Individualization of Instruction by Vocational and Nonvocational Teachers: Self-Reports Compared with Observations.


NOTE

143p.; Appendix A (pages 108-122 of the original document), containing two testing instruments for individualization of instruction, is copyrighted and therefore not available. It is not included in the pagination.

EDRS PRICE

MF-$0.83 HC-$7.35 Plus Postage.

ABSTRACT

Observations by trained observers using the Descriptive Observation Record of Individualized Instruction were compiled on 335 teachers, who also completed a self-report of the Individualization of Instruction Inventory. Both inventory results were then examined with the purpose of determining the relationship between self-reports of the teachers and reports made by the trained observers, with this relationship having implications for inservice programs. The teachers' group was composed of 200 academic, 86 vocational, and 49 special education teachers from 20 school districts located near Austin, Texas. Major findings showed that (1) teachers' self-reports yielded higher scores than observer recorded scores; (2) discrepancies between scores derived from teachers' self-reports and observer-reported scores increased as the degree of individualization of instruction, reflected from observations, decreased; (3) scores derived from teachers' self-reports varied in a systematic way from observer-reported scores, which permitted the conversion of teacher self-reports into scores that were reasonable predictors of observer reported scores; (4) teachers who were observed to be individualizing instruction more tended to have more realistic perceptions than did teachers who were observed to be individualizing instruction less, and (5) teachers who were from districts which had offered inservice education in individualization in instruction seemed more enthusiastic as evidenced by invitations to view materials and student work. (TA)
INDIVIDUALIZATION OF INSTRUCTION
BY VOCATIONAL AND NONVOCATIONAL TEACHERS:
SELF-REPORTS COMPARED WITH OBSERVATIONS

by

RICHARD J. HARDEBECK, CARL R. ASHBAUGH,
AND KENNETH E. McINTYRE

Published by
The Office of School Surveys and Studies
Department of Educational Administration
The University of Texas at Austin

in cooperation with

Division of Occupational Research and Development
Department of Occupational Education and Technology
Texas Education Agency

