. Thomas P. Evans, Science Education Department,
Oregon State University, Corvallis, Oregon 97331
Phone (503) 754-4031.

APPENDIX VI

THE CLOCK-TIMER

As part of the pilot study, a timer was devised to facilitate in the recording of the nonverbal behaviors. The specific purpose of the timer was to provide an easy means of dividing the one minute of video record of behavior into five second intervals and to help identify which of the five second intervals was being viewed at any particular point in the observation process.

To accomplish this the timer provided an audio signal, consisting of clicks every one second and a short tone every five seconds, and a visual display that counted the five second intervals, so that each one could be identified by the number displayed on the video portion of the videotape record. Figure VI-1 shows a simple diagram of the audio and visual output of the timer.

The digital display used a 0.6 inch LED (light emitting diode) display that was quite clearly visible on the video record. In addition, the visual display included two LED flashers to mark the second and five second intervals. The audio tone marking the five second intervals was about 28 milliseconds in duration and consisted of a 3.7 KHZ audio signal. A block diagram of the circuit used is included in Figure VI-2, and a listing of the major components is given in Figure VI-3.

During the videotaping of the main study, the timer was placed on the table in front of the student and in the bottom of the video field of view. The audio output was mixed with the audio input from the microphone and fed into the videotape recorder.

Appendix VI (continued)

