An investigation of the effect of verbal teacher-student interaction of similar and different ability groups in secondary classrooms.

H. Jerome Freiberg

University of Massachusetts Amherst

Follow this and additional works at: http://scholarworks.umass.edu/dissertations_1

Recommended Citation

http://scholarworks.umass.edu/dissertations_1/2593
AN INVESTIGATION OF THE EFFECT OF VERBAL TEACHER-STUDENT INTERACTION OF SIMILAR AND DIFFERENT ABILITY GROUPS IN SECONDARY CLASSROOMS

A dissertation Presented
By

H. Jerome Freiberg

Submitted to the Graduate School of the University of Massachusetts in partial fulfillment of the requirements for the degree of

DOCTOR OF EDUCATION

September 1972

Major Subject: Teacher Education
ACKNOWLEDGEMENTS

This dissertation is dedicated to Dr. William V. Fanslow, and Dr. Jimmie C. Fortune who guided me through the tribulations of my graduate career and Dr. Mary Alice Wilson and Dr. R. Mason Bunker for their expertise and assistance with my dissertation. I also wish to thank the teachers and students who participated in the study.

Finally, I wish to thank my high school counselor (Northeast High, Philadelphia, Pennsylvania) who recommended I become an auto mechanic.
Table of Contents

Introduction - Chapter I

| a. Statement of the Problem | 4 |
| b. Pilot Study | 8 |
| c. Discussion on the Statistical use of Flanders Interaction Analysis | 11 |
| d. Statistical Analysis of Pilot Study | 15 |
| e. Conclusions - Pilot Study | 24 |

Rational - Chapter II

| a. Historical Background | 26 |
| b. Existing Research | 27 |
| c. Summary | 36 |
| d. Hypotheses - Introduction | 37 |
| e. Hypotheses | 42 |

Procedures - Chapter III

| a. Location | 45 |
| b. Subjects | 47 |
| c. Methodology | 47 |
| d. Summary | 52 |
| e. Analysis of Data | 52 |
| f. Teacher Effectiveness Questionnaire | 53 |
| g. Training and Reliability of Observers | 56 |
| h. Validity of Observations | 56 |
| i. Summary | 57 |

Findings - Chapter IV

| Summary of Findings | 62 |
| Analysis of Findings | 73 |
| Summary | 74 |

Conclusions - Chapter V

| iv |
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>84</td>
</tr>
<tr>
<td>Recommendations</td>
<td>85</td>
</tr>
<tr>
<td>Summary</td>
<td>90</td>
</tr>
<tr>
<td>Bibliography</td>
<td>92</td>
</tr>
<tr>
<td>Appendix A</td>
<td>95</td>
</tr>
</tbody>
</table>
Charts, Tables and Figures

<table>
<thead>
<tr>
<th>Chart/Table/Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chart I</td>
<td>Teacher-Student Interaction With Two Different Ability Groups</td>
<td>7</td>
</tr>
<tr>
<td>Table I</td>
<td>Comparison of Phases Taught by Teachers A and B Using Darwin's Chi-Square Analysis</td>
<td>16</td>
</tr>
<tr>
<td>Table II</td>
<td>Comparison of Phases Taught by Teacher A Using Darwin's Chi-Square Analysis</td>
<td>18</td>
</tr>
<tr>
<td>Table III</td>
<td>Flanders Interaction Analysis Frequency Comparisons of Matrices for All Phase Levels (Teacher A)</td>
<td>21</td>
</tr>
<tr>
<td>Table IV</td>
<td>Flanders Interaction Analysis Frequency Comparisons of Matrices for All Phase Levels (Teacher B)</td>
<td>23</td>
</tr>
<tr>
<td>Table V</td>
<td>Classroom Statistics</td>
<td>48</td>
</tr>
<tr>
<td>Table VI</td>
<td>Classroom Taped Observations</td>
<td>49</td>
</tr>
<tr>
<td>Figure I</td>
<td>Observation Validity Sheet</td>
<td>60</td>
</tr>
<tr>
<td>Figure II</td>
<td>Teacher Effectiveness Questionnaire</td>
<td>61</td>
</tr>
<tr>
<td>Figure III</td>
<td>Flanders Interaction Analysis Indirect/Direct Ratios</td>
<td>63</td>
</tr>
<tr>
<td>Figure IV</td>
<td>Flanders Interaction Analysis Revised Indirect/Direct Ratios</td>
<td>64</td>
</tr>
<tr>
<td>Figure V</td>
<td>Flanders Interaction Analysis Category Three-Use of Student Ideas</td>
<td>65</td>
</tr>
<tr>
<td>Figure VI</td>
<td>Teacher-Student Responses on the Teacher Effectiveness Questionnaire</td>
<td>68</td>
</tr>
<tr>
<td>Figure VII</td>
<td>Student Flexibility In Ability Grouped Sample Population</td>
<td>70</td>
</tr>
</tbody>
</table>
CHAPTER 1

AN INVESTIGATION OF THE EFFECTS ON VERBAL TEACHER-STUDENT INTERACTION OF SIMILAR AND DIFFERENT ABILITY GROUPS IN SECONDARY CLASSROOMS

Introduction

The problem of ability grouping in the American educational system has created much heated discussion and debate regarding its success for developing a conducive learning environment. Many educators and parents have criticized the criteria (usually IQ) for assigning students to a particular ability level. In recent years the focus on IQ as a tool for measuring student ability has received additional criticism from minority groups. They perceive IQ tests as a discriminatory measure which inaccurately describes lower socio-economic students with middle class terminology, concepts and values. A recent study of Langston Bannister (1971) found black students representing a disproportionate number of lower ability groups. Using IQ to determine ability level, therefore, has the effect of creating de facto segregation within a legally integrated school.

Ability grouping instinctively appeals to educators and laymen alike" "Allow those students to work at their particular level" was the response of one teacher and a school principal in a rural school system.
The concept aids the principal and his staff in that it is convenient to place those students who work at the same level, as determined by IQ tests, in one class. At a recent faculty meeting in a large urban school system, the principal threatened the faculty with heterogeneous grouping if they did not stop complaining about working conditions. This particular group of teachers were intimidated more by the threat of heterogeneous groups than their need for improved working conditions.

Although ability grouping appeals to teachers and administrators, the process of transferring the idea of ability grouping into a workable school program often destroys the original purpose of allowing students to work at their own academic levels. The self-fulfilling prophecy, lack of movement of students from different ability groups, and negative self images have been noted to deter the student's freedom to work at their own level (Rosenthal and Jacobson, 1968). Even a typical homogeneous class (Borg, 1965) contains a diverse population of ability levels. Many teachers in the schools feel that up to a quarter of their students belong in a lower ability group while another quarter belong in higher ability groups.

The above problems have developed through misunderstanding, ignorance, and lack of communication between those educators who have developed the theoretical concepts and the teachers who implement
those concepts in the classroom. This malady is not uncommon in the field of education and is supported by many government agencies, but solutions cannot be sought until those involved in the use of ability grouping review significant evidence and realize that their approach needs revision, reevaluation and reexamination of different view points.
Statement of the Problem

This study will attempt to present a systematic view of specific relations among variables with the intent of elucidating and predicting the phenomena. (Kerlinger 1964). Value judgements regarding the question of whether ability grouping is inherently good or bad will not be answered, but rather the directional relationship among specific variables will be observed as ability grouping is implemented in one particular secondary school.

To talk about ability grouping in general seems idle since it makes a difference "how rigid or uniform the learning situation is..." (Cohn, Ruth, and Isenberg, 1967) for the student in the classroom and his teacher. The term "ability grouping" presents different images for different members of the educational community. Some teachers see it as a means of separating the intelligent and "good" students from those who are "dumb" and cause problems in the classroom. Administrators view ability grouping as a vehicle for increasing the achievement level of the students (Borg et al). Some educators view ability grouping as a way of meeting the needs of individual students by keeping the pace of learning constant throughout the class. Many parents see ability grouping as a status symbol for those upper ability grouped students and as a means of keeping the stereotyped slower and troublesome students from interfering with the achievement of their own college
bound children.

Most of the rationale around ability grouping (Borg et al) stems from the belief that ability grouping increases achievement, through differentiated curriculum, thus meeting individual student needs (Ekstrom, 1959). Selection criteria for ability grouping, as previously stated, have been drawn into controversy as educators begin to question the reliability of standard IQ tests for certain socio-economic groups in our society. The "side effects" of ability grouping include such problems as:

1. Negative self-concepts (Rosenthal and Jacobson et al)
2. Stereotyping of students' abilities by their teachers (Schrank, 1970)
3. Predetermined expectations of teachers which may build or destroy the self esteem of the child (Smith, 1969)
4. Tracking students into ability groups without the flexibility of movement from one level to another (N.E.A., 1968)

Three of the four side effects focus on the teacher as the crucial variable in the success of ability grouping. This study will examine those variables which can be observed in actual classroom settings. Specifically, individual teacher's verbal interaction with two different ability grouped classes will be observed and analyzed to see if any change in verbal patterns between the two ability groups exists. (See Table I,
p. 16, which shows the comparisons between the first observation 
(O\textsubscript{1}) and second (O\textsubscript{2}) for each phase of teacher A and B.)

A pilot study was constructed to generate hypotheses regarding 
the effects of teacher-student interaction of similar and different ability 
groups in secondary classrooms. The hypotheses from the pilot study 
were directional in nature and are tested in this dissertation to ascer-
tain their appropriateness in developing conclusions regarding the effects 
of teacher-student interaction in secondary ability grouped classrooms.
CHART 1

Teacher-Student Interaction With Two Different Ability Groups

TEACHER A

CLASSROOM

INTERACTION

HIGH ABILITY GROUP

CLASSROOM

INTERACTION

LOW ABILITY GROUP
Pilot Study

The pilot study was conducted at a regional junior-senior high school in western Massachusetts with the cooperation of two female social studies teachers. The two teachers conducted classes which included both the upper ability groups or as the school referred to them, Phases IV and III, and the lower ability groups, or Phases II and I of seventh, eighth, and ninth grade students. The school identified Phase IV as the "best" students and the other phases correspond in descending order to the lower ability groups.

Each of the teacher's ability groups was observed twice during a three month period. The first set of observations for teachers A and B (O₁) were concluded during a one week period. Three months later a second set of observations for teachers A and B (O₂) was completed during another one week period. In order to keep the variables of content somewhat equal, the classes viewed by the researcher were designated in advance by the classroom teacher as test review sessions. The teachers were not chosen at random for the following reasons: first, there were only a few teachers willing to participate in the study. Second, teachers with both upper ability groups (Phase IV) and lower ability groups (Phase II) were needed to give the broadest possible range of phases for statistical analysis.
With the teachers as the focal point of the study, verbal interaction patterns between the teacher and the various levels of ability grouped students were analyzed using Flanders Interaction Analysis' Ten Category System (Flanders, 1965). Flanders Interaction Analysis enables the researcher to quantify teaching behaviors so he can produce relatively unbiased judgements as to the verbal interaction patterns in the classroom. The need for various statistical research tools was investigated, and the data obtained with Flanders gave an adequate picture of the problems and directions needed to synthesize the pilot study into this doctoral thesis.

Utilizing the percentage and ratio statistics of Flanders (See Appendix A which shows the computations of Indirect/Direct Ratios from Flanders Interaction Analysis.), ten variables of classroom interaction were measured.