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

1974

PERMISSION TO REPRODUCE THIS COPY-
RIGHTED MATERIAL HAS BEEN GRANTED BY

RICHARD J. HARDEBECK

TO ERIC AND ORGANIZATIONS OPERATING
UNDER AGREEMENTS WITH THE NATIONAL IN-
STITUTE OF EDUCATION. FURTHER REPRO-
DUCTION OUTSIDE THE ERIC SYSTEM RE-
QUIRES PERMISSION OF THE COPYRIGHT
OWNER
### TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>I. INTRODUCTION</strong></td>
<td></td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td>1</td>
</tr>
<tr>
<td>Purposes of the Study</td>
<td>2</td>
</tr>
<tr>
<td>Assumptions</td>
<td>3</td>
</tr>
<tr>
<td>Limitations</td>
<td>4</td>
</tr>
<tr>
<td>Theoretical Framework</td>
<td>5</td>
</tr>
<tr>
<td>Definition of Variables</td>
<td>6</td>
</tr>
<tr>
<td>Research Questions</td>
<td>7</td>
</tr>
<tr>
<td>Instruments</td>
<td>8</td>
</tr>
<tr>
<td>Procedures for Data Collection and Analysis</td>
<td>9</td>
</tr>
<tr>
<td>Plan for Remainder of the Study</td>
<td>10</td>
</tr>
<tr>
<td><strong>II. REVIEW OF LITERATURE</strong></td>
<td>11</td>
</tr>
<tr>
<td>Observations Contrasted with Surveys</td>
<td>12</td>
</tr>
<tr>
<td>Teacher Self-Reporting</td>
<td>13</td>
</tr>
<tr>
<td>Research Relating Self-Reports to Observations</td>
<td>14</td>
</tr>
<tr>
<td>Number of Observations--Differing Views</td>
<td>15</td>
</tr>
<tr>
<td>Observations for Improvement of Teaching</td>
<td>16</td>
</tr>
<tr>
<td>Classroom Observation and Student Performance</td>
<td>17</td>
</tr>
<tr>
<td>A Needs Assessment</td>
<td>18</td>
</tr>
<tr>
<td>The Measuring Instrument</td>
<td>19</td>
</tr>
<tr>
<td>Summary</td>
<td>20</td>
</tr>
<tr>
<td><strong>III. PROCEDURES</strong></td>
<td>21</td>
</tr>
<tr>
<td>Selection of Subjects</td>
<td>22</td>
</tr>
<tr>
<td>Criterion Variable</td>
<td>23</td>
</tr>
<tr>
<td>Predictor Variable</td>
<td>24</td>
</tr>
<tr>
<td>Intervening Variables</td>
<td>25</td>
</tr>
<tr>
<td>Multiple Linear Regression Analysis</td>
<td>26</td>
</tr>
<tr>
<td>Summary</td>
<td>27</td>
</tr>
<tr>
<td>Chapter</td>
<td>FINDINGS OF THE STUDY</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------</td>
</tr>
<tr>
<td></td>
<td>The Sample</td>
</tr>
<tr>
<td></td>
<td>The Three Studies</td>
</tr>
<tr>
<td></td>
<td>Multiple Linear Regression Analysis</td>
</tr>
<tr>
<td></td>
<td>The First Hypothesis Tested</td>
</tr>
<tr>
<td></td>
<td>The Second Hypothesis Tested</td>
</tr>
<tr>
<td></td>
<td>The Third Hypothesis Tested</td>
</tr>
<tr>
<td></td>
<td>Tests of Equal Difference by Levels</td>
</tr>
<tr>
<td></td>
<td>Tests for Equality by Level</td>
</tr>
<tr>
<td></td>
<td>Test for Interaction</td>
</tr>
<tr>
<td></td>
<td>Self-Scores Minus Observed Scores Plotted</td>
</tr>
<tr>
<td></td>
<td>Against Self-Scores</td>
</tr>
<tr>
<td></td>
<td>Tests for Linearity</td>
</tr>
<tr>
<td></td>
<td>Comparison of F-Tests for Sub Groups</td>
</tr>
<tr>
<td></td>
<td>Summary of the Chapter</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter</th>
<th>SUMMARY AND CONCLUSIONS</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Limitations</td>
<td>103</td>
</tr>
<tr>
<td></td>
<td>General Conclusions within the Limitations of the Study</td>
<td>103</td>
</tr>
<tr>
<td></td>
<td>Implications for Future Research</td>
<td>105</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>APPENDICES</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Instruments</td>
</tr>
<tr>
<td>B</td>
<td>Letters</td>
</tr>
<tr>
<td>C</td>
<td>Reproductions of Computer-Plotted Scores</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td></td>
</tr>
</tbody>
</table>
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>The Organizational Context for In-Service</td>
<td>8</td>
</tr>
<tr>
<td>1-2</td>
<td>Decision-Making Settings</td>
<td>11</td>
</tr>
<tr>
<td>4-1</td>
<td>Self-Observed Scores Summary</td>
<td>56</td>
</tr>
<tr>
<td>4-2</td>
<td>Correlations of Observed Scores with Self-Reported and Observed Scores</td>
<td>60</td>
</tr>
<tr>
<td>4-3</td>
<td>Correlations of Self-Reported Scores with Self-Reported and Observed Scores</td>
<td>61</td>
</tr>
<tr>
<td>4-4</td>
<td>Predicted Mean Scores of Vocational Academic and Special Education Teachers</td>
<td>63</td>
</tr>
<tr>
<td>4-5</td>
<td>Tests of Equal Difference by Levels—Total</td>
<td>69</td>
</tr>
<tr>
<td>4-6</td>
<td>Tests of Equal Difference by Levels—Vocational</td>
<td>71</td>
</tr>
<tr>
<td>4-7</td>
<td>Tests of Equal Difference by Levels—Special Education</td>
<td>72</td>
</tr>
<tr>
<td>4-8</td>
<td>Tests of Equal Difference by Levels—Academic</td>
<td>73</td>
</tr>
<tr>
<td>4-9</td>
<td>Tests of Equal Difference by Levels—In-Service 1 and 2</td>
<td>74</td>
</tr>
<tr>
<td>4-10</td>
<td>Tests of Equal Difference by Levels—In-Service 3</td>
<td>75</td>
</tr>
<tr>
<td>4-11</td>
<td>Test for Equality of Levels—Vocational and Academic</td>
<td>77</td>
</tr>
<tr>
<td>Table</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>4-12</td>
<td>TEST FOR EQUALITY OF LEVELS--VOCATIONAL AND SPECIAL EDUCATION</td>
<td></td>
</tr>
<tr>
<td>4-13</td>
<td>TEST FOR EQUALITY OF LEVELS--SPECIAL EDUCATION AND ACADEMIC</td>
<td></td>
</tr>
<tr>
<td>4-14</td>
<td>TEST FOR EQUALITY OF LEVELS--IN-SERVICE 1 AND 2 AND IN-SERVICE 3</td>
<td></td>
</tr>
<tr>
<td>4-15</td>
<td>TEST FOR INTERACTION--VOCATIONAL AND ACADEMIC</td>
<td></td>
</tr>
<tr>
<td>4-16</td>
<td>TEST FOR INTERACTION--VOCATIONAL AND SPECIAL EDUCATION</td>
<td></td>
</tr>
<tr>
<td>4-17</td>
<td>TEST FOR INTERACTION--SPECIAL EDUCATION AND ACADEMIC</td>
<td></td>
</tr>
<tr>
<td>4-18</td>
<td>TEST FOR INTERACTION--IN-SERVICE 1 AND 2 AND IN-SERVICE 3</td>
<td></td>
</tr>
<tr>
<td>4-19</td>
<td>TEST FOR LINEARITY--TOTAL SAMPLE</td>
<td></td>
</tr>
<tr>
<td>4-20</td>
<td>TEST FOR LINEARITY--VOCATIONAL</td>
<td></td>
</tr>
<tr>
<td>4-21</td>
<td>TEST FOR LINEARITY--SPECIAL EDUCATION</td>
<td></td>
</tr>
<tr>
<td>4-22</td>
<td>TEST FOR LINEARITY--ACADEMIC</td>
<td></td>
</tr>
<tr>
<td>4-23</td>
<td>TEST FOR LINEARITY--IN-SERVICE 1 AND 2</td>
<td></td>
</tr>
<tr>
<td>4-24</td>
<td>TEST FOR LINEARITY--IN-SERVICE 3</td>
<td></td>
</tr>
<tr>
<td>4-25</td>
<td>COMPARISONS OF F-TESTS FOR SUB-GROUPS BY ALL REPORTS, SELF-REPORTS, AND OBSERVER REPORTS</td>
<td></td>
</tr>
<tr>
<td>5-1</td>
<td>FACTORS FOR ESTIMATING OBSERVED SCORES FROM SELF-REPORTS ON THE INDIIDUALIZATION OF INSTRUCTION INVENTORY</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

Statement of the Problem

As categorical funding for education undergoes close scrutiny by funding sources, the need for accountability in education becomes more evident. Evaluative data are essential to accountability and are dependent upon reliable information.

Information as to the degree of individualization of instruction is one type of evaluative data which has been used in a systems analysis approach to accountability in educational organizations (Harris, 1971). The Individualization of Instruction Inventory developed by Harris and Coody (1971) has been established as an acceptable instrument for use by trained observers in describing five aspects of individualization of instruction. Harris, Bessent, and McIntyre (1969) emphasize the point that reliability increases with multiple observations. While such reliability is desirable and may be sufficiently cost-effective for use in a local setting, the cost would be prohibitive in statewide evaluative efforts. A Teacher
The Questionnaire (TQ) form of the Individualization of Instruction Inventory has been utilized in conjunction with observations. Positive relationships between self-report and observer-completed versions of the Individualization of Instruction Inventory were predicted by Harris, but no systematic method had yet been developed to predict observed levels of individualization of instruction based on teachers' self-reports.

**Purposes of the Study**

The purposes of this study were:

1. to submit to systematic inquiry the self-reported estimates of individualization of instruction as compared with observations of the same teachers by trained observers.