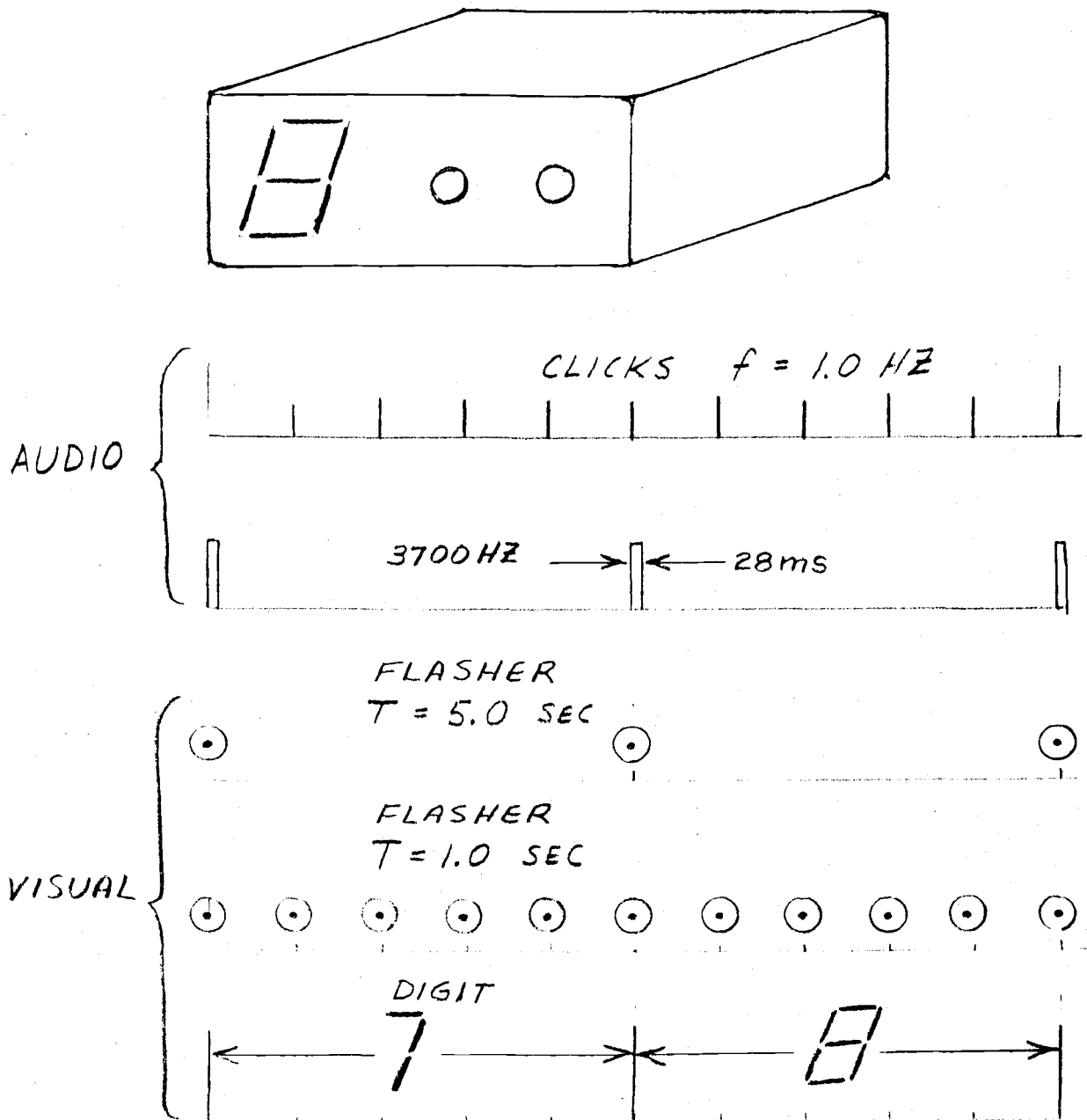


Figure VI-1. Diagram of timer output - audio and visual.

Appendix VI (continued)