1. Percentage of teacher talk as compared with student talk
2. General pattern of indirectness or directness in the classroom
3. Types of motivation used by the teacher
4. Differences in motivation used by the teacher
5. Degree of reinforcement the teacher employs after student responses
6. Extent to which the class is motivated without the use of content
7. Type of responses the teacher reinforces
8. Types of questions the teacher asks the students
9. Extent to which the teacher uses the students' ideas
10. General teaching pattern within the classroom

These variables were used to test the notion that: ability grouping creates a direct interaction pattern as measured by Flanders between the teacher and her lower ability groups, while an indirect interaction pattern is created between the teacher and her upper ability groups.

The ten Flanders categories were divided into indirect/direct ratios. Categories Number 1 through Number 4 have been labeled indirect teacher responses to students (Flanders, 1965) and categories Number 5 through Number 7 have been labeled direct teacher responses to students. (See Appendix A, which described the Ten Flanders Categories.) In general, teachers who use more indirect responses foster higher achievement rates than teachers who use more direct teacher responses (Flanders, 1965).

The classroom interactions were recorded via a tape recorder placed in the rear of the room. A baseline time was allowed for the elimination of any novelty effect before the actual recording of data was begun.

After the recording sessions were completed tallies were made while playing back the recordings using Flanders Interaction Analysis.
The tallies recorded during the observations were recorded onto data cards and read into Program Interact (Whightman, 1970). Program Interact produces a ten-by-ten matrix for the ten categories of the Flanders Interaction Analysis System. Program Interact produces all Indirect/Direct ratios and Chi-Square statistics for each matrix.

The data collected from the pilot study which is explained in the following pages indicated that ability grouping may have effected teacher behavior. This motivated the author to design a study and pursue the question further.

Discussion on the Statistical Use of Flanders Interaction Analysis' Ten Category System

The use of Flanders Interaction Analysis presents some problems in that, "the exact usefulness of I.A. in predicting pupil achievement remains to be seen" (Rosenshine, 1970). The use of Flanders Interaction Analysis (Flanders, 1965) assumes certain givens; first, that an indirect teacher increases the achievement level of her students and that a direct teacher's behavior limits the pupil's freedom, as defined by Flanders. Rosenshine's review of Flanders Interaction Analysis and in this author's own readings there appears to be a deficiency of valid statistical tools to analyze observational systems. The validity errors include the use of students for the unit
number of analysis rather than the teachers who were the sampling unit (Rosenshine, 1970) which would increase the population size considerably. In many of the studies less than five teachers were used. The authors in these cases were generalizing past the boundaries for a study with a limited population. In the review of the *Journal of Research and Development in Education* (Fall, 1970) the observational systems which were explored used simple t-tests for expanded Flanders categories without considering adjustment of t-tests. The simple mean comparisons ignored the fact that a closed matrix system created the problem of didactic variables; i.e., the cell which receives a specific number of tallies affects the outcome of the remaining categories. The problem also exists that the relationship between Flanders categories and student achievement may not represent a linear relationship (Soar, 1968). The use of Chi-Square statistics within and between teachers compounds the problems of didactic variables and the comparison of dependent variables. Most of the statistics used for Flanders and other observational systems assume a normal distribution, but the closed matrix system prevents that approach. Some systems avoid these problems by using frequency counts of specific variables, leaving the system open; however, these create other problems in rater reliability.
Since Flanders developed his interaction system (Flanders, 1960), some 400 independent observational systems have been created, but fewer than 15 have been developed and used to relate classroom behaviors to student growth (Rosenshine and Furst, 1971). Many of the 15 observational systems face the problems of inadequate or inappropriate statistical analyses (Greenberg, 1970). The research problem of ability grouping is being approached with a tool that needs special and perhaps newly developed statistical methods. Since most of the other observational systems investigated suffer from similar maladies, the use of Flanders Interaction Analysis is justifiable.

Although numerous studies report significant findings using Flanders Interaction Analysis or other observational systems based on it, present statistical analysis beyond graphic representations are both inappropriate and invalid. At present no statistical techniques exist which will compensate for interviewing variables which accompany observational systems, without violating assumptions about the statistics being used for the analysis. For example, in the Flanders' system, category 5 usually receives the greatest number of tallies, with category 1 receiving the least amount of tallies. It is assumed with a normal distribution that each category would have an equal opportunity to receive a tallie. With a closed matrix system this assumption is invalid.
The use of Flanders Interaction Analysis and its ten categories in this study presents the researcher with established value loaded words such as praise (category 2), empathy (category 1), directions (category 6) and criticism (category 7). These categories present the reader with positive and negative connotations regarding teaching behaviors. Criticism is not inherently a negative or inhibiting action nor is empathy necessarily a positive or freeing action. A teacher who never uses criticism may be as effective as a teacher who freely uses criticism.

The statistical analyses (e.g. Indirect/Direct ratios) developed by Flanders incorporates the first seven categories into a formula for analyzing the directness or indirectness of teaching. Flanders expresses his value preferences toward indirect teachers who he feels have a positive effect on achievement by freeing the learning environment, (Flanders, 1965).

Value judgements regarding the categories and statistical analyses utilized in Flanders Interaction Analysis will be avoided, with the results of this study being expressed in terms of the direction and consistency of specific relationships among the variables tested. This approach will possibly counterbalance observational instruments which have certain value weighted designs (e.g. Flanders Interaction Analysis).
Statistical Analysis of Pilot Study

The ten variables presented in the pilot study were analyzed using a "goodness of fit" statistical approach (Wyatt and Bridges, 1966). This approach was selected after investigating the statistical problems in the use of Flanders'. A goodness of fit analysis enables the researcher to compare matrices with each other. Using a total matrix for each of the classes of teacher A and B, eight matrices are presented for teacher A and six matrices for teacher B. (See Table I below.) Keeping in mind the previous discussion on matrix analysis and its problems relative to Flanders', Darwin's Chi-Square goodness of fit analysis was used between teacher A and B. Analysis was also made between ability levels for each of the teachers. (See Table II, p. 18, which shows the comparisons between each phase for teacher A for the first (O₁) and second (O₂) observations. It also describes the cross comparisons for O₁ and O₂ for the four phase levels). The first analysis tested the similarities and differences between the first observation (O₁) and the second observation (O₂) along identical phases. (See Table I above.) Since three months elapsed between O₁ and O₂, analyses were made to test similarities and differences between the two observations. For teacher A, only Phase IV developed difference at the .005 level;
TABLE I

Comparison of Phases$^1$ Taught by Teacher A. and B. Using Darwin's Chi-Square Analysis$^2$

<table>
<thead>
<tr>
<th>Teacher A.</th>
<th>Teacher B.</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td>Second</td>
</tr>
<tr>
<td>Observation</td>
<td>Observation</td>
</tr>
<tr>
<td>IA.</td>
<td>2A.</td>
</tr>
<tr>
<td>Phase IV</td>
<td>.005</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Phase III</td>
<td>Not sig.</td>
</tr>
<tr>
<td></td>
<td>at .10 level</td>
</tr>
<tr>
<td>Phase II</td>
<td>Not sig.</td>
</tr>
<tr>
<td></td>
<td>at .10 level</td>
</tr>
<tr>
<td>Phase</td>
<td>Not sig.</td>
</tr>
<tr>
<td></td>
<td>at .10 level</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 Phase IV represents the top ability group, Phase III the average ability group, Phase II the lower ability group and Phase I the lowest ability group. Phases IV, III, II and I average a class size of thirty students, while Phase I usually limits its class size to ten students.

2 Darwin's Chi-Square is utilized by Dr. Lawrence Wightman as his statistical analyses for his Program Interact, which I used to run and analyze my data.

3 The significance between Teacher A. and B. represents the comparing of matrix IA. and 2A. of Teacher A., with IB. and 2B. of Teacher B.
the other three phases of teacher A did not develop any significant differences. Teacher B showed significant differences with Phase II. The other two phases did not show significant differences at the .10 level. The generalization is made that over a period of time the two teachers interacted identically with similar phases. (See Table II.) Between teacher A and B, Phases IV and III showed significant differences in classroom interaction at the .005 level, with Phase II for both teachers developing no significant differences. From this analysis one has evidence that teacher A and teacher B interact differently with their classes except on the lower ability level of Phase II. Teacher B did not have a Phase I class. Therefore, comparison for Phase I was limited to teacher A.

One important variable tested was the change in teacher-student interaction which occurred as the teacher interacted with different phase levels. Using Darwin's Chi-Square (Darwin, 1959) on the first set of observations \( (O_1) \) of teacher A, matrix comparisons were tested between Phases IV and III, III and II, II and I. Significant differences were obtained in all four phase levels. (See Table II) The matrix comparisons were repeated for the second set of observations \( (O_2) \) for teacher A. A confidence level of .05 was set as the significant bounds for this study. The differences
Table II

Comparison of Phases Taught by Teacher A. Using Darwin's Chi-Square Analysis

Teacher A.

<table>
<thead>
<tr>
<th>First Observation</th>
<th>Second Observation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A</td>
<td>1A</td>
</tr>
<tr>
<td>Phase IV</td>
<td>Phase IV</td>
</tr>
<tr>
<td>(sig.) .010-.025</td>
<td>(sig.) .05-.10</td>
</tr>
<tr>
<td>Phase III</td>
<td>Phase III</td>
</tr>
<tr>
<td>(sig.) .025-.05</td>
<td>(not sig.) at .10 or less</td>
</tr>
<tr>
<td>Phase II</td>
<td>Phase II</td>
</tr>
<tr>
<td>(sig.) beyond .005</td>
<td>(sig.) beyond .005</td>
</tr>
<tr>
<td>Phase I</td>
<td>Phase I</td>
</tr>
</tbody>
</table>

Cross Comparison of Phases Taught by Teacher A. Using Darwin's Chi-Square Analysis

- Phase IV (1A.) Significant between .025-.05  
  Phase III (2A.)
- Phase IV (2A.) Significant between .05-.10
  Phase III (1A.)
- Phase III (1A.) Not significant at .10 level
  Phase II (2A.)
- Phase III (2A.) Significant between .05-.10
  Phase II (1A.)
- Phase II (1A.) Significant beyond .005 level
  Phase I (2A.)
- Phase II (2A.) Significant beyond .005 level
  Phase I (1A.)
between Phase IV and III approach very closely to a significant level. The differences between Phase III and II show no significance while comparison between Phase II and I achieved significance beyond the .005 level. (See Table II) Of the six possible measures between individual matrices, four showed significant differences, one approached significance and a comparison of matrices between O1 and O2 showed significance only at the Phase IV level. This analysis directed the author to conclude that teacher A interacted with each of her four phases differently, but over a period of three months (O1-O2) teacher A interacted with identical phases in a similar manner.

The next statistical comparison followed the procedures of the analysis described above, with a cross comparison of O1-O2 of Phases IV and III, Phases III and II, and Phases II and I. (See Table II.) The comparisons either achieved or approached significance except for the comparison of Phase III (O1) with Phase II (O2). The greatest significance was achieved with the comparison of the two lower phases.