2. to develop a set of factors for converting self-reports of individualization of instruction obtained from the Teacher Questionnaire of the Individualization of Instruction Inventory into scores which would be reasonable predictions of scores recorded by trained observers.
Such a set of conversion factors would then permit administrators of the Teacher Questionnaire of the Individualization of Instruction Inventory to provide data from large numbers of vocational, academic, and special education teachers for use in decision making by educational planners at the regional, state, or national level.

Assumptions

Two assumptions provided the basis for the proposed study. The first of these assumptions was that the teaching process can be submitted to systematic inquiry. Classroom teaching involves human beings in group situations. Observation techniques can be applied to the classroom teaching process.

The second of these assumptions was that teaching involves behaviors which can be identified, and which teachers can systematically acquire. These behaviors should be regarded as skills which can be acquired through the study and analysis of teaching. Skills can be practiced and controlled in and out of the classroom.

Limitations

This study was subject to the usual constraints which confront researchers in their attempts to publish
findings of significant impact on the educational community: namely, time, resources, and geography.

Although predictive model development was desired, it was anticipated that, in the strictest sense, it would only predict scores of individual teachers utilized in the sample considered in this study. Any generalizations made from information contained in this paper are done on the user's cognizance.

While reliability of data increases as the number of observations by trained observers increases, practical considerations limited this study to one observation and one self-report per teacher.

**Theoretical Framework**

The Texas Education Agency is the state agency charged by the Texas Legislature to direct and evaluate the nature and effects of public elementary and secondary education in Texas. A recent review of the Texas Education Code surfaced the following specific mandates from the Legislature to the Agency:

exercise "general control of the system of public education at the state level." Sec. 11.02(a)
serve as "the policy-forming and planning body for the public school system for the state" Sec. 11.24(a)

review "periodically the educational needs of the state," for adopting or promoting, "plans for meeting those needs," and for evaluating, "the achievements of the educational program" Sec. 11.26(a)

accept responsibility for "promoting efficiency and improvement in the public school system of the state." Sec. 11.52(b)

Thus, the Agency is accountable to the Texas Legislature and to the people of Texas for the wise investment of the public resources to produce educational benefits for Texas' citizens.

In carrying out these functions of setting directions for elementary and secondary education, assessing needs, and evaluating progress, the State Board of Education adopted in October, 1971, a set of Goals for Public School Education in Texas. Part I of the statement deals specifically with the outcomes for learners which the State Board desires to be achieved. Part II refers to "Organizational Efficiency," that is, to desired improvements in school programs and services which will increase learner benefits. Part III centers upon strengthening the capability of the system (the Agency, local districts, regional education service centers, etc.) to plan systematically and to be accountable for results produced.
In the State Board's Goals for Public School Education in Texas, Part II prescribes: "A program of continuing evaluation should be provided for measuring the performance of the public system in terms of the competency of its products and the efficiency of its structure and processes."

Section 16.310 of the Texas Education Code has provided for Texas public school teachers to be compensated for ten nonteaching days per year to be used for in-service education and preparation for the beginning and ending of school. This law has been in effect since 1970.

Vocational, academic, and special education teachers have been receiving in-service training in the individualization of instruction as part of the ten days of in-service training required by the State of Texas for all teachers. Educational decision makers need to know whether this in-service training in the individualization of instruction has had an effect upon the practices of these teachers and whether the effect can be measured as effectively through self-report as by observation by others.

Harris, Bessent, and McIntyre (1969) have the following to say about in-service education:

In summary, we have asserted that to put instructional change in its proper perspective the processes
for achieving that change must be clearly in view. Change may be brought about by use of authority, by changes in physical environment (facilities, materials, buildings), through use of rules and regulations, through changes in functional specialization, and through in-service development of personnel. Though it cannot stand alone, in-service development is the most fundamental of the change processes, since it is concerned directly with the individual, is aimed at some change in his knowledge and behavior, and can affect his willingness to accept the change.

Harris, Bessent, and McIntyre (1969) offer the following schematic which shows in-service education in the total context of change.

Individualization of instruction is the element to be changed by the in-service programs of concern to this study. Individualization of instruction can be viewed as a series of strategies to bring about a desired result.

Harris and McIntyre (1971) have broken these strategies into five classes. These strategies are (1) Intraclase Grouping, (2) Variety of Materials, (3) Pupil Autonomy, (4) Differentiated Assignments, and (5) Tutoring.

Intraclase grouping can be viewed as being made up of several aspects. These are (1) flexibility of groupings, (2) frequency of changes in grouping, (3) length of time students remain in groups, (4) variety of size and number of groups, and (5) freedom of movement.
TABLE 1-1
THE ORGANIZATIONAL CONTEXT FOR IN-SERVICE

The Organization
- Organizational Maintenance
  - Unplanned Change
  - Planned Change
    - Physical Change
    - Rule Change
    - Structural Change
    - Functional Change
    - Personnel Change
      - Replacement
      - Reassignment
      - In-Service Education
A variety of materials would consist of the following: (1) library books, (2) reference books, (3) teacher-made materials, (4) newspapers and magazines, and (5) audio-visual aids. These materials should have variations in levels of difficulty and interest.

Economy can be measured by examining the following items: (1) pupil self-direction, (2) involvement of pupils in planning, (3) involvement of pupils in the leading of activities, (4) self-grading work, (5) working with other students, and (6) unique learning situations.

The substrategies of differentiated assignments are (1) interest in assignments, (2) challenge and stimulation of pupils, (3) participation of pupil, (4) significant variations in assignments, (5) evidence of advanced-level or enrichment work, and (6) basing of assignments upon specific, diagnosed learning needs.

Tutoring can be done by the classroom teacher, special teachers, students, parents, or volunteers.

If the Texas Education Agency is to account for the expenditure of revenues for in-service education directed toward individualizing instruction, cost effective measures are needed for determining the extent to which
teachers practice individualization of instruction. As mentioned earlier in this paper, Harris, Bessent, and McIntyre (1969) have pointed out that reliability of observations made by trained observers using observation guides increases with multiple observations. If the purpose of observation is to improve practices of an individual teacher, several observations would be highly desirable.