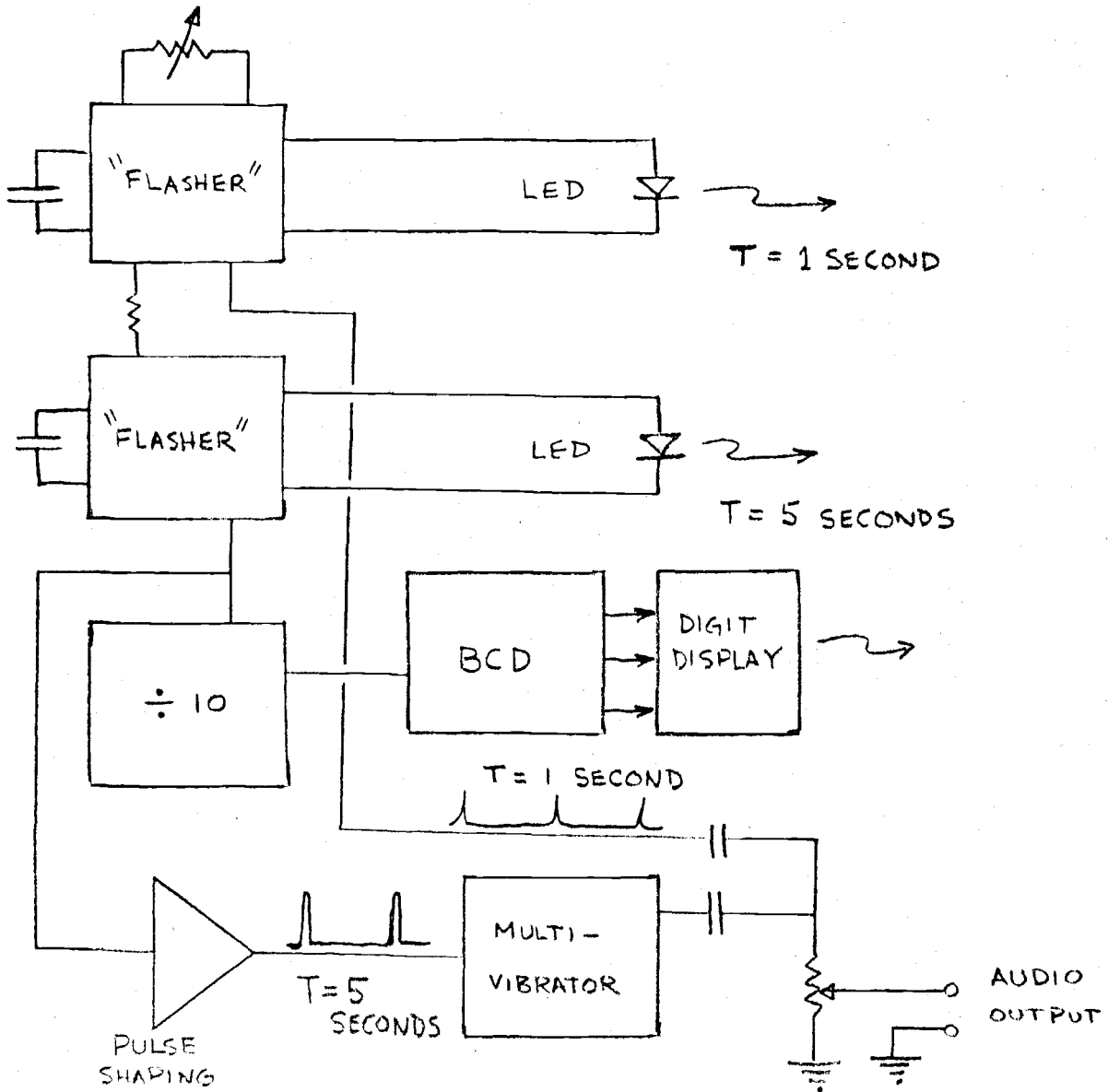


Figure VI-2. Block diagram of timer circuit.

Appendix VI (continued)

<u>Generic Number</u>		<u>Radio Shack Part Number</u>
7805	5VDC Regulator	276-1770
	0.6 inch LED display	276-056
7447	BCD to 7 segment display driver	276-1805
3909	LED flasher oscillator	276-1705
	Red LED type TLR107 fresnel lens	276-033
7490	TTL decade counter	276-1808

Figure VI-3. Major components parts list.

APPENDIX VII

THE OBSERVATIONAL INSTRUMENT

The observational instrument consists of a listing of behaviors to be observed and tabulated and the instrument record sheet. The listing of behaviors includes behavior descriptions and definitions to be used in recording and tabulation of the behaviors. In addition, this behavior listing includes instructions on the use of the record sheet. The observation instrument record sheet is given below as Figure VII-1.

The list of behaviors and their definitions is provided below as well as a listing of "other behaviors". These other behaviors were observed during the main study and recorded under the "comments and other observations" section of the record sheet. However, none of these other behaviors was observed with sufficient frequency to be included in the statistical analysis of the data from the main study.

Behavior Definitions and Criteria

In the listing of behaviors given below, the circled reference numbers are intended to help identify the correct row of the record sheet to which the instructions or definitions apply.

1 Sequence - In the row marked sequence, the numbers from the timer are recorded. See the description of the timer in Appendix VI. These numbers help to identify which of the five-second intervals is being viewed on the videotape recording.

2 & 3 Examiner and Subject - The verbal sequence of the testing is recorded with a few symbols in these two rows of the record sheet. The purpose of this verbal record is to help facilitate the recording of the behaviors in the correct five-second interval and to help identify coincidence of nonverbal behaviors with events in the verbal dialogue.

Symbols used in these two rows are listed and defined below.

Appendix VII (continued)

Recorded under Examiner:

Q - indicates the examiner asking a question from the test

C - indicates a comment by the examiner such as "good".

Recorded under Student:

A - indicates the subject responding to the previous question. This may include answers such as "I don't know".

C - indicates that the subject is making a comment or asking a question. For example, the subject may ask for clarification on part of the question, e.g. "(Did you say)...five times five?"

M - indicates a filled pause (See filled pause below). Basically a filled pause consists of unintelligible filler such as, 'aaah...' or "ummm...".

During the recording of the behaviors, clarification of the above symbols may be important. In most cases, a record of the verbatim comments made by both the examiner and subject will be included in the comments and observations section of the record sheet. Other verbal events may also be recorded in the comments and observations and under the paralanguage section of the record sheet below.