Program Interact output produces a matrix for each phase and calculates the number and percent of tallies in each phase and calculates the number and percent of tallies in each of the ten categories. It also generates ratios between the Indirect/Direct categories and generates elementary statistical analysis. (See Appendix A) If teacher responses are tallied in categories Number 1
through 4 with greater frequency than categories Number 5 through 7, Flanders would consider the teacher to be more indirect. The sum of the indirect categories divided by the sum of the direct categories produces an Indirect/Direct ratio of classroom interaction.

Each of the ten Flanders categories represents defined interaction. A graphic comparison (See Table III, p. 21, which graphically compares the upper, Phase IV and III, and lower, Phase II and I, phases for teacher A with each of the ten Flanders categories) of the percent of direct and indirect verbal behaviors used by the same teacher with different phases dramatically shows how the teacher tends to be more indirect with the higher phases and more direct with the lower ones.

Table III represents the mean percents between $O_1$ and $O_2$ for each phase of teacher A. The upper phases (Phase IV and III) were grouped into one by taking the mean percents of two phases. The same procedures were followed for the lower phases (Phase II and I). The differences are very distinct as the upper phases received a higher percentage of tallies in the indirect categories and a lower percentage of tallies in the direct categories. The lower phases received a lower percentage of tallies in the indirect categories and a higher percentage of tallies in the direct categories.
Table III

Teacher A.

Flanders Interaction Analysis Frequency Comparisons of Matrices For All Phase Levels

\[\begin{array}{cccccccccccc}
\text{Flanders Categories} & \text{X Phases IV and III} & \% & \text{X Phases II and I} & \% \\
\hline
\text{Upper Phases) Phase IV and III} & 0.0\% & 3.76\% & 3.35\% & 13.74\% & 35.83\% & 3.79\% & 3.53\% & 15.66\% & 30.44\% & 14.54\% \\
\text{Lower Phases) Phases II and I} & & & & & & & & & & \\
\end{array}\]
Table IV represents the mean percents between $O_1$ and $O_2$ for each phase of teacher B. These graphs present a picture of how the percentages relate to the Flanders categories. Appendix A gives the parameters for considering each of the ten categories. In the first three categories for teacher B, the distributions seem fairly close except for Phase III which fluctuates from each of the other phases. Teacher B's differences are greater between Phase IV and II, with a substantial amount of indirect interaction given to her upper phase, and a higher percentage of direct interaction for Phase II. In the Flanders system, an interaction 9-10-9 categorizes the degree of student-student interaction. For teacher B nearly 40 percent of Phase IV and 30 percent of Phase II in categories Number 9 and Number 10 represent this student-student interaction. An interesting side note is that teacher B was not rehired because of her lack of "control" of her classes especially with her lower phase classes. One cannot infer a cause and effect relationship, but it does seem to be an "interesting" relationship.
Table IV

Teacher B.

Flanders Interaction Analysis Frequency Comparisons of Matrices For All Phase Levels

<table>
<thead>
<tr>
<th>Phase</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase IV</td>
<td>.16</td>
<td>1.88</td>
<td>3.46</td>
<td>13.62</td>
<td>27.00</td>
<td>.16</td>
<td>.67</td>
<td>2.62</td>
<td>39.81</td>
<td>10.31</td>
</tr>
<tr>
<td>Phase III</td>
<td>0.0</td>
<td>.54</td>
<td>7.15</td>
<td>20.39</td>
<td>54.20</td>
<td>.94</td>
<td>2.50</td>
<td>11.40</td>
<td>13.55</td>
<td>3.79</td>
</tr>
<tr>
<td>Phase II</td>
<td>0.0</td>
<td>2.08</td>
<td>4.73</td>
<td>18.29</td>
<td>9.44</td>
<td>4.12</td>
<td>3.29</td>
<td>15.12</td>
<td>22.96</td>
<td>19.09</td>
</tr>
</tbody>
</table>

Phase IV
Phase III
Phase II
Conclusions From the Pilot Study

The pilot study has refined the thinking of the author regarding the role of the teacher in dealing with ability groups and the structure within which the teacher must operate. Further exploration includes an examination of the flexibility of ability groups, and a focus, in greater detail, on teacher interaction in the classroom. Ability grouping must be seen in the context of specific criteria and processes. For example, a system that bases its ability grouping or phases, on IQ tests and grade point averages, and proceeds to keep the student phased into one particular ability group for the duration of the school year is different from a system which allows the student the flexibility to be in numerous ability groups for an undesignated amount of time.

The role of the teacher in ability grouping is explored in relation to the school structure, classroom interaction and teacher-student perceptions of teacher effectiveness.

Flanders Interaction Analysis was used to observe teacher-student verbal interaction in the classroom. Graphic analysis at the present time is the only viable statistical tool available to deal with the problems of closed observational systems. Although there is a divergence of opinion regarding the relationship between achievement and Indirect-Direct teachers (Flanders, 1965 and Rosenshine and Furst, 1971) the
direction of this study is not to test that relationship. Rather, this study observes teacher behavior with children in ability groups using Flanders as an observation tool with supplemental information gained through student questionnaires.

From the pilot study hypotheses are generated regarding the effects of teacher-student interaction on different and similar ability groups in secondary classrooms.
CHAPTER II
RATIONALE

Historical Background

The concept of ability grouping began with the one-room school house where a mixture of subjects, ages, grades and learning tasks composed the classroom milieu. This situation gave way to the construction of city schools of the 1900's with their grade and subject oriented classrooms. The multi-age, -grade and -subject level one room school house developed into departmentalized subject oriented homogeneous classrooms. The number of students seeking education in the city schools contributed to the need for separation of students according to their ability in the classroom.

From this quiet beginning the practice of ability grouping has gained the attention of educational researchers (Borg, 1965) and the educational community. Most of the controversy around ability grouping stems from the belief that grouping according to ability increases achievement through differentiated curriculum which meets individual student needs (Ekstrom, 1959). The criteria for ability grouping however, has created a great controversy with teachers, parents, and educators as they begin to question the reliability of standard intelligence tests for the many socio-economic groups in our society.
Existing Research

The extent of research pertaining to the general area of ability grouping is overwhelming. Although the quantity is prodigious, the quality of the reported research leaves a great deal to be desired. The Rosenthal study, *Pygmalian in the Classroom* (Rosenthal and Jacobson, 1969), has been widely circulated and quoted as a source on the "Self-Fulfilling Prophecy and Ability Grouping," yet it suffers from numerous statistical and design errors. Critical discussions on the Rosenthal study were presented by Thorndike (1968) and Gumpert and Gumpert (1968). The study was recently replicated (Jose, 1970) with no significant findings in the area Rosenthal and Jacobson found significant differences. However, the concepts raised by Rosenthal and Jacobson seem valid and should be explained in future studies.

In a recent comprehensive review of the literature of ability grouping since 1922, Borg (1965) found many of the studies had flaws in their statistical analyses, design, or generalizability. Borg conducted a four year study including some 4,000 students from the fourth, sixth, seventh, eighth, and ninth grades from two school districts. Although he states: "ability plays a large role in all secondary schools in determining which students will take the more advanced electives in science and mathematics," he stops his study at the ninth grade.
Borg analyzes ability grouping in terms of general achievement, overachievers and underachievers, study methods, attitudes, socioeconomic choice, pupil attitudes, pupil problems, self-concept and personality. Using achievement as the measured variable for the ninth grade math and science classes Borg concludes:

...there is some tendency for ability grouping to lead to greater mathematics achievement among superior pupils and greater science achievement among average pupils. Among the slow pupils, the differences between the two group treataments were small, but tended to favor the non-ability group (heterogeneous group) in both mathematics and science.

Average students gained in study methods and attitudes in heterogeneous groups for the secondary level. A summary of Borg's analysis tends to show, at least for the junior high level, that results were incomplete and found little significant effect on the achievement of students who were ability grouped. However, the analyses of self-concept did show less favorable attitudes associated with ability grouped classes. Although short-range achievement scores are not affected, the possibility of long range effects through negative self-concepts needs to be explored.

The Borg study contains one of the most comprehensive review of literature on ability grouping, and presents an extensive study encompassing a number of variable. Although his study is quite extensive, it does encounter a major problem which Borg does not confront, that of the role of the teacher, Further studies reviewed in this section will therefore look at the role of the teacher from different perspectives.
In another series of studies concerned with the effects of ability grouping (Schrank, 1968-70), grade point average was used to determine the effects of ability versus non-ability grouped classes. Schrank found significant difference by randomly assigning ability-level labels to randomly grouped mathematics classes. In another study (Schrank, 1970), teachers were notified as to the groups randomness. No significant differences appeared in the second study. Schrank concluded that the teachers in the first study had been influenced to give the students in the higher ability groups better grades.

In a similar study Flowers (1966) employed fictitious ability groups to learn about the effects of pupil performance on teaching expectancy. Using two schools, the matched pairs of classes were arbitrarily labeled as a top group in the school and the other as a low group in the school. Compared to the teachers of the control group the teachers in the supposed superior groups:

1. Referred more often to what the children could do rather than what they could not do.

2. Found virtually no discipline problems in the class although discipline problems were reported by almost all teachers of the control group children.

3. Referred more often to efforts to motivate their pupils and less often to inadequacy of teaching materials.

4. Preferred teaching the "higher" ability group.

Although Flowers, using matched pairs, created some statistical
problems, his four observations have merit. In an actual school setting with five ability groups, a questionnaire was presented to the faculty. On the question of preference for ability groups, 25 percent of those teachers in lower ability groups preferred to teach upper ability groups (Evaluation Report, Vol. 1, 1968). The remaining teachers were satisfied with their present teaching assignments.

In another study Beez (1970) matched sixty graduate students with sixty preschool students in a one to one tutoring situation. The "teachers" (graduate students) received biased reports on each preschool student prior to their teaching a set of symbols to their charges. The report concludes that the "high ability" group (students with positively biased reports) learned 10.43 words while the "low ability" group learned 5.66 words. His statistics support his hypothesis that the teacher is influenced and student achievement is affected by biased information regardless of the actual ability of the students to the .001 level; however, he does not describe the statistical tests he used.

The Beez study attempts to generalize from a one-to-one tutoring situation to a relationship between achievement and ability grouping. Although the study seems valid for a tutoring situation, he generalizes beyond the bounds of his study.

investigated teacher behaviors in teachers' classes who had two
different ability groups. She analyzed teacher-pupil verbal interaction
and teacher cognitive goals in the eleventh grade English classes of
a large urban high school.

Six successive class sessions of ten English classes
representing two different ability levels for each of the
five teachers in the study were observed using Flanders' Interaction Analysis technique. The teachers were inter-
viewed to obtain estimates of time emphasis on cognitive
goals, teacher talk, and student talk in the class. They
were asked about differentiation of their teaching and
their preference of classes: enriched classes (for pupils
having IQ's of 130 and above), adjusted classes (for pupils
who were two grades below grade level on English achieve-
ment tests and received below average English marks), or
average classes (for the remaining pupils in the eleventh
grade). Two for each class were analyzed according to the
cognitive goals inferred from the test items, using the six
major categories of Bloom's Taxonomy of Educational Objectives: cognitive domain (1956), knowledge, comprehen-
sion, application, analysis, synthesis, and evaluation.

Pfeifer (1966 and 1967) found:

1. Teachers did not differentiate their patterns of
teacher-pupil verbal interaction in classes of
different levels of ability. Verbal interaction was
significantly similar in classes of different ability
levels taught by the same teacher.