If, on the other hand, the purpose of observation is to provide general information for statewide evaluation, assessment, or decision making, actual on-site observation would not be necessary, if data from teacher self-reports would be treated so as to yield results which would not be unreasonable to expect to have come from actual classroom observation by trained observers. Stufflebeam's (1971) modification of Braybrooke and Lindblom's (1963) diagram of decision-making settings (Table 1-2) illustrates this point. Degree of change is displayed on the horizontal axis, with magnitude increasing from the vertex. Similarly, rising from the vertex along the vertical axis is information grasp, representing increased precision in data collection methodologies. State departments of education and regional education service centers
<table>
<thead>
<tr>
<th>High Information</th>
<th>Low Information</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HOMEOSTASIS</strong></td>
<td><strong>METAMORPHISM</strong></td>
</tr>
<tr>
<td>Activity: Restorative</td>
<td>Activity: Utopian</td>
</tr>
<tr>
<td>Purpose: Maintenance</td>
<td>Purpose: Complete Change</td>
</tr>
<tr>
<td>Basis: Technical standards and quality control</td>
<td>Basis: Overarching theory</td>
</tr>
<tr>
<td><strong>INCREMENTALISM</strong></td>
<td><strong>NEOMOBILISM</strong></td>
</tr>
<tr>
<td>Activity: Developmental</td>
<td>Activity: Innovative</td>
</tr>
<tr>
<td>Purpose: Continuous Improvement</td>
<td>Purpose: Inventing, testing, and diffusing solutions to significant problems</td>
</tr>
<tr>
<td>Basis: Expert judgment plus structured inquiry</td>
<td>Basis: Conceptualization, heuristic investigation, and structured inquiry</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Small DEGREE OF CHANGE</th>
<th>Large</th>
</tr>
</thead>
</table>
traditionally seek continuous improvement of education through developmental activities. These descriptors are found in the lower left portion of the diagram. Therefore, a decision for a state department of education to place a high priority on allocating resources for teachers' in-service training in specific aspects of individualization could be justified through needs-assessment sampling, utilizing teacher self-reports of individualization of instruction.

Definition of Variables

In order to increase communication between the author and the reader, terms for which precise definitions are needed are presented in the following alphabetical listing:

1. Academic teachers: Those teachers who devote at least one-half time to activities classified as English (05 series), mathematics (11 series), science (13 series), or social studies (15 series) of the Texas Education Agency Assignment Code Sheet, TEA 637.

2. Elementary teachers: Those teachers in grades Kindergarten through six who devote at least
one-half time to activities classified as Elementary Education (19 series) of the Texas Education Agency Assignment Code Sheet, TEA 637.

3. Individualization of instruction: Those strategies used by teachers to promote initiative and self-direction on the part of students. There are five classes of these strategies. These classes are (1) intraclass grouping, (2) variety of materials, (3) pupil autonomy, (4) differentiated assignments, and (5) tutoring.

4. Special education teachers: Those teachers who devote at least one-half time to activities classified as Special Education (20 series) of the Texas Education Agency Assignment Code Sheet, TEA 637.

5. Vocational teachers: Those teachers whose time is reported to the Texas Education Agency on the Secondary Vocational Education Class Organization Report, VOC-066, R73.

Research Questions

Research Questions to be tested in the present study were:
1. Teachers' self-reports of individualization of instruction practices on the Individualization of Instruction Inventory yield higher scores than observations made using the same instrument.

2. Discrepancies between scores derived from teachers' self-reports of individualization of instruction practices on the Individualization of Instruction Inventory and scores from observations made using the same instrument will increase as the degree of individualization of instruction, reflected from observations, decreases.

3. Scores derived from teachers' self-reports of individualization of instruction practices will vary in a systematic way from scores derived from observations made using the same instrument, so as to permit development of a set of factors for converting self-reports of individualization of instruction obtained from teacher self-reports into scores which would be reasonable predictors of scores recorded by trained observers.

Instruments

This study used the Descriptive Observation Record for Individualization of Instruction (Appendix A)
for each classroom version. The abovementioned instrument was developed by Ben M. Harris and Kenneth E. McIntyre and has been used in several schools to analyze instruction.

The instrument that was used for the self-report of individualization of instruction is the Individualization of Instruction Inventory (Appendix A). This instrument was developed by Betty Coody and Ben M. Harris.

**Procedures for Data Collection and Analysis**

Letters were sent to superintendents of school districts within 100 miles of Austin inviting participation in the study. Twenty school districts were identified from which a sample elementary, academic, vocational and special education teachers were invited by letter to participate in the study. At least two vocational and special education teachers were chosen from each district.

One observation was made in the classroom of each teacher in the sample. After observations had been completed, each teacher was asked to complete and return the self-report form.

The observations were made by a trained team of observers. An interrater reliability score of .80 was
required by the team of observers during the training sessions.

The data were punched onto cards to permit their analysis through the use of the computer. All analyses were carried out with the statistical computer program LINEAR developed by Jennings (1971).

Plan for Remainder of the Study

In the following chapter a theoretical framework for presentation of hypotheses will be discussed. The literature was reviewed to identify criteria relevant to the hypotheses.

Chapter III contains a discussion of the procedure used in collecting data. References will be made to data gathering instruments and letters found in the Appendix.

Chapter IV will report the methods used in analyzing the data and the findings of the study. The final chapter will summarize the findings and report conclusions drawn as a result of the research. A discussion of implications of this study for further research will conclude the report.
CHAPTER II

REVIEW OF LITERATURE

Much has been written and even more has been spoken about the need for schools to meet the needs of each individual student and to be held accountable for doing so. The present study includes direct observation of individualization of instruction and comparison of those observations with teachers’ self-reports of individualization of instruction; thus this review will limit its scope to literature which bears most directly on the problem.

Ornstein (1971) pointed out that research on teacher behavior is not only voluminous but also contradictory.

Systematic classroom observation is defined in the Encyclopedia of Education (Deighton, 1971:168) as:

A set of procedures which uses systems of categories to code and quantify classroom behaviors of teachers and students. These procedures require that observed behaviors be coded or classified by the use of non-evaluative, relatively objective sets of categories which describe specific behaviors or actions.
An extensive and growing list of observation instruments is available (Simon and Boyer, 1967).