4 Questions per minute (QPM) - indicates the number of question-answer pairs per minute of observation, to the nearest half-question. Each Q-A pair will be counted as one question. When the last question is not followed by an answer within the one minute observation, judgement may be used but this will usually be counted as one-half question for the question per minute score. The number of questions per minute will be estimated to the nearest one-half question.

5 Time to Answer (TTA) - The time to answer each question will be recorded to the nearest second. The time to answer will be defined as the time from the end of the question to the start of the answer. The number of seconds will be recorded in the frame corresponding to the answer (A). Vocalization between the asking of the question and the answer will be ignored in the timing even when some clarification or repeat of part of the question is involved. If the answer is drawn-out, timing may be made to the specific word or phrase that answers the

Appendix VII (continued)

question. For example, in "I -- think - the - answer - is number -- four", the time to answer would be measured from the end of the question to the word "four". If the answer is given and then changes, the TTA will be defined as the time to the student's final answer.

The score for TTA is the average of the TTAs recorded. In the calculation of the average, if there are less than five questions, ($QPM < 5$) then the TTA for the last question will be included in the average even if the last answer is given after the 60 seconds of observation is complete.

Body Movement

6 Position - The position data is used to get rough data on the frequency and magnitude of movement of the subject. For this, the coordinate position of the subject's nose is taken from a grid placed over the video monitor. The coordinates of position in terms of "X" and "Y" displacement are recorded. The coordinate position is noted to the nearest whole number from the grid for both "X" and "Y". The position is noted at the start of each 5 second interval. Thus from the 60 second observation, 13 positions are recorded; one at the start of each interval and one at the end of the last interval. Standardized placement of the camera, and controlled lens settings allow calculation of approximate distance from the grid data. Details are given in Appendix X.

7 Movement (MAMT) - Movement is scored as a one (1) for each interval during which the position coordinates, described above, change. Any change in the "X" or "Y" coordinate position is given a score of zero (0) or one (1). Any actual motion that occurs during the interval is ignored unless it causes changes in the recorded coordinate positions. The score for MAMT is the sum of the number of movements.

Appendix VII (continued)

8 Movement Magnitude (MDST) - A rough calculation of the movement magnitude may be calculated from the X and Y position data. MDST is calculated by the following formula.

$$\sum_{j=1}^{12} [(x_{j+1} - x_j)^2 + (y_{j+1} - y_j)^2]^{1/2}$$

9 Posture Shift (PSFT) - PSFT is identified in terms of a body shift that appears to require some shift in the placement of the buttocks on the seat. Commonly this may be judged by some sitting-up motion or seat readjustment.

10 Lean Back (LEAN) - This specific behavior is defined as the fairly rapid and continuous motion of the head and/or upper body, back or back and slightly sideways. A one (1) is scored for such motion that appears to move the eyes back at least 5 centimeters. Note: From the pilot study it was observed that such motions often occur nearly coincidental with the subject's answering of a question. That is, the subject may slowly lean forward somewhat during the asking of the question and while considering alternative answers, then the subject leans back fairly rapidly just before, during, or just after giving the answer.

11 Head Shake (HSHK) - Negative head shake is scored as one (1) for each occurrence. The head shake is defined by one or more complete cycles of the head in side to side motion.

12 Manipulation (MANP) - MANP indicates active contact of the hands with some part of the body. In this study, camera position limits this category to the manipulation of the head or neck area with the hand or some instrument held in the hands. Examples include twirling the hair with the fingers, scratching the head, pulling on a chain or necklace around the neck, or scratching the head with a pencil. A score of one will be given for each 5-second interval in which some manipulation occurs.

13 Hands to Head (HTH) - HTH indicates some touching of the head or neck with one or both hands. A score of one (1) will be given for each interval in which some touching occurs.

The categories of manipulation and hands to head will be scored independently but a score of one for manipulation will usually imply a score of one or two for hands to head.

14 Gaze Direction - The direction of gaze is scored with the following symbols:

0 - indicates the direction of gaze for the subject was toward the testing materials during the entire 5 second time frame.