2. Four of the five teachers utilized more than half of
the class time for their own talk; they used direct
influence (lecture, criticism, directions) more
than indirect influence (use of student ideas, praise,
acceptance of student feelings, questions). Teachers'
emphasis on content was greater in their classes of
higher ability.

3. Student-initiated talk in a class generally was greater
when the teacher followed such talk with indirect verbal influence. The lower ability class usually did not initiate as great a percentage of ideas as the other group.

4. Most of these teachers had a preference for teaching classes of certain ability level or combinations of classes of different ability levels. Three of the four teachers indicating a preference used more indirect influence with the class preferred. This finding has potential value in assignment of classes since research by Flanders and his associates (5) and Furst (7) has indicated a positive relationship between superior student achievement and indirect teacher influence.

5. Teachers' estimates of time spent on cognitive goals indicated a differentiation of goals between average and adjusted classes. The emphasis for adjusted classes was on less complex skills: knowledge, comprehension, and application. Goals of average classes included the higher cognitive processes: analysis, synthesis, and evaluation.

6. Teachers' test items revealed a differentiation of operational cognitive goals for their classes of different ability levels. There was more concern with the less complex mental skills; the more complex skills received little attention.

7. Teacher estimates of time spent on cognitive goals were not significantly related to goals inferred from tests given to that class. Teachers indicated one pattern of cognitive goals was important and tested for another. See Table 2.

8. Teachers in this study did not realistically classify classroom talk. They especially misjudged how much they had talked. They underestimated student initiated talk of adjusted classes.

9. Three of the four teachers teaching adjusted classes believed little progress could be expected from this lower group. The teachers indicated that these expectations were reflected in limited assignments and restricted goals.
Pfeifer did not find significant differences in the verbal interaction of different ability levels taught by the same teacher using Flanders' Interaction Analysis. However, a careful analysis of the study procedures creates some basic questions regarding design, statistical analysis, and generalizability. The Pfeifer study is discussed in depth because of its similarity to the study to be proposed. Problems in design are analyzed first.

Pfeifer eliminated the lower ability groups from the study, using instead the upper and middle ability groups. The fact that Pfeifer did not find significant differences between ability groups could be linked to the lack of "extremes" in the sample population. In the pilot study discussed in the previous section (Freiberg, 1971) few significant differences were found between the upper and middle ability groups. However, significant differences occurred between the upper and lower ability groups. The elimination of the lower ability group in Pfeifer's study was compounded by the use of inappropriate statistical analyses.

Pfeifer used the Spearman Rank Order to analyze the Flanders' data. The use of the Spearman Rank Order analysis ignored the statistical problems inherent in closed observational system in general, and specifically with Flanders' Interaction Analysis. (See page 11 for a comprehensive discussion on the problem.) The Spearman Rank Order does not deal with didactic variables in closed observational systems and
therefore, seem inappropriate.

Pfeifer attempts to legitimize in her dissertation the small sampling population (Teacher N=5) by referring to previous research by:

a. Anderson and Brener (1965) who included five teachers in their study.

b. Mitzel and Robwinowitz (1963) who used four teachers in their study.

c. Hughes whose conclusions were based on the data from four teachers (Medley and Mitzel, 1963)

d. Withall (1963) who involved three teachers in a major project.

Perhaps Pfeifer is justified in using a small sample as the other researchers did, but she should realize that a small sample affects the generalizability of a study. To use the limitations of authorities as a justification for her own study dilutes the results of the study.

A research summary presented by the National Education Association (1968) on 158 studies of ability grouping reports that:

Although there is diversity in evaluation, opinion, and practice with regard to ability grouping, there appear to be three major areas of agreement.

1. Ability grouping has yet to prove itself as an administrative device to meet both effectively and efficiently the individual needs of all pupils in most areas of educational concern.

2. More and better research is needed to measure or control a larger number of the variables involved in ability grouping.

3. Objectives, materials, curriculum, and teaching
methods should also change when instructing groups at different ability levels.

From the evidence gathered, it appears that ample opportunity for flexibility in grouping is among the most important guiding principles... The skilled teacher is still the crucial factor.

The above study (N.E.A. et al) recognizes the teacher as a crucial factor in studies on ability grouping; however, this variable is omitted from most of the research on ability grouping. In this proposed study the teacher and her interaction with different ability groups are the focal point around which this study revolves.
Summary

Most of the research in the area of ability grouping is quite extensive but the studies give few significant results and those reporting significant results often defy replication by other researchers. The conditions and criteria for the term ability grouping vary with each study. The researchers have not attempted to standardize their terms or set realistic boundaries regarding the generalizability of their studies.

The area of ability grouping presents numerous facets and complexities. The studies on self-fulfilling prophecies (Rosenthal and Jacobson, 1969), tracking, teacher expectations (Flowers, 1966 and Beez, 1970), ability grouping, achievement (Borg, 1965 and N.E.A. Report, 1968), and teacher-student interaction on ability levels (Pfeifer, 1966) represent only a small part of the total problem of ability grouping.

Utilizing the results and recommendations of the above studies, the proposed study will focus on: (1) flexibility of ability grouping, and (2) verbal interaction patterns of teachers with classes grouped on the dimension of ability. These two areas are neglected in most of the research on ability grouping, although these variables are crucial to any discussion of the subject.
Hypotheses -- Introduction

Many discussion areas are not explored in this paper due to limited resources and the need to define a researchable problem. That there is a scarcity of resources is an economic fact, and a decision must be made on the allocation of available resources to priority areas. The selection of these four discussion points follows that economic principle.

The purpose of the pilot study was to test the feasibility of the proposed study and to generate hypotheses. From the pilot study and the literature in the field relating to ability grouping, four discussion points have been developed. These points are being presented here for further exploration and elaboration:

1. The effects of similar and different ability groups interacting with the same teacher under actual classroom conditions. Most educators concede that the actual classroom is the most important area of study in ability grouping. Yet, many of the studies (N.E.A. et al) look at problems peripheral to the classroom. These areas include the administrative and philosophical components of ability grouping. Also included in this area is the controlled experiment, which tends to generalize past the boundaries of the study (Beez et al).
Few studies actually observe the process of teacher-student interaction in ability grouped classrooms without some external experimental controls from the researcher. The data collected in this study will be obtained through the use of Flanders' Interaction Analysis. The effects of upper and lower ability groups interacting with the same teacher are observed under normal classroom conditions without external controls by the researcher.

2. The effects of ability grouping on secondary school students as it relates to direct or indirect teaching. A substantial quantity of information on ability grouping for the elementary grades is available, but the information on the secondary level remains scarce in comparison. Therefore, this proposed study focuses on the secondary classroom.

The direct teacher in the secondary classroom may inhibit the students' responses (Flanders, 1955). This pattern of direct teaching is observable in most secondary classrooms. Thus the change of this pattern to coincide with one particular ability group alerts the researcher to the subtle effects of ability grouping. The effects of ability grouping on direct or indirect teaching are observed as they affect the responses of secondary classroom teacher.
3. The flexibility or lack of flexibility of ability grouping as it is presently implemented. After reviewing the literature, particularly the findings and conclusions of the National Education Association Summary (1968), the author realized the need to explore the flexibility of ability grouping and the actual movement of students from one ability group to another. Since the problems of ability grouping are both complex and numerous, it would seem that solutions to such a problem would contain many facets. The flexibility of the student in ability groups appears to be an integral variable in the structure and implementation of ability grouping. The flexibility of students in the proposed study is calculated from the information attached to the teacher effectiveness questionnaire. (See Figure II) The information is used to set the generalizability of this study to other ability group situations.

4. The students' and teachers' perceptions of classroom interaction in relation to upper and lower ability groups.

If the ability group of a student affects the teacher-student interaction then a series of observational cross-checks must be utilized to examine the initial results. One
procedure examines the perceptions of the people involved in the experiment. In this case the teacher, student, and observer offer important information regarding the interaction occurring in the classroom. Both the students and teachers' perceptions are compared by giving both groups a teacher effectiveness questionnaire. The students rate their teachers' effectiveness and the teacher rates his own effectiveness with each of his classes. This process enables the researcher to examine the correlation between teacher effectiveness ratings and particular ability groups.

The four discussion points represent a direction which leads to four hypotheses presented below. The hypotheses are directional in nature and were generated from the results of the pilot study.

To avoid ambiguities regarding the hypotheses, the following definitions are offered:

**Ability Grouping** - "A practice wherein the total student population is divided into instructional groups according to some criteria of likeness" (Cohen, Ruth, Isenberg et al.).

**Phase** - The specific ability level to which a student is assigned while attending a class or classes for a definite amount of time.

**Track** - The student is permanently assigned to a particular phase level (e.g. Phase II) in most
of his subject areas and is unable to obtain entrance into another phase level during the school year.

**Heterogeneous Group** - Any group which is not divided into instructional groups according to some criteria of likeness.

**Homogeneous Group** - Used interchangeably in this study with the term ability grouping.

**Validity** - To measure what we think we are measuring (Kerlinger, 1964).

**Reliability** - The accuracy or precision of a measuring instrument (Kerlinger, 1964).
Hypotheses

The hypotheses are stated in specific terms with annotations and references made to the four discussion points in the preceding section.

I. Teachers are more direct, as measured by Flanders' Interaction Analysis, with lower ability groups and less direct with upper ability groups.

Categories Number 1, 2, 3, and 4 of Flanders' Interaction Analysis represent indirect teacher responses, and categories Numbers 5, 6, and 7, direct teacher responses (See Appendix A). Utilizing the Flanders I/D ratios, the indirect categories are divided by the direct categories (I/D sum of categories 1, 2, 3, 4 â by the sum of categories 5, 6, 7) to obtain the degree of indirect or direct teaching patterns for each teacher corresponding to his upper and lower ability group classes. The same procedure is followed with content categories Number 4 and 5 being removed from the calculations (I/D sum of categories 1, 2, 3 â by the sum of categories 5, 6). This process enables the researcher to calculate the degree of indirect or direct teaching patterns with teacher content removed from the results (Category Number 4 asking questions and category Number 5 lecturing). The hypothesis contends that teachers accept feelings (Number 1), give praise (Number 2),
and use student ideas (Number 3) a greater percentage of the time in the upper ability groups than in the lower ability groups. It also states the lower ability groups receive a greater percentage of directions (Number 6) and criticism (Number 7) than the upper ability groups.

II. Upper ability grouped students rate their teachers more favorably than lower ability grouped students.

This hypothesis contends students perceive a difference in the effectiveness of their teachers corresponding to their respective ability group. The hypothesis enables the researcher to compare student perceptions with the data collected from Flanders' Interaction Analysis and teacher perceptions. (See discussion point Number 4.)

III; Teachers perceive themselves to be more effective with their lower ability groups than the students' ratings of their teachers.

This hypothesis contends the teacher overestimates his effectiveness with the lower ability groups. This hypothesis enables the researcher to compare teacher perceptions with student perceptions on the teacher effectiveness questionnaire.

IV. Upper ability grouped and lower ability grouped students have 50 percent or more of their other classes in the same ability group.
This hypothesis contends the students are tracked into one particular ability group without either upward or downward mobility. The need for flexibility has been discussed in previous sections of this paper. The effect of inflexibility in ability grouping has been recognized by researchers as an important variable in the success of ability grouping. (See discussion point Number 3.)