Early systematic studies on teacher behavior included those by Anderson, Lippitt and White, Withall, Flanders, Perkins, Cogan, Mitzel and Rabinowitz, and others (Amidon and Flanders, 1963). Beginning in the late 1930's, Anderson and his colleagues systematically observed "dominative" or authoritarian and "integrative" or democratic behaviors of five teachers over a period of years. They demonstrated that teacher behavior sets the climate of the class. Lippitt and White conducted laboratory experiments to analyze the effects of adult leaders on groups of five boys. Their conclusions confirmed and extended the findings of Anderson. Withall classified teachers' verbal statements into seven categories. These categories formed an index of teacher behavior which related to the integrative-dominative (I/D) ratio of Anderson. Flanders studied the effects of integrative and dominative behaviors on individual students in laboratory situations. Favorable reactions were recorded when students were exposed to integrative behaviors. Perkins used Withall's technique to demonstrate that groups led by integrative leaders were able to learn more about a given topic, child growth and development, more frequently than groups led by dominative leaders.
Cogan found that students reported doing more school work under teachers who could be classed as integrative. Mitzel and Rabinowitz used multiple observations from which they concluded that teachers' behavior patterns vary according to the immediate situation.

Harris, Bessent, and McIntyre (1969) discussed three purposes for systematic classroom observation. Administrative decision making involves use of data for such purposes as merit pay determinations, reprimands, promotions, and contract renewals. Program evaluating and planning focuses on the program rather than the individual teacher. Direct in-service experience may be gained by those being observed as well as by the observers. Professional skill is required for systematic classroom observation. They also pointed out, "Teaching is a very complex process. It is not possible for an observer to see everything while observing a classroom." Classroom observation should be for a specific purpose or purposes in order for the observer to focus on the most relevant occurrences.

McGaw, Wardrop, and Brunda (1972) listed three sources of error in classroom observation schemes. These include differences between observers, imprecise definitions of behaviors, and variations in teacher behavior.
Kaplan (1964) summarized the status of scientific observation. "Scientific observation is deliberate research, carried out with care and forethought, as contrasted with casual and largely passive perceptions of everyday life."

**Observations Contrasted with Surveys**

In this study, observations, made by trained observers, were contrasted with teacher self-reports obtained through surveying the same individuals who were observed.

According to Hayman (1968), the objective of descriptive research is to assess and describe certain characteristics of a particular situation at one or more points in time. Methods equal observation, survey, and content analysis. In observation, subjects do not deliberately furnish information about themselves. In survey, subjects help the researcher to gather data. Hayman recommended that the observation form should be limited to twenty-five categories. He also cautioned that the order of questions may affect respondents.

Hillway (1969) cautioned educational researchers that surveys help to pin down facts but cannot reveal what conditions are necessarily ideal.
Both observations and surveys are acceptable methods to assess and describe characteristics or situations. Observations by trained observers have the advantage of potentially high inter-observer reliability. Surveys have the advantages of low relative cost and lack of geographical constraints.

Teacher Self-Reporting

Allen and others (1970) made several interesting points in the volume entitled Teacher Self Appraisal: A Way of Looking over Your Own Shoulder. First, they pointed out, "In self-appraisal, any change in teaching behavior begins with a need to know yourself." Secondly, they stated, "Changes in teaching behavior are most likely to occur in a threat-free atmosphere." Thirdly, they stated, "The participation of teachers in a program of self-appraisal must be completely voluntary. Freedom of choice to enter and to leave the program must be an absolute guarantee for each teacher." A series of assumptions undergirding teacher self-appraisal have also been made: (1) "Teaching is more than mental processes, more than thinking. Basically it involves human interaction--where learning is the objective." (2) "The chance for teaching
to improve will occur only when the teacher behaves differently." (3) "Teaching is not a single pattern of 'most successful' behavior. Each teacher, therefore, should be free to develop his own unique style of teaching." (4) "Teaching behavior can be changed by one person--the teacher. No amount of command or exhortation from another can actually change one's behavior." and (5) "Teaching behavior most readily changes when the teacher is provided objective data of his own teaching."

One argument which may be made for teachers' self-reporting their perceptions of classroom behavior is that they are more likely to accept changes resulting from data which they were involved in developing. The American Association of School Administrators (1964) reported:

During his tenure at Columbia University, President Eisenhower engaged a firm of consultants to study the University's organization. He is reported later to have commented that the resulting document was "the most expensive and least-read document the University's library ever acquired."

They further reported:

Every teacher knows that involvement is critical to learning. The effective survey is one that involves every person whose behavior or attitude counts in a problem.
Motivation is another critical concept. In this context, it implies that those concerned must be aware of a problem in the real studies and want to see it solved. Want it enough to help solve it.

Research relating Self-Reports to Observations

Data collected for studies concerning classroom behavior are from direct or indirect measures. Indirect measures provide indicators from which inferences may be made. Student attendance rates, student achievement, disciplinary referrals, and the number of library books checked out by students are among a myriad of such indicators. Direct measures call for input either from the teacher or from someone else observing teaching-learning interactions. Observers may include students, fellow teachers, administrative or supervisory personnel from within the school district, or observers from sources external to the school. Such data collection may or may not be systematic. With a wide variety of variables and an abundance of literature available, a conscious decision was made to devote the following paragraphs to recent literature relating systematic observation to self-reports.

Several recent researchers have approached the question, "How do teachers see themselves and their
classrooms comparison to the observations of trained observers? The value of observing and being observed has been discussed. If self-reports could be used for making administrative decisions relative to program evaluating and planning, self-reports could also provide direct in-service experience.

Relevant often gather data by sending questionnaires, censuses, or other self-reporting forms to the population being studied. Is it reasonable to expect reliability in responses which come from individuals untrained in use of the particular instrument and possibly uninformed as to the intended meaning of terminology used in the instrument? A partial answer to this question was presented by Worle (1972), who studied the effects of training on the variance of teacher ratings. Variance depends on two factors. First, greater variance existed when raters used unscaled items than when they used scaled items. This supported the view of Harris, Bessent, and McIntyre (1969) that when a rater had specific behavior items upon which to focus his attention, the ratings tended to become more uniform than when only general terms were offered. Secondly, training did significantly alter variance; however, it would be difficult
to obtain reliable ratings using an extremely short training period.