1 - indicates some gaze was directed toward the examiner during the time frame.

2 - indicates some other gaze direction during some part of the time frame; i.e. the subject looked away for a moment or two.

Notice that the score of zero (0) indicates continuous behavior during the frame but one (1) or two (2) may represent a change in gaze for only a fraction of the 5 second frame. Where both one (1) and two (2) may apply to the same time frame, both are noted "1-2". The score equals the total number of "ones" and "twos".

15 Eyebrow Flash (EBFL) - EBFL indicates a rapid raising of the eyebrow. The eyebrow flash is usually a behavior that occurs for only a fraction of a second but here is defined to include extended eyebrow raising as well. Such occurrences may be noted in the "comments and other observations". A score of one (1) is given for each occurrence.

16 Eyebrow Lowering (EBLO) - EBLO is an observed lowering of the eyebrow. It will be scored as a one (1) for each occurrence.

17 Wrinkled Nose - Any wrinkling of the nose will be noted here with a score of one (1) for each occurrence.

18 Mouth Twitch (MTWT) - MTWT represents a fairly rapid motion of one or both sides of the mouth back. A score of one (1) will be made for

Appendix VII (continued)

each occurrence during the time frame. From the pilot study it was noted that such twitches may be very brief and often occur almost coincidental with the subject answering the question.

19 Smile (SMIL) - A score of one (1) will be given for each frame during which some smiling occurred. Other frames will be given a score of zero (0).

20 Lip Biting - Mouth movements that indicate the biting of either the upper or lower lips will be noted here. A one (1) will be scored for each occurrence.

21 Lip Licking (LICK) - Mouth movements that indicate some moistening of the lips will be noted here. Judgement will be used and special attention will be paid to mouth movements where the tongue can be seen. A one (1) will be scored for each occurrence.

22 Mouth Drop (MDRP) - A one (1) will be scored for each frame during which the subject's mouth opens without vocalization. When the mouth drops open prior to vocalization, judgement must be used as to whether the mouth opens further in advance of any vocalization than is necessary. Note that vocalization or lip movement during the mouth drop may change the score to "filled pause", "subvocalization" or some other category below.

23 Filled Pause (FLPA) - Each occurrence of a filled pause is scored one (1). The filled pause is defined above under the sequence coding.

24 Subvocalization (SUBV) - SUBV indicates obvious lip movements that may or may not be accompanied by some audible vocalization. Such subvocalization indicates that the subject is rereading the question or that the subject is thinking out loud. Judgement must be used to separate such subvocalization from other "comments" which will simply be noted in the sequence coding as C_s. Judgement will be based usually on the basis of vocal clarity, volume, and intent to communicate.

25 Question Repeat - Each time that some part of the question is repeated by the subject a score of one (1) will be given. Such repeats will also be noted under the coding C_s in the sequence coding and the actual comment may be noted in the "comments and other observations". Examples would include, "... (Did you say), 'times five'".

26 Questioning Answer - Indicates that the answer is given in a questioning tone of voice. Judgement must be used here but such questioning intonation is usually associated with a rising vocal tone or perhaps a drawn out response. In writing a transcript of such a response a question mark would be used. A score of one (1) is given for each such answer.

27 Sigh - A sigh indicates a pronounced exhaling of breath which may or may not be accompanied by some vocalization. A score of one (1) is given for each occurrence.

28 Don't Know - The subject's answer was simply a statement that he (she) does not know the correct answer and that no guess was made. Such an answer will be noted as "A" in the sequence coding but will be noted here with a score of one (1). Unusual response statements may be noted in the "comments and other observations" section.

29 Comments and Other Observations - Any unusual behavior may be noted here. In addition, as stated above, this section will include some clarification of coded observations made above.