Each of the above hypotheses analyze a particular segment of ability grouping. Together, they examine the effects of teacher-student verbal interaction of similar and different ability groups in secondary classrooms.
CHAPTER III
PROCEDURES

The Location

The study was conducted at a regional junior-senior high school (grades 7-12), located in western Massachusetts. The school enrolls around 1,000 students who are required to take English, social studies, mathematics, and science each year. The students are placed in one of the four ability groups for these subjects as they enter the seventh grade from their elementary schools. The placement into these four ability groups, or as the school terms them, phases, depends upon the student's score on group administered Otis Lennon and Stanford achievement standardized IQ tests, previous semester's grade in the subject and the teacher's evaluation of the student's present ability and future potential. The ability grouping or phases begin with the "upper" students denoted by a IV after the student's grade level (e.g., 8-4, eighth grade, phase four). The "average" student is placed in Phase III (e.g., 8-3, eighth grade, phase three), the lower ability group is represented by Phase II (e.g., 8-2, eighth grade, phase two), and the special education student is assigned to Phase I (e.g., 8-1, eighth grade, phase one). This study concerns itself with the students in Phase IV and Phase II.
Following a student's first semester at the school, his ability level is assigned according to his previous semester grades and the teacher's written evaluation. According to the school administration, the student could possibly be assigned to more than two different ability levels depending on the subject area. Once he is placed at a particular level, he will remain in that class for the duration of the semester. Each school year has two semesters.

A teacher may request a low ability group assignment, but generally low ability groups are assigned to teachers with more than two years teaching experience. The administration feels that experienced teachers are better prepared to cope with lower ability students who tend to present discipline problems to many of the new teachers.

The daily schedule is divided into fifteen twenty-minute modular units or modules. This schedule is designed to give the student one module of social studies on Monday and three modules of social studies on Tuesday.

The school presents a clean, modern environment for the student and faculty. The student population is over 99 percent white, representing the farming and small town communities within a thirty-mile radius of the school. Over 85 percent of the faculty live within the regional boundaries of the school, and, in many instances, have lived in the area most of their lives.
Subjects

Six teachers are used in this study, with teaching experiences ranging from one year to twenty-five years. The average experience of the six teachers is 9.5 years. The teachers' class size varies as noted in the table on page 48.

The student population in this study consists of 135 females and 124 males. Of the 259 students in the above six subject areas, 107 are in low ability groups and 152 in upper ability groups.

Methodology

The six teachers' upper and lower ability groups were observed for an entire week using an audio tape recorder to register the verbal interaction in the classroom. The tapes were analyzed by trained raters using Flanders Interaction Analysis ten category system (Flanders, 1965). A sixteen question teacher effectiveness questionnaire was given to the teachers to evaluate their perceptions regarding their effectiveness with their two ability-grouped classes. The same questionnaire was given to the student to evaluate their teachers. The questionnaire and classroom interaction protocols have been statistically analyzed and compared to observe the effect of ability grouping on verbal interaction in the classroom.
Table V
Classroom Statistics

<table>
<thead>
<tr>
<th>Teacher</th>
<th>Sex</th>
<th>Grade Level</th>
<th>Phases</th>
<th>Subject Area</th>
<th>Students per class</th>
<th>Total number of students in both phases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>(4 and 2)</td>
<td></td>
<td>Phase 4</td>
<td>Phase 2</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>10th</td>
<td>Social Studies</td>
<td></td>
<td>12 M 11 F</td>
<td>09 M 08 F</td>
</tr>
<tr>
<td>1*</td>
<td>M</td>
<td>10th</td>
<td>Social Studies</td>
<td></td>
<td>12 M 11 F</td>
<td>09 M 08 F</td>
</tr>
<tr>
<td>2</td>
<td>F</td>
<td>7th</td>
<td>Social Studies</td>
<td></td>
<td>17 M 21 F</td>
<td>13 M 11 F</td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>7th</td>
<td>Mathematics</td>
<td></td>
<td>14 M 14 F</td>
<td>18 M 06 F</td>
</tr>
<tr>
<td>4</td>
<td>M</td>
<td>9th</td>
<td>Science</td>
<td></td>
<td>03 M 10 F</td>
<td>05 M 07 F</td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>11th</td>
<td>English</td>
<td></td>
<td>02 M 17 F</td>
<td>06 M 08 F</td>
</tr>
<tr>
<td>6</td>
<td>M</td>
<td>11th</td>
<td>Social Studies</td>
<td></td>
<td>16 M 14 F</td>
<td>09 M 08 F</td>
</tr>
</tbody>
</table>

N=6
N=259

*Note--The number beneath Teacher (1-6) refers to the alphabetic listing of the last name.
The limitation of the study to those teachers with upper and lower ability groups is supported by the authors' analysis of the Pfeifer study (1966) and the pilot study (1970) developed in the previous section of this study. Six teachers met the criteria of teaching both an upper and lower ability group. All six teachers agreed to participate in this proposed study. Since only six teachers, three males and three females, teach both upper and lower ability groups, a total population of teachers in one school was studied. The combined junior-senior high school kept the criteria of selection of students for ability grouping constant for the entire population.

<table>
<thead>
<tr>
<th>Teacher</th>
<th>Phase IV class</th>
<th>Phase II class</th>
<th>Total Observations per teacher</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(4) 20-min. tapes</td>
<td>(4) 20-min. tapes</td>
<td>(8) 20-min. tapes</td>
</tr>
<tr>
<td>2</td>
<td>''</td>
<td>''</td>
<td>''</td>
</tr>
<tr>
<td>3</td>
<td>''</td>
<td>''</td>
<td>''</td>
</tr>
<tr>
<td>4</td>
<td>''</td>
<td>''</td>
<td>''</td>
</tr>
<tr>
<td>5</td>
<td>''</td>
<td>''</td>
<td>''</td>
</tr>
<tr>
<td>6</td>
<td>''</td>
<td>''</td>
<td>''</td>
</tr>
</tbody>
</table>

N = 6 Total: (24) 20-min. tapes  (24) 20-min. tapes  (48) 20-min. tapes
To keep the content of the lessons of each teacher somewhat equal, classes were observed at the beginning of a new lesson. The forty-eight classes for the six teachers were observed within a two-week period to reduce the effect the variable of time may have. A total of eight 20-minute classes per teacher were observed; four upper ability group and four lower ability group for each of the six teachers.

The teachers and students only knew that research of classroom teaching was being conducted, and the exact nature of the study was not divulged until all analyses had been completed. This procedure was followed to avoid any reactive effects prior knowledge could have. Due to the modular scheduling system and other innovations utilized by the school, the students and teachers have become accustomed to observers and recording equipment in the classroom.

Rosenshine and Furst (1970), in their review of educational research, found both low-inference and high-inference measures absent in most studies. They felt "the use of high-inference and low-inference measures in future studies may be most advantageous." A low-inference measure is one which two or more observers may readily agree upon. An example might be category number two (praise) in the Flanders system. Praise given to a student by a teacher would receive a higher percentage of agreement by two observers than the teacher behavior of warmth. The observation of warmth would represent a high-inference measure.
Rosenshine and Furst (1970) concludes by stating:

One procedure for combining the two observational procedures would be to use student questionnaires to describe high-inference behaviors, and outside observers and recordings to describe the low-inference behaviors.

The use of Flanders Interaction Analysis and the student questionnaire on perceived teacher effectiveness incorporate both the low-inference and high-inference measures recommended by Rosenshine and Furst (1970).
Summary

Only verbal interaction was recorded. Non-verbal interaction in the classroom was not recorded, and this observational parameter should be noted as the reader develops conclusions from the study. Flanders' Interaction Analysis was utilized to quantify the verbal interaction and to determine any similarity or differences in the teachers' reactions to students in different ability groups. The teacher effectiveness questionnaire was used by the students and their teachers to determine if similarities and differences exist between the teacher and his students' perceptions of the teacher. This instrument is also used to determine if any differences exist regarding the students' perception of their teacher between different ability groups.

Analysis of Data

Statistical analysis on the Flanders categories must reflect past statements on analysis of observational system in general, specifically with Flanders' Interaction Analysis (See page 11). Each teacher's upper and lower ability group receives frequency distributions graphed for the categories discussed in the section of hypotheses. General observations based on the frequency distributions are presented in the section on the results and findings of this study.
Teacher Effectiveness Questionnaire

Reliability. The teacher effectiveness questionnaire presented to the students \( N = 259 \) and their teacher \( (N = 6) \) describes behaviors of teachers which are to be rated along an open ended scale (See Figure II p. 61). The scale was originally used by Theodore Cromack (1971) in his study on teacher effectiveness. His scale focused on twelve items which received a reliability rating of \( r > .30; p < .01 \). An additional four items have been added to his scale reflecting the recommendations of Rosenshine and Furst (1971), N.E.A. report (1968), and Wightman (1971), who found fairness, warmth, flexibility and enthusiasm to have high correlation with teacher effectiveness. The addition of the four items requires another reliability test using the Hoyt Reliability Test which was significant at the \( r = .929 \) level.

Analysis. The teacher effectiveness questionnaire (See Figure II) data was analyzed by the Statistical Package For The Social Sciences (Nie; 1971) for all appropriate statistical analyses. The SPSS performs eleven basic statistical analyses which was utilized to observe possible interaction of variables on the teacher effectiveness questionnaire. The sixteen questions on the questionnaire were statistically compared (Rank Biseral Correlation) with the six teachers and their two phase levels (IV and II); these questions were also compared with the grade
the student expects to receive by the end of the year, the student's
grade level, the time of day, and with the response of the students by
the two phases by the teacher's response for those two phases.

Although student perceptions of teacher effectiveness is a key
variable in this study, the other variables listed below must be analy-
zed to observe their possible significance in producing any observed
results. Comparisons will be made in the following areas:

1. Each of the 6 teachers and their phases with each of
the 16 concepts in the questionnaire.

2. Total phases (4 and 2) by each of the 16 concepts on the
questionnaire.

3. The expected grade point average the student expects
to receive by each of the 16 concepts.

4. Phase in which student received questionnaire by his
other phase levels.

5. Student grade level by each of the 16 concepts in the
questionnaire.

6. Phase by mod (time of day) by the 16 concepts.

7. The students' response for each of the 16 concepts by
the teacher's response for each of the 16 concepts.

8. Total teacher response for the 16 concepts by the total
students' response for the 16 concepts.

9. The students' response for each phase (4 and 2) for the
16 concepts by the teacher's response for each phase for the
16 concepts.

The above comparisons were made to determine the patterns of
response by the student and their teachers and to isolate those significant variables (at .01 level) which support or refute the hypotheses.
Training and Reliability of Observers

An integral part of this study utilizes forty-eight 20-minute tapes of six teachers' classes. These tapes are rated by four observers using the Flanders Interaction Analysis ten category system. The raters receive approximately thirty hours of training in the use of Flanders' system. Near the conclusion of their training period they are presented with a prerated tape (from the Flanders training packet) which they all rate. Their ratings are analyzed by Lawrence Wightman's Program Interact (Wightman 1971) and are subjected to the Scott Reliability Test (Gregory, 1969). A rating of .80 or better on the Scott Reliability Test gives sufficient confidence to their inter-reliability using the Flanders' system. The raters are randomly assigned eight tapes out of the forty-eight tapes to code for use in analyzing the interaction in the classroom. The four raters received a Scott Reliability rating of .899 on the Flanders Interaction Analysis tapes.