There are significant differences among the instructional activities of teachers (Cohen, 1973). Classroom observation instruments may be divided into category systems and rating systems. Category systems are classified as low-inference measures (Gage, 1969), in that the items focus on specific, denotable, relatively objective behaviors. Rating systems are considered high-inference measures, because the observer must make inferences from a series of events (Rosenshine, 1970).

The major disadvantages of category systems are the cost of using observers and the difficulty of specifying behaviors before they can be included in a category system. Rosenshine (1970) and others have noted that the disadvantages of rating systems include (1) the halo effect, (2) the error of central tendency, (3) leniency error (Power, 1973), (4) difficulty in calibrating factors, such as "excellent," (Moutly, 1969), and (5) difficulty in recording high-inference items systematically (Gage, 1969).

Emmer (1968) discovered that teachers overestimated, by as much as 50 percent, the amount of time their
pupils had talked during the twenty-minute discussion that each teacher had conducted. Once teachers were able to estimate the amount of student participation better than others.

Using a randomly selected sample of participants, it was found that there was a discernible gap between teacher perceptions of their classroom practices and behaviors and their actual performances, as revealed by actual classroom observations (Young, 1973).

Kviding (1973) compared student evaluations of teachers with teacher self-evaluations. Data were obtained from a ten-trait scale, the Teacher Rating Scale, from both teachers and students, and from the Trait Ranking form, used by each teacher and student to rank the ten traits in order of importance. A tendency was noted for composite self-evaluations of all teachers to be similar to their students' composite evaluations of their teacher's effectiveness. Teachers and students agreed that "Knowledge of the Subject" is the most important trait, while "Personal Appearance" is least important.

Gwinn (1970) compared the perceptions of teacher competence, within and among three groups of educators, using the Instrument for the Observation of Teaching
activites IOTA. Classroom teachers, administrators, and college supervisors were involved in reacting to data gathered from thirty-minute video tapes of six respondents. Owens concluded that teachers are not as reliable judges of classroom behavior as college supervisors. Administrator reliability ranked between that of the other two groups. Owens' research raises a question for investigators who rely on teacher self-reports of their own behavior. If reliability is lacking in the more objective activity of measuring competencies of others, what justification is there to assume greater reliability as teachers approach the more subjective area of self-reporting?

Savage (1970) studied the validity of teacher self-reports for predicting pupil and observer ratings of teacher performance. Involved in his study were 100 elementary teachers, language arts/social science teachers, and 100 mathematics/science teachers. The correlations between predicted and actual scores were very low for all groups. Only scores of mathematics/science teachers using the Classroom Observation Record (COR) substantiated the hypothesis that self-reported responses can be used as significant predictors of pupil effectiveness, as judged by pupil and trained-observer responses.

34
In a study by Peek (1970), classroom teachers made self-analyses of their verbal behavior in the classroom and a trained observer made analyses of the same behavior using Flanders' Categories for Interaction Analyses System. Peek concluded that teachers did not appear to be able to self-analyze their overall verbal behavior except in the area of praising and encouraging students.

One basic difficulty in relating self-perceptions to observed behavior lies in determining the relationship between self-perception and observed behavior. Reynolds (1972) found that teacher self-acceptance and self-satisfaction are related to observed classroom behavior.

Zimmer (1976) studied the relationships between self-evaluations and supervisory ratings of first-year Lutheran elementary school teachers who had graduated from Concordia Teachers College, River Forest, Illinois. He found that teachers and supervisors agree on some things and disagree on others. Supervisor ratings of such factors as personal qualifications, commitment, professional qualifications, teaching competence, and classroom management correlated positively with teacher self-evaluation factors such as adjustment to the school organization, difficulty with class control, compatibility of assignment,
and liking of work at present location and negatively with using own ideas and plans for teaching and satisfaction with progress being made by students.

Fillbrandt (1971) used nine dimensions of teacher behavior, defined by Ryan in *Characteristics of Teachers*, to compare systematic observations with ratings by teachers and their supervisors. These dimensions included Democratic, Responsive, Understanding, Kindly, Stimulating, Original, Responsible, Systematic, and Optimist behavior. No significant relationships between ratings or self-rating and the interaction of rater's priorities for classroom behavior could be found.

**Number of Observations**

**Differing Views**

Researchers differ on opinions as to how many observations are needed on the same teacher before such information may be used with confidence. Harris, Assent, and McIntyre (1969) emphasized the point that reliability increases with multiple observations. The uses to be made of the information gathered through systematic observations will, in part, determine the information grasp required.
Medley and Mitzel (1963) contended that it is better to increase the number of observations than the number of observers. Instability of behavior over occasions gives rise to sampling error.

Voege (1970) determined, in a study of the procedures for evaluating classroom teachers in Washington, that effectiveness of evaluation lay in the manner of its performance more than in the frequency of teacher-evaluator contacts. The number of observations is not important.

Wells (1970) stated, "A more representative picture of the predominant teaching pattern being displayed by a teacher is found by combining at least three fifteen-minute sample episodes of classroom verbal interaction."
The instrument used in Wells's study was Flanders' Classroom Interaction Analysis.

Adachi (1971) reported that the third observation by individuals who had completed training in the Instrument for the Observation of Teaching Activities (IOTA) were consistent with IOTA consultants' scores of the same teacher's performance.

More Comprehensive Observation Guide items were identified as significantly reliable at the .05 level by pooling scores from three visits than were obtained from single-observation visits (Power, 1973).
If one use of classroom observation is to improve teaching, then a case can be made for training teachers in observation techniques. Teachers so trained would be in a better position to analyze and improve practices of themselves and of fellow teachers. Johnson (1971) suggested that teacher-colleague analysis of questioning skills might be an economical method of improving classroom teaching practices.

Moller (1968) found that first-year teachers listed fellow teachers most frequently as the one source from which they received the most help. Fellow teachers were followed by department chairman, 14 percent; principal, 7 percent; assistant principal, 4 percent; and supervisor, 3 percent.

Thomas (1972) found that teacher self-analysis of video tapes of own behavior leads to improvement. Teachers viewed daily video tapes of their own classroom behavior and counted the number of times they engaged in desired behaviors. The frequency of each behavior category increased as the teacher counted it.