Other Nonverbal Behaviors Observed

The following nonverbal behaviors were not included in the behavior list of the observational instrument record sheet. These behaviors were observed during the analysis of main study videotapes. This list is limited to behaviors that were observed for more than one student. These behaviors were recorded on the score sheet under "Comments and Observations". None of the following behaviors were observed in more

than 25 percent of the student observations, and usually the behavior was observed for only two to four of the students.

Description of the Behaviors

Rocking foot - This behavior may often be detected even when the foot is not in view. The behavior may be defined and detected by the rapid vibration of the body due to the rapid rocking of the foot.

Shoulder Shrug - The shrug is denoted as a rapid movement of the shoulders up and sometimes forward. The subject may lower the head simultaneously.

Head Cocked - A head cock is a tilt of the head sideways. It may be accompanied by some shoulder shrugging.

Head Wag - A head wag is a movement of the head side to side at least one complete cycle.

Head Toss (Hair Toss) - This is a throwing of the head back with a quick motion. Often the intent may seem to be to toss the hair back and out of the face.

Wide Eyes - This is an opening of the eyes wide without significant other eye or eyebrow movement. In stop motion or still picture the gesture may appear to be a stare but seen in videotape motion this subtle nonverbal behavior may last less than a second and is seldom longer than several seconds.

Closed Eyes - This behavior is defined as a closing of the eyes longer than for blinking. This behavior seems to occur while the student is attempting to think of an answer.

Tongue in Cheek - This behavior is denoted when the tongue pushes out on the cheek.

Appendix VII (continued)

Tight Smile - This behavior may occur rapidly enough that it would be called a mouth twitch or be included under the behavior "smiling". However, in this case the smile is not relaxed. The mouth is tightly closed and pulled back on both sides. This behavior definitely seems tense and controlled and may communicate a message different than the smile.

Swallowing - This may be difficult to detect in some subjects due to camera angles or clothing styles. Differences in male and female anatomy of the throat complicate technique. Swallowing may be indicated or defined by the motion seen in the Adam's apple.

Laugh - Laughing is defined in the ordinary sense. Laughing may range from nervous laughing and may be associated with students guessing difficult questions or laughing at the simplicity of the easy questions.

Quiet Answer - This is a response that is barely audible. Certainly instrumentation could be used to quantify and objectify the process of measuring loudness of student responses although the additional technology may not be practical except in the laboratory research situation.

Drumming the Table - This behavior is defined in terms of the common usage of the expression. The behavior may be indicated by tapping fingers on the table but could be detected by both audio and visual record. This behavior may be a specific sub-classification of manipulation as already defined.

Appendix VII (continued)

Subject Name _____ Date of Testing _____ Test _____ Number _____

Age: Yr _____ Mo _____ Tape Number _____

Grade _____ Start _____ Stop _____ Start Phrase _____

Sex _____ Start _____ Stop _____ Start Phrase _____

	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12
1 Sequence.....																								
2 Examiner.....																								
3 Subject.....																								
4 Questions per min..																								
5 Time to Answer.....																								
Motion:																								
5 position X.....																								
position Y.....																								
7 Change.....																								
8 Magnitude.....																								
9 Posture Shift...																								
10 Lean Back.....																								
11 Head Shake.....																								
Manipulation:																								
12 Manipulation....																								
13 Hands to Head...																								
Eyes:																								
14 Gaze.....																								
15 Eye Brow Flash..																								
16 Eye Brow Lower..																								
Nose:																								
17 Movement.....																								
Mouth:																								
18 Mouth Twitch....																								
19 Smile.....																								
20 Lip Bite.....																								
21 Lip Lick.....																								
22 Mouth Drop.....																								
Paralanguage:																								
23 Filled Pause....																								
24 Subvocalization.																								
25 Question Repeat.																								
26 Questioning Ans.																								
27 Sign.....																								
28 Don't Know.....																								
Other: _____																								

29 Comments and Observations _____

APPENDIX VIII

RELATIVE FREQUENCY OF BEHAVIORS OBSERVED

The table below gives the relative frequency of each of the behaviors observed during the main study. The frequency is reflected in the percentage of one-minute behaviors observations during which each of the behaviors was observed at least one time. This calculation included both the pre-frustrational and frustrational behavior observations and both the PIAT and PPVT testing conditions.