Validity of Observations

The use of observational tools presents the researcher with some unique perceptual problems, specifically from the viewpoint of the raters and students' interpretation of the teachers' behaviors. A single message may be communicated (verbally) to
a group, but each person may decode the signal in a different manner. This "human" factor may alter the reporting of an actual behavior in the classroom. Raters are tested for reliability to assure some agreement between the different raters, but the differences between the raters and the students perceptions of what is occurring in the classroom is very difficult to measure.

To gain some feel for the presence or absence of validity between what the raters perceive as praise (e.g. Category number 2 in the Flanders system) and what the student perceives as praise, the following procedure is used:

The researcher questions students participating in two recorded mathematics classes and randomly selects two ten-minute segments of tapes to be rated by the trained raters and two groups of four students each. These four students are randomly selected from each of the mathematics classes. The students listen to the tapes and explain what the teacher is saying to the class and how they interpret the teacher's words. For example, if the Flanders rater records a category 3 (use of student idea), the students listening to the same tape should also perceive the teacher's use of a student idea.
This process enables the researcher to observe any similarities or differences in the responses of the two groups of students and to judge these responses with those of the raters' who used Flanders Interaction Analysis system. A simple frequency comparison between the group from the mathematics class and the Flanders raters is made to achieve validity in observations (See Figure I for validity format.) The students and Flanders Interaction Analysis raters achieved an 80 percent level of agreement using this validity format.

Summary

Reliability studies have been made on: (1) Flanders Interaction Analysis raters using the Scott Reliability Test, and (2) the teacher effectiveness questionnaire using the Hoyt Reliability Test. A validity study was also made on the agreement between the Flanders raters and students perceptions of the same teacher's behavior.

This study represents one step in a long staircase. Numerous problems exist regarding the grouping of students according to ability which this study will not touch. The concepts of the labeling effect, self-fulfilling prophecy, tracking, curriculum, and pupil achievement need to be studied in the light of classroom interaction, and hopefully some of the questions regarding this interaction will be answered by this proposed study.
Since there are certain limitations to this study, the researcher may only generalize on specific research groups which follow the criteria developed for this study. However, it is felt this study leaves the way open for other studies to replicate and expand the sample population.
Figure I
Observation Validity Sheet

<table>
<thead>
<tr>
<th>Number of Students and Classes</th>
<th>Number of Topics Noted</th>
<th>Number of Topics Agreed with %</th>
<th>Number of Topics Disagreed with %</th>
</tr>
</thead>
<tbody>
<tr>
<td>8th Grade Math</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group I (4 students)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8th Grade Math</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group II (4 students)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10th Grade Geometry</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group I (4 students)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10th Grade Geometry</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group II (4 students)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure II

Teacher Effectiveness Questionnaire

Teacher: __________________________  Expected Final Grade: ________  Grade Level: ________

Subject: __________________________  Phase: ________  Please check ( ) Male; ( ) Female

Please list your other subjects and phase levels.
1. __________________________________________  2. __________________________________________
3. __________________________________________  4. __________________________________________

Instructions: The following lines represent traits commonly noted by students when describing their teachers. Please place a check mark (√) on that part of the line which would indicate how you would rate your teacher. Each line should be checked.

1. Good Organization  Poor
2. Thorough Preparation  Unprepared
3. Current Subject matter knowledge  Limited
4. Interesting Presentation  Dull
5. Open minded Attitude  Biased
6. Good Sense of Humor  No
7. Interesting Personality  Poor
8. Encourages Discussion  Prohibits
9. Effective Speaker  Boring
10. Respects Student  Belittles
11. Recognizes Student's Needs  Ignores
12. Clear What is Expected of Student  Unclear
13. Fair Fairness  Unfair
14. Very Warmth  Not
15. Very Flexible  Not
16. Very Enthusiastic  Not
CHAPTER IV
FINDINGS

Hypothesis I states: Teachers will be more direct, as measured by Flanders Interaction Analysis, with lower ability groups and less direct with upper ability groups. The indirect/direct (I/D) teacher verbal interaction was measured using Flanders Interaction Analysis (Flanders, 1965) ratio statistics, which are expressed in terms of the sum of the observations of the first four Flanders categories divided by the sum of the observations of categories five through seven. This procedure enables the researcher to obtain a general view of the teachers' indirectness or directness in the classroom. According to Flanders and others who have used the system, the indirect teacher allows greater verbal student interaction in the classroom than the direct teacher.

The above procedure was followed to observe the difference between the verbal interaction of the same teacher with the combined I/D ratio scores of four upper and lower ability grouped classes. The results graphically demonstrate (See Figure III) that for each teacher, the upper ability groups receive from 1 percent to 28 percent more indirect verbal interaction than the lower ability groups. Observing the entire sample of six teachers and their forty-eight ability grouped classes, the upper ability grouped students received $\times$ 16 percent more
Flanders Interaction Analysis I/D Ratios  Figure III

These percents are Indirect/Direct (I/D) ratios expressed as \((I/(I+D)) \times 100\)
*These are Revised I/D Ratios Expressed as \((I/(I\;D))100\)
### Flanders Interaction Analysis - Category Three - Use of Student Ideas Figure V

<table>
<thead>
<tr>
<th>T_1</th>
<th>T_2</th>
<th>T_3</th>
<th>T_4</th>
<th>T_5</th>
<th>T_6</th>
<th>(T_1+...+T_6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hi Lo</td>
<td>Hi Lo</td>
<td>Hi Lo</td>
<td>Hi Lo</td>
<td>Hi Lo</td>
<td>Hi Lo</td>
<td>Hi Lo</td>
</tr>
</tbody>
</table>

- **Hi** - Upper Ability Group
- **Lo** - Low Ability Group
- **T** - Teacher
- * Percentage of tallies from the total ten Flanders Categories
indirect teacher interaction than the lower ability groups (See Figure III). The I/D ratios represent an aggregate picture of teacher responses within the Flanders system. However, content categories (numbers 4 and 5) may have weighted the results. A second analysis using a revised Flanders I/D ratio, eliminating the content categories is graphically presented (See Figure IV). The revised I/D ratios indicate that the upper ability groups receive 7 percent to 57 percent more indirect verbal teacher responses than the lower ability groups. (See Appendix A for calculation of I/D ratios.) Analyzing the entire sample using the revised I/D ratios, the upper ability groups receive 35 percent more indirect verbal teacher responses than the lower ability groups.

The I/D and revised I/D ratios analyze seven Flanders categories which examine verbal teacher interaction in the classroom. Category number 3 (use of student ideas) in particular has been expressed by Flanders (1960) to be important in developing student freedom in the classroom. Category number 3 has been graphically analyzed (See Figure V) to observe the similarities or differences between the teachers' use of student ideas in the upper and lower ability groups. The upper groups received .7 percent to 6.8 percent more use of student ideas than the lower ability groups. The upper ability groups received 2.7 percent more use of student ideas than the
lower ability groups. The increase in use of students' ideas by teachers of upper ability groups was not due to a higher percentage of upper ability group student initiated responses. The student initiated questions (category 9 on the Flanders scale) received equal distribution for both upper and lower ability groups.

Hypothesis II states: Upper ability grouped students rate their teachers more favorably than the lower ability grouped students. Each student in the six teachers' classes responded to a sixteen question Teacher Effectiveness Questionnaire (T.E.Q.) The responses for each of the sixteen items on the teacher effectiveness questionnaire from one (effective) to seven (least effective). The individual ratings for each of the sixteen questions were totaled giving a composite score for every student's responses on the teacher effectiveness questionnaire. The total scores were deemed to represent the students' perceptions of their teachers' effectiveness. The lower scores represented more effective teaching than the higher scores on the Teacher Effectiveness Questionnaire. The teachers received ratings from 16 (effective) to 114 (not effective) The upper and lower ability grouped student scores were graphically compared (See Figure VI) to observe the effect ability grouping has on the student effectiveness ratings of their teachers. The number of students ratings their teachers from effective to not effective were plotted from the upper and lower ability groups. The
Teacher-Student Responses On The Teacher Effectiveness Questionnaire Figure VI

Less Effective

Teacher Effectiveness Scale

Effective

T - Teacher Self-Evaluation  S - \bar{x} Student Evaluation of Teacher Effectiveness
graph indicates the upper ability group students rated their teachers more effective than the lower ability group students.

A mean of the students' responses to the teacher effectiveness questionnaire for the upper and lower ability groups was computed and compared with the teachers' own responses to the teacher effectiveness questionnaire (See Figure VI).

Hypothesis III states: Teacher perceive themselves to be more effective with their lower ability groups than the students' ratings of their teachers. The teachers utilized the teacher effectiveness questionnaire to rate their own effectiveness with their classes, which were later divided into the appropriate ability groups. The teacher self-evaluation scores were plotted on the student-response teacher effectiveness graph to develop a comparative picture of student-teacher perceptions of teacher effectiveness. The lower ability group students found their teachers less effective than the teachers' ratings of their own effectiveness with lower ability group students. The upper ability grouped students' (See Figure VI) teacher-effectiveness ratings correspond more closely with the self-evaluation ratings of their teachers.

Hypothesis IV states: Upper ability grouped and lower ability grouped students have 50 percent or more of their other classes in the same ability group. The students listed all their other classes and ability levels on the teacher effectiveness questionnaire. The ability
## Student Flexibility in Ability Grouped Sample Population Figure VII

Phase in which the Student Received the Teacher Effectiveness Questionnaire

<table>
<thead>
<tr>
<th>Phase and Subject Areas</th>
<th>Low Ability Students Other Phase and Subject Areas</th>
<th>High Ability Students Other Phase and Subject Areas</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low Phase</td>
<td>Middle Phase</td>
</tr>
<tr>
<td>Science</td>
<td>31.0</td>
<td>04.0</td>
</tr>
<tr>
<td>English</td>
<td>24.0</td>
<td>09.0</td>
</tr>
<tr>
<td>Math</td>
<td>68.0</td>
<td>01.0</td>
</tr>
<tr>
<td>Social Studies</td>
<td>128</td>
<td>29.0</td>
</tr>
<tr>
<td>Other (e.g. Art and Music)</td>
<td>06.0</td>
<td></td>
</tr>
</tbody>
</table>

Percentage of Students enrolled in various Phases

- Low Ability Students: 86% 14% 00%
- High Ability Students: 01% 26% 73%
group in which the students received the teacher effectiveness questionnaire was compared with the student's other subject and ability areas. The students who received the questionnaire in a lower ability group had 86 percent of their other classes in a low ability group; 14 percent had middle ability group classes while 0 percent were in upper ability group classes. The students who received the questionnaire in an upper ability group had 73 percent of their other classes in the upper ability groups with 26 percent in the middle ability group and 1 percent in low ability group classes.

The following variables were analyzed to observe extraneous interaction effects and to produce graphic teacher-student patterns regarding the teacher effectiveness questionnaire. The expected grade point average, grade level, and time of day were analyzed using the Pearson product moment correlation to observe the possible interaction effects with the teacher effectiveness questionnaire. The remaining variables listed below were analyzed to develop graphic information regarding Hypotheses II, III, and IV.