Weeks (1972) developed an innovative approach to classroom observation. Each user of Weeks' system
selects ten behaviors for a series of observations to be made on his teaching in his own classroom by an observer of his choice. The ten behaviors are selected from a master list of 199 classroom behaviors including cognitive, effective, and psychomotor behaviors of teachers and students. Lists are pooled when a group of teachers are to observe one another.

Coombs (1966) added a new dimension to describing such teacher-student interactions as defining, explaining, comparing, and classifying in what he called the logic of teaching. He discussed eleven types of "moves" or verbal activities which introduce concepts. Moves may be made by the teacher, by a student, or by the teacher and one or more students together. Knowledge of such moves, Coombs contends, enables the teacher to be more critically conscious of the progress of classroom discussions; but no evidence is presented which would indicate that these concepts had been validated or that they could be reliably recorded, either by the teacher or by outside observers. The eleven moves were: (1) criterion description, (2) classification, (3) analysis, (4) analogy, (5) differentiation, (6) positive instance, (7) negative instance, (8) instance substantiation, (9) instance production, (10) enumeration, and (11) metadistinction.
Classroom Observation and Student Performance

Medley (1972) said the following to say about the relationships between classroom observation of teachers and performance of students under the tutelage of those teachers:

There is considerable research evidence to show that teachers who are rated most effective are not really the most effective in this (that pupils show the greatest gains while in her class) sense of the term. I have searched the literature and every study I have found which compared supervisor's ratings or judgments of teacher effectiveness with measured gains of pupils (adjusted to equalize ability levels, etc.) has found no appreciable relationship between them.

Although Rosenhine (1970) made a valid argument for relating teacher behavior to pupil achievement, Cohen (1973) pointed out that interactions among students and between the teacher and individual students may also relate to student achievement:

Sociologists point to structural factors in the organization of teaching producing so much variability in how teachers carry out instructional activities that the possibility of coming up with systematic understanding of the conditions for teaching effectiveness through inductive studies of teacher-talk and student learning would appear very limited. Secondly, studies of the participation rates of students in typical classrooms suggest that a tutorial
model of teacher-student interaction will prove quite inadequate to explain student learning. There are typically too few students who have the chance to interact with the teacher. Thirdly, studies of the effect of classroom status systems on the teacher and on the students' learning point to the necessity for conditionalizing any statements concerning teacher effectiveness: learning partly depends on the formal and informal structure of the classroom.

Literature cited thus far helps to establish the need for further research concerning the relationship between reported teacher self-perceptions and reports by trained observers.

A Needs Assessment

Selection of individualization of instruction as the area in which to collect data was the result of a study which included assessment of the needs for future in-service teacher training in Texas.

The study on in-service teacher training was conducted in the Spring of 1973, by the Division of Evaluation of the Texas Education Agency (Nutt, 1973). Participating in the study were independent school districts, education service centers, other divisions of the Texas Education Agency, and two professional teacher organizations. These two organizations were the Texas Classroom
Teachers association (TCTA), which represents classroom teachers, and the Texas State Teachers Association (TSTA), representing a broad spectrum of Texas educators.

As part of the study, respondents were asked to assess the need for future in-service education among various types of activities. Analysis of data from school districts involved a stratified random sample. A weighed frequency count was employed with high priority = 3, middle priority = 2, low priority = 1, and not needed = 0. From the pooled total of all districts, individualization of instruction was the most needed area for in-service. In Strata I, districts with over 50,000 ADA, individualization of instruction ranked number 22 among over 100 possible choices, while needs assessment study ranked number one. In Strata II, districts with 10,000 to 49,999 ADA, and in Strata V, districts with less than 500 ADA, individualization of instruction ranked number two. Training to help teachers to work with special education students having emotional problems was the prime concern in Strata II; while in Strata V, emphasis was given to the broad area of elementary school teaching skills. Individualization of instruction was the priority area of respondents from Strata III, districts with 1,000 to 9,999 ADA and
Strata IV, districts with 500 to 999 ADA. Texas Education Agency Officials ranked individualization as the sixth highest needed area for teacher in-service, with career education assuming highest priority. Education Service Center directors also saw career education as the priority in-service need, but ranked individualization of instruction a close second.

Questionnaires from Texas State Teacher Association members, saw individualization of instruction listed as third priority, behind developmental reading and development of skill in working with children who do not learn using traditional instructional methods.

Members of the Texas Classroom Teacher Association indicated that they considered in-service in individualization of instruction to be the highest priority.

School officials also reported areas in which in-service had been offered during the 1972-73 school term. Training in individualization of instruction seemed to be a function of size. All districts in Strata I reported that individualization of instruction was an in-service education topic as did 94 percent of the districts in Strata II, 69 percent of the districts in Strata II, 61 percent of the districts in Strata IV, and 45 percent of the districts in Strata V.
Instructional Leadership Materials, through the Office of School Studies and Surveys of The University of Texas at Austin, has made available self-reporting and observation guides for use in describing classroom behaviors in a systematic manner. One instrument, The Descriptive Observation Record of Individualization of Instruction (Harris and McIntyre, 1971) allows trained observers to record evidence as to the amount of five types of individualization of instruction:

I. Intra-class Grouping
II. Variety of Materials
III. Pupil Autonomy
IV. Differentiated Assignments
V. Tutoring

Another form of the instrument, the Individualization of Instruction Inventory, allows for teachers to report their perceptions of the degree to which they individualize instruction in each of these five categories.

The present forms of these instruments evolved from those found in a set of "Comprehensive Observation Guides and Inventories" (Harris and McIntyre, 1964).
key step in the development of these instruments was the use of the "Basic Teaching Procedures Scale" and the "Classroom Behavior Scale" in a study of the effects of demonstration teaching upon experienced and inexperienced teachers (Harris, 1966; Coody, 1967). In this study, trained observer-analysts observed classroom teachers at the beginning and the end of the school year. Each observation was for a duration of thirty minutes. Teachers were subjected to three supervisory approaches in order to test the effect of demonstration teaching on the teacher's individualization of instruction. Factors which were reported to be the best indicators of individualization of instruction were incorporated into the revised instruments.