Table VIII-1. Behavior frequencies (%).

Behaviors Listed*		Other Behaviors	
PSFT	29%	Sigh	10
LEAN	41	Don't Know	1
HSHK	36	Shoulder Shrug	17
MANP	51	Head Cocked	14
HTH	51	Wide Eyes	10
GAZE	87	Hair (Head) Toss	10
EBFL	56	Swallow	13
EBLO	34	Rocking Foot	6
**NOSE	10	Pursed Lips	4
MTWT	96	Drums Table	2
SMIL	50	Tight Smile	6
**LIP BITE	23		
LICK	70		
MDRP	50		
FLPA	34		
SUBV	66		
**Q-Repeat	11		
**Q-Answer	24		

*These behaviors were listed on the observational instrument and specifically identified and defined before the main study.

**These behaviors were not seen with sufficient frequency to be used in the analysis of the data from the main study. In each case the behavior was seen in less than 25% of the observations.

Appendix VIII (continued)

Table VIII-2. Average values for nonverbal behavior variables.

Behavior	PPVT Data		PIAT Data	
	Pre-frustration	Frustration	Pre-frustration	Frustration
Variable				
TTA	***2.13	4.39	***2.06	10.93
AMTM	8.89	9.05	5.13	4.80
DSTM	1.24	1.31	.56	.58
PSFT	.24	.16	.28	.22
LEAN	** .13	.55	.25	.50
HSHK	.08	.26	.18	.25
MANP	*1.32	2.16	*1.13	1.84
HTH	2.05	2.54	1.25	1.47
GAZE	1.00	1.26	***2.47	1.22
EBFL	*** .37	1.26	* .28	.63
EBLO	* .26	.53	* .25	.63
MTWT	***1.74	3.08	***1.47	2.72
SMIL	.61	.53	**1.47	.53
LICK	.58	.74	.82	.75
MDRP	** .16	.55	* .44	1.03
FLPA	.45	.50	.41	.44
SUBV	*** .55	1.68	*** .50	1.81

*p < .05

**p < .01

***p < .001

APPENDIX IX
NOTES ON THE CLUSTERING PROGRAM

*CLUSB

PROGRAM PURPOSE

This program clusters observations of multivariate data. All observations are assigned to the first cluster and the mean of this cluster is computed. The observation with the largest scaled distance from the cluster mean is selected to start the next cluster; the observation serves as the new cluster mean. Each point is examined to see whether its distance to the new cluster mean is less than its distance to the old cluster mean. If the distance is less, the point is moved to the new cluster, altering the new cluster mean. For subsequent steps, the observation with the largest distance from its cluster means initializes the new cluster, and all points are examined for reclassification until the desired number of clusters is created. There is a limit of 20 clusters, 20 variables (dimensions) in an observation and 100 observations.

Scaling factors are calculated once, in the beginning of the program, and maintained throughout the clustering procedure. The purpose of scaling is to eliminate the dominating effect of variables with great variation and to allow the investigator to weight some variables more heavily than others. The scale factors, c , are applied in the following manner.

Let c_j represent the scale factor for variable j . Then the scaled distance, d , of observation i from a mean is calculated by

$$d^2 = \sum c_j (x_{ij} - \bar{x}_j)^2$$

where $c_j = w_j^2 / s_j^2$

s_j^2 is the variance among the observations of variables, and w_j is the prescribed weight for variable j .

Data records must contain an identifier (Species I.D.) for each observation as the first item read from the record. The user supplies a data format for the identifier and the variable vector. A few simple transformations are available.

More than one set of data may be analyzed by repeating the header records for each data set and stacking the data sets in one file. One set of data may be analyzed more than once by repeating indicated parts of the header records.

APPENDIX X

DIAGRAMS AND NOTES ON THE EQUIPMENT SETUP
DURING THE MAIN DATA COLLECTION