1. Each of the six teachers and their phases with each of the 16 concepts in the questionnaire.

2. Total phases (4 and 2) by each of the 16 concepts on the questionnaire.

3. The expected grade point average the student expects to receive by each of the 16 concepts.

4. Phase in which student received questionnaire by his other phase levels.
5. Student grade level by each of the 16 concepts in the questionnaire.

6. Phase by mod (Time of day) by the 16 concepts.

7. The students' response for each of the 16 concepts by the teacher's response for each of the 16 concepts.

8. Total teacher response for the 16 concepts by the total students' response for the 16 concepts.

9. The students' response for each phase (4 and 2) for the 16 concepts by the teacher's response for each phase for the 16 concepts.

The first two variables were tested for internal consistency within the teacher effectiveness questionnaire. All the correlations were positive with significance at the .01 level. The expected grade point average and the grade level of the students received no significant correlations at the .01 level of significance. Variable number four was utilized to obtain the other subject and phase levels of students who received the questionnaire in either an upper or lower ability grouped class. The time of day did not receive significant correlations at the .01 level. Variables number seven, eight and nine were utilized in Figure VI to develop a graphic picture of the teachers' and students' responses to the T.E.Q.

The first two variables were tested for internal consistency within the teacher effectiveness questionnaire using the Pearson product-moment correlations coefficient. All correlations were positive with significance at the .01 level. The expected grade point average developed
no significant correlations at the .01 level. The expected grade point average assigned by the students were in close agreement with the grades the teachers expected to give their students. Variable number four was utilized to obtain the other subject and phase levels of students who received the teacher effectiveness questionnaire in either an upper or lower ability grouped class. The time of day did not reach significant differences using Chi-Square statistical analysis. Variables seven through nine were utilized in Figure VI to develop a graphic analysis of the ability groups responses to their teacher's effectiveness and their teacher's self-evaluation of their own effectiveness.

**Summary of Findings**

1) Teachers were more direct (as measured by Flanders Interaction Analysis) with lower ability groups and less direct with upper ability groups; 2) The teachers appeared to be even more direct with the lower ability groups with the content categories removed; 3) Teachers used a greater percentage of the upper ability groups' ideas in the classroom as compared to the ideas of the lower ability groups; 4) The lower ability groups rated their teachers less effective than the upper ability groups; 5) Teachers perceived themselves to be more effective with the lower ability groups than the student ratings of those same teachers; 6) A student in a lower ability group class would have 86 percent of his other
classes in the lower ability group; 7) A student in the upper ability group would have 73 percent of his classes in upper ability groups.

Analysis of Findings

The findings of the first hypothesis emphasizes the importance of the teacher's verbal interaction with students in the classroom. Using Flanders I/D ratios the researcher was able to observe (See Figure III) greater use of indirect verbal interaction with the upper ability groups and direct verbal interaction with the lower ability groups by the same teacher. With two of the six teachers the differences were small, but the remaining four teachers showed observably greater differences.

The elimination of the content categories (numbers 4 and 5) on the Flanders scale produced a greater observed effect (See Figure IV) with regard to the teacher-student verbal interaction. The revised I/D ratios emphasized the teachers use of empathy, praise, and use of student ideas relative to teacher directions and criticism. The lower ability group students received less verbal empathy, praise, and use of their ideas than upper ability group students with identical teachers.

The isolation and analysis of Flanders category three, use of student ideas, (See Figure V) followed the pattern developed with the above two analyses. The upper ability groups received a greater percentage of teacher use of student ideas than their corresponding lower ability group
peers with the same teacher.

The findings of this study are in direct contrast with a similar study conducted by Pfeifer (1966) who found: "Teachers did not differentiate their patterns of teacher-pupil verbal interaction in classes of different ability levels taught by the same teacher" (See page 33 regarding discussion on statistical problems and the elimination of the lower ability groups in the Pfeifer study). The analysis of the first hypothesis gives support to the contention that teacher-student interaction is effected by the ability group of the student with which he or she is identified. These findings in conjunction with the following analysis will further support the first hypothesis.

The teacher effectiveness questionnaire was utilized to observe the differences in student perceptions of their teacher's effectiveness relative to the students' ability group. The lower ability group students perceived their teachers to be less effective than did the upper ability group students. The assumption that lower ability group students would rate their teacher less effective is based on the observation that the same teacher gives the lower ability group students more direct verbal teacher responses while giving the upper ability group students more indirect verbal responses. Holding the teacher variable constant the researcher is able to observe the response of two different ability level students to the same teacher. The graphic representation (See Figure VI) of the
lower ability group ratings of their teachers' effectiveness seems to support the assumption that most lower ability group students (in this study) are rating their teachers less effective. This is possibly due to the type of verbal responses from their teachers. The lack of significant correlations between the interacting variables (See page 55) gives additional support to these results regarding ability grouping.

The teachers' perceptions of their effectiveness with their two ability groups, as stated in Hypotheses III, focuses on the teacher as a crucial variable in this study. The apparent agreement the teacher and his upper ability classes is observably close (See Figure VI). However, the same teacher's perception of his effectiveness with the lower ability groups represents an observably distinct (See Figure VI) difference. Teachers apparently overestimated their teaching effectiveness with their lower ability students. Lower ability grouped students scored their teachers to be less effective than the upper ability groups' ratings of the same six teachers. When one looks at the student ratings together with the teacher self-evaluation ratings one sees a distinct difference between the two ability groups and their perceptions of their teacher's effectiveness.

The need for students to move freely between ability groups has been stated previously as an important variable in the success or failure of an ability grouped program. Students who are locked into one particular ability group are limited in their possible contact with other students which
tends to limit the enrichment of their educational environment (N.E.A., 1968). The fact that 86 percent of those in one lower ability grouped class have their remaining subjects in other lower ability grouped classes emphasizes a definite lack of flexibility in this program. Those students in the upper ability groups had some flexibility, but 73 percent of their classes were in the same ability group. Hypothesis IV is important in defining the boundaries of this study in terms of its generalizability to other studies. The lack of flexibility in this particular ability grouped sample has possibly made this ability grouping program less effective. The results of this study would not necessarily find duplication in an ability grouping program where the students had free movement through the different ability levels.

**Summary**

The four hypotheses and their underlying assumptions are supported by the following: 1) Teacher differentiated their pattern of teacher-pupil verbal interaction in classes of different ability levels; 2) Indirect/Direct verbal interaction and use of student ideas were not significantly similar in classes of different ability levels taught by the same teacher; 3) The upper ability grouped students rated the same teacher more effective than the lower ability grouped students; 4) Teachers perceived themselves to be more effective with the lower ability groups than the students
perceived them; 5) Lower ability group students had 86 percent of their other classes in the lower ability group; 6) Upper ability group students had 73 and 86 percent of their classes respectively in the same ability group.
CHAPTER V
CONCLUSIONS

This study attempted to observe under actual classroom conditions the process of teacher-student verbal interaction as it related to different ability levels. This was accomplished by combining two observational techniques suggested by Rosenshine and Furst (1970). Student questionnaires (teacher effectiveness) were given to two hundred and fifty-nine secondary students to rate the effectiveness of their teachers. The same questionnaire was given to their teachers to evaluate their own effectiveness with their students. The questionnaires were used to develop a measure for high inference behaviors. Outside observers and recorders using Flanders Interaction Analysis techniques were utilized to describe the lower inference behaviors in the classroom.

The above approach was developed to observe the process of transferring the concept of ability grouping into a workable school program. The key variable in this transfer was the teacher and his role in the daily operations of the ability grouped secondary classroom. Three of the four hypotheses focused on the teacher and his verbal interaction with his students. The findings of this paper reemphasize the importance of the teacher in implementing educational programs.

The effects of ability grouping on secondary students as it relates to
direct or indirect teaching was explored to observe the development of indirect/direct teaching patterns along the lines of specific student ability levels. Although the issue of achievement and its relationship to indirect/direct teaching has not been settled, consistent direct teacher behavior toward a particular group of students appears to inhibit the verbal interaction of the students.

The results show that the teachers in this study are consistently more indirect with their upper ability group students. This factor leads the author to question the advisability of using a system which offers more indirect verbal teacher responses to the high ability groups and more direct teacher responses to the low ability group students. The elimination of ability grouping suggests one solution, but the question of grouping will continue to be resolved. Alternate approaches to grouping should be explored and analyzed before any final decisions are made regarding ability grouping.

Students who receive a greater percentage of direct verbal teacher responses seem to rate their teacher less effective than those students who receive greater indirect teacher responses. The training of indirect teachers for the lower ability groups would not be recommended as a solution for balancing the verbal interaction in these classrooms. The Indirect/Direct ratios are utilized to observe verbal teacher patterns.

Once the patterns have been established (as described in this study) the
directions are noted and analyzed to ascertain their effect on such variables as student perceptions of teacher effectiveness. The direction of those variables should be viewed in terms of possible differences with explanations offered to account for the variations. The results and findings of this study should be utilized to predict the possible occurrence of patterns and directions in future studies under similar situations.

The indirect/direct teacher responses seem to correspond to the high and low ability groups respectively. The lower ability group students rated their teachers less effective than the corresponding upper ability group students. The results of the teacher effectiveness questionnaire and the classroom observation support these arguments. Although there seems to be a correlation between ability grouping, indirect-direct teacher responses and student perceptions of teacher effectiveness, the relationship should be termed directional in nature.

Accurate information in many instances fosters the greatest change in teacher behaviors. The teachers will be offered the results of this study, stated in terms of pattern differentiation rather than value statements on the positive or negative aspects on the degree of criticism used in their classrooms. Adequate information is crucial to the decision making process which the teachers in this study should make regarding their teaching in general and specifically with the ability grouping structure they use. The teacher may decide to abandon the ability grouping programs and construct a system which allows students to develop their
maximum learning capabilities.

A conducive learning environment would require all of those involved (i.e. teachers, students, parents, school administrators) to be in close agreement with most of the stated educational goals and objectives of the learning environment, which requires assessment of the teacher's effectiveness in implementing those objectives. Presently, students are assessed on their ability to meet the teachers' objectives. Teacher-student perceptions regarding the effectiveness of the learning situation should reflect some commonality. Two groups of observers who consistently disagree on a given event reduces the validity of their findings. The validity of an assessment of a learning environment is questionable when the participants consistently disagree on their observations of teacher effectiveness.

The teacher effectiveness questionnaire was utilized in this study to bridge the possible gap between the perceptions of the students and their teachers. Teacher-student perceptions should be made available to assist the teacher and students in their decision making processes regarding their learning environments.

The crucial point of this discussion focuses on the teacher and his perceptions of his effectiveness with his ability grouped students. Teacher-student perceptions regarding teacher effectiveness in the lower ability groups did not correspond. These differences in perceptions indicate
that the effects of ability grouping as seen in the classroom are concealed from the teacher. The verbal interaction assessments combined with the teacher effectiveness questionnaires had led this author to conclude that ability grouping in this particular school has not aided the educational environment of the students.