Williams (1968) used a self-report form of the instrument in his study of practices of teachers in graded and nongraded classrooms. The degree to which teachers in nongraded classrooms reported their practices of individualization of instruction proved not to be significantly higher than those in graded classrooms. Another form of the instrument, the Special Classroom Observations Guide, was used by Murray (1971) in a study entitled "Individualization of Instruction in Special Learning Disabilities Resource Classrooms: A Comparative Study."
Murray's study involved fifteen elementary and fifteen junior high school learning disabilities resource teachers. Each teacher was observed during at least one thirty-minute period by three trained raters simultaneously. Subsequent to the observations, teachers completed a packet of reports which included a self-report form of the instrument. Significant differences were found between observations and self-reports in all aspects of individualization of instruction, with teachers viewing their own practices in a more favorable light.

Data gathered by observers were also used in a study in which teachers in districts where either laboratory methods or traditional approaches to individualization of instruction in-service were offered were compared with observations of teachers where individualization of instruction was not a major area of in-service training. Teachers in districts where in-service training in the area of individualization of instruction had been offered were observed to individualize instruction to a greater extent than the other teachers (Heeney, 1973).

Summary

The interactions between teachers and students in classroom situations may be observed systematically
by trained observers. Attempts to relate observed teacher behaviors to other factors such as student achievement or teacher self-perceptions have been inconclusive.

The need for teacher in-service training in individualization of instruction was affirmed by superintendents, education service center directors, state education agency personnel, and individuals from two major associations of professional educators. Lastly, the use of the Descriptive Observation Record of Individualization of Instruction to describe the degree of five varieties of individualization of instruction has been reviewed.

The review of the literature found in this chapter, information included in the theoretical framework in Chapter I, and references appropriately located throughout the paper, will serve as a point of departure for future researchers.
CHAPTER III

PROCEDURES

This chapter will describe the methods and procedures used in collecting the data for this study. Methods used in selecting subjects and characteristics of the subject will be described. Techniques used to measure the criterion variable, observed individualization of instruction, will be discussed, followed by an explanation of the measurement techniques used to obtain the predictor variable, self-reported individualization of instruction. Finally, the statistical procedures used in the study will be described briefly. A complete explanation of statistical techniques will be found in Chapter IV.

Selection of Subjects

Superintendents of 167 school districts within a 100-mile radius of Austin were contacted by letter and invited to participate in the study. These superintendents were also asked if their district had provided teachers with in-service education specifically designed to help them to individualize instruction. Those interested in
participating in the study were asked to identify a contact person within the district. Of the districts from which responses were received, the ten largest where in-service for individualizing instruction had not taken place were selected as well as ten districts with comparable average daily attendance (ADA) where in-service for individualizing instruction had taken place.

Copies of the Texas Education Agency Assignment Code Sheets, TEA-637, for the 1972-73 school term of the twenty participating districts were obtained from the management information center of the Texas Education Agency through the cooperation of Jerry T. Barton, Research Director. Teachers listed as being at least half time in areas of language arts, mathematics, social studies, science were identified and placed in a pool from which a random sample could be drawn. Similarly, special education teachers were identified. Directories listing vocational or occupational education teachers were obtained from the Division of Public School Occupational Programs of the Texas Education Agency, L. V. Ballard, Director. A sample of vocational, academic, and special education teachers, including at least two vocational and special education teachers, was selected from each district.
Criterion Variable

A team of observers was trained by William C. Heeney, using Harris's and McIntyre's Descriptive Observation Record of Individualization of Instruction. Heeney reported an interobserver reliability of .82. Training sessions included discussions of the instrument and instructions on completing it. Observer trainees used a training film as a simulated observation, with observers filling in checklists independently followed by discussions of what observers recorded. Actual classroom observation as part of the training was conducted in the Oak Hill Elementary School of the Austin Independent School District.

Observation schedules were established in each participating district. District protocol was observed in that the contact person in some districts set observation schedules while in other districts observation schedules were developed in conference with campus principals or their assistants.

Predictor Variable

In four of the districts the Individualization of Instruction Inventory Self-Report was completed by
teachers prior to classroom observation. Goodlad (1970) had noted a discrepancy between teacher knowledge of technique and individualization of instruction and classroom practices. Lamb (1970) revealed that the quality of teacher question was unaffected by prior knowledge that classroom observers would be using the Teacher Question Inventory.

Conversely, an article in The Encyclopedia of Education by Furst (1971:181), entitled "Classroom Observation, Systematic," cautioned against the use of classroom observation for teacher evaluation:

Several researchers object to the use of their instruments for evaluation because they believe that teachers may soon learn to falsify behavior patterns when they know they are being observed. Conversely, other teachers might become so self-conscious that they could not act normally.

It was hoped that collection of self-reports from some teachers before observation and from other teachers after observation would provide a particle of additional evidence for future researchers to consider in designing studies.

Self-report forms of the Individualization of Instruction Inventory were mailed to each of the 364 teachers participating in the study. The forms were accompanied by a letter and a preaddressed stamped envelope. A
A follow-up letter was sent to teachers who delayed in mailing back their self-reports. Where no response was received, an attempt was made to contact each teacher by telephone. As a final resort, during the week prior to keypunching of data, school districts where three or more self-reports had not been received were visited in person by the researcher. The resultant return netted 335 self-reports of the 364 teachers who had been observed.

**Intervening Variables**

The following information was obtained by observers and keypunched in addition to scores on the Individualization of Instruction Inventory:

<table>
<thead>
<tr>
<th>Observer</th>
<th>School level</th>
</tr>
</thead>
<tbody>
<tr>
<td>District</td>
<td>Teacher's ethnic background</td>
</tr>
<tr>
<td>In-service program of district</td>
<td>Number of Anglo students</td>
</tr>
<tr>
<td>Teacher identification</td>
<td>Number of Black students</td>
</tr>
<tr>
<td>Teacher assignment</td>
<td>Number of Mexican-American students</td>
</tr>
<tr>
<td>Teacher sex</td>
<td>Total number of students in class</td>
</tr>
</tbody>
</table>
Multiple linear regression was chosen as the major statistical analysis technique to be used in this study because of its versatility and power as a research tool. It is basically a technique in which a criterion (or dependent) variable $Y$ is predicted from a set of predictor (or independent) variables