The educational environment is further constrained by the lack of flexibility between ability groups for this sample. The National Education Association report (1968) underscores the importance of flexibility in the movement of students between different ability levels. The students in this particular study are apparently "tracked" into one particular ability level for all their major subject areas. It is inconceivable to think that 86 percent of the lower ability grouped students would perform at the same ability level for all their other subject areas. Is is equally disturbing to think that 73 percent of the upper ability grouped students would perform in a similar manner in their other subject areas. It has been noted in the research in this paper that students in homogeneous classes vary in their achievement and ability performances (Borg, et al). It has also been stated (N.E.A. et al.) that the situation described above is not unique in schools which group according to ability. Therefore, the generalizability of this study may be interpreted to include situations in many school systems.
Summary

1) The teacher is the crucial variable in the implementation and operation of ability grouping.

2) Teachers should receive empirical data on their effectiveness with students in different learning situations.

3) A number of learning situations and structures should be available for students and teachers to explore in order to obtain maximum learning-teaching effectiveness.

4) All participants should be involved in the development and assessment of stated educational goals and objectives.

5) Classroom variables should be measured to ascertain patterns and directions with the objective of changing those patterns which the teachers and students feel inhibit the learning environment.

6) Students should have the flexibility to be in a variety of learning environments which allow their free movement to achieve maximum learning effectiveness.
Recommendations

The following recommendations are directed toward the teachers involved in this study with broader implications to those teachers who see themselves in similar situations. Value judgements regarding ability grouping should not be made without a clear understanding of the interacting educational variables. Ability grouping is an effective tool when implemented by teachers or administrators who have an understanding of these variables. In this particular study the flexibility which was assigned a high priority by the school administration was absent during the implementation of the program. The structural defect in this program resulted from the lack of an effective evaluation of the model as it operated in the school.

The need for continuous evaluation is a constant problem in programs which attempt to make the educational environment conducive for learning. Although the research in the field emphasizes the need for continuous evaluation, most school systems continue to implement ability grouping without thorough planning and evaluation.

The teachers who are responsible for the operation of ability grouping programs have been given in most cases minimal or no training in the operation and evaluation of grouping in their classrooms. Given adequate information regarding the pitfalls of ability grouping, teachers may
operate more effectively. The information gathered in this report would undoubtedly surprise the participants who felt that they did not differentiate their verbal interaction with classes of different ability levels. Teachers should not be expected to modify their behavior without accurate information regarding their teaching effectiveness. A continuous teacher-student evaluation should be implemented particularly in situations where changes in the learning environment are occurring.

The examination of the role of the teacher in this study should offer a new perspective from which teacher expectations and student achievement could be explored. Although these areas were not explored in this study, the role of the teacher and his effect on the success of ability grouping should be considered in further studies. Replication would produce important evidence which could possibly expand the generalizability.

Replication of this study should include: 1) Exploration of the effects of ability grouping on non-verbal teacher-pupil interaction in the classroom. 2) A number of schools randomly selected from a national sample utilizing the variables discussed in this study. 3) The directional relationships presented in this study should be tested. 4) Compare an ability grouped model with a free-school model utilizing the variables discussed in this study.
The recommendations regarding this study represent one aspect for improving research in the field of ability grouping. The assumptions underlying the philosophy and implementation of grouping need to be reevaluated and analyzed in greater detail. Presently, we group students in many of our schools in terms of a product. We decide where a student should be placed according to his output in a learning situation. Grouping based on output places the responsibilities and burdens on the learner. In many instances teachers should be grouped with specific students. Although this suggestion has been expressed by numerous educators, its implementation in schools which utilize ability grouping is practically non-existent.

Given the information of this study, some teachers might agree that they should not have been teaching one particular group of students. The effectiveness of the teacher with particular ability grouped students should be considered prior to the formation of classroom assignments.

After the first six months every new teacher should be assessed according to his or her effectiveness with different ability groups. This would include student ratings of their teachers using a teacher effectiveness questionnaire similar to the one utilized in this study. Teachers would also evaluate their own perceived effectiveness with their classes. The teachers presently teaching in the system would receive a similar type of assessment twice a year. The classroom assignments for the
following year would be based on the evaluation of the previous year.
The data gathered from the teachers and the students would be compared, and decisions for realigning classes would be based on the assessment information. This procedure would enable teachers to utilize their skills with greater effectiveness.

The only criteria utilized by the school administration in this study was the amount of time the teachers have had in teaching. They gave to the more experienced teachers the lower ability groups. Many school systems give no thought to the needs of the teachers or students in ability grouping programs. Many urban systems give first year teachers the low ability groups. In some of these systems grouping is used to isolate the discipline problems. In other systems ability grouping is utilized to segregate the school by placing a disproportionate number of black students in the lower ability groups (Bannister, 1971). Without adequate evaluation and participation of teachers in the development of educational programs, problems of this nature will continue to plague our schools.

The concept of grouping grew out of the need to meet the individual abilities of students. This procedure focused on what the learner produced rather than the process used to develop the product. Part of the problem is determined by the educational system which finds it easier to assess a product rather than a process. We have just begun to explore learning areas in terms of process. It would seem appropriate to
explore the possibilities of grouping students according to their learning processes, and teachers by their ability in various teaching processes. If we learn by our senses, a logical educational system would utilize those senses in the teaching process. Presently, we teach through the auditory sense, and group students according to their ability to produce from auditory input. From this author's observations in classrooms using Flanders Interaction Analysis, teachers talk approximately 70 percent of the class time. Placing students who produce on a similar auditory level in the same classroom does not substantially change the students learning abilities in the other sensory areas.

The exploration of grouping according to the assessed learning processes of students is recommended. This would enable students to be grouped according to their innate learning abilities. This procedure would begin in the early grades with the general objective of obtaining learning competencies in most of the sensory areas. The process would require an entirely new approach to the training of teachers and the development of new curricula. Some work in this area has been developed in the multi-media centers in some innovative schools, but these centers usually supplement traditional classroom teaching.

Pilot studies should be developed to amplify and test areas of sensory learning. The ultimate goal would be met when an entire generation of teachers and students would learn and be taught through a total sensory
The entire process of education develops a new approach when viewed from the perspective of the teacher. Ability grouping represents a small segment of the educational picture in America, but in the classroom it may mean the difference between success and failure for an individual student. The foundations on which ability grouping is based represent a basic philosophical expression in our society. The work-ethnic in American society is evident in America's classrooms. This ethic has met with diminishing success as our society (particularly urban society) changes at a rapid rate. The need for a total revamping of our educational system is evident as each school day ends.

Summary

The recommendations obtained from the evidence in this study are as follows: 1) Teachers and administrators receive extensive training prior to the implementation of educational programs; 2) Educational programs should be assessed on a continuous bases to monitor the long range objectives; 3) Students should rate the effectiveness of their teachers; 4) Teachers should evaluate their own effectiveness in the classroom; 5) Teacher-student evaluation feedback should be utilized to assign teachers to classes in which they have demonstrated effectiveness; 6) Pilot studies should be developed around the grouping of students according
to their assessed learning processes; 7) Further exploration of the role of the teacher regarding teacher expectations and achievement in ability grouping situations; 8) Replication of this study exploring the effects of ability grouping on non-verbal teacher-student interaction; 9) Selection of a number of schools from a national sample replicating the variables utilized in this study.

Ideally students should walk into a school and decide which learning areas they wish to explore. They would pace themselves according to their ability to grasp concepts and skills. Students could pretest out of a learning area if they demonstrate competencies in achieving the objectives stated for that area. Alternatives would be offered for students to utilize all their learning capabilities. Teacher-student patterns would be varied and flexible, alleviating the tracking situations which exist in many ability grouping programs.

The above recommendations represent just one step in the long staircase of our educational system. The implementation of these suggestions would lay the foundations for needed change in many of our educational programs which seem to continue without proper monitoring systems. The effectiveness of our educational system will depend on our ability to assess the needs of the system on a continuous basis.


*Journal of Research and Development in Education*. Vol. 4, Number 1, Fall, 1970, Athens, Georgia.


Appendix A
The individual frequency counts for the ten Flanders categories are tallied on a ten by ten matrix. The ratios below are the result of different combinations of rows, individual cells and columns.

<table>
<thead>
<tr>
<th>Teacher Talk%</th>
<th>sum of columns #1-7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total # of tallies</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Student Talk%</th>
<th>sum of columns #9 and 9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>total # of tallies</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Big I/D</th>
<th>sum of columns #1-#4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indirect/Direct Differences</td>
<td>sum of columns #5-#7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Little i/d</th>
<th>sum of columns #1-#3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indirect/Direct Differences with Content Removed</td>
<td>sum of columns #6-#7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Extended i/d</th>
<th>sum of cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motivating Behaviors / Controlling Behavior</td>
<td>(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3), (3, 1), (3, 2), (3, 3)</td>
</tr>
<tr>
<td></td>
<td>(6, 6), (6, 7), (7, 6), (7, 7)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>8/9 I/D</th>
<th>sum of cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher Responses Immediately After Student Responds</td>
<td>(8, 1), (8, 2), (8, 3), (8, 4), (9, 1), (9, 2), (9, 3), (9, 4)</td>
</tr>
<tr>
<td></td>
<td>(8, 5), (8, 6), (8, 7), (9, 5), (9, 6), (9, 7)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Revisited 8/9 i/d</th>
<th>sum of cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher Response Immediately After Student Responds With Content Removed</td>
<td>(8, 1), (8, 2), (8, 3), (9, 1), (9, 2), (9, 3)</td>
</tr>
<tr>
<td></td>
<td>(8, 6), (8, 7), (9, 6), (9, 7)</td>
</tr>
</tbody>
</table>

(Flanders, 1960)*
<table>
<thead>
<tr>
<th>INDIRECT INFLUENCE</th>
<th>DIRECT INFLUENCE</th>
<th>STUDENT TALK</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. * ACCEPTS FEELING: accepts and clarifies the feeling tone of the students in a nonthreatening manner. Feelings may be positive or negative. Predicting or recalling feelings is included.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. * PRAISES OR ENCOURAGES: praises or encourages student action or behavior. Jokes that release tension, but not at the expense of another individual; nodding head, or saying &quot;um hm?&quot; or &quot;go on&quot; are included.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. * ACCEPTS OR USES IDEAS OF STUDENTS: clarifying, building, or developing ideas suggested by a student. As teacher brings more of his own ideas into play, shift to Category 5.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. * ASKS QUESTIONS: asking a question about content or procedure with the intent that a student answer.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. * LECTURING: giving facts or opinions about content or procedures; expressing his own ideas, asking rhetorical questions.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. * GIVING DIRECTIONS: directions, commands, or orders with which a student is expected to comply.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. * CRITICIZING OR JUSTIFYING AUTHORITY: statements intended to change student behavior from nonacceptable to acceptable pattern; bawling someone out; stating why the teacher is doing what he is doing; extreme self-reference.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. * STUDENT TALK-RESPONSE: a student makes a predictable response to teacher. Teacher initiates the contact or solicits student statement and sets limits to what the student says.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. * STUDENT TALK-INITIATION: talk by students, which they initiate. Unpredictable statements in response to teacher. Shift from 8 to 9 as student introduces own ideas.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. * SILENCE OR CONFUSION: pauses, short periods of silence, and periods of confusion in which communication cannot be understood by the observer.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There is NO scale implied by these numbers. Each number is classificatory; it designates a particular kind of communication event. To write these numbers down during observation is to enumerate—not to judge a position on a scale.
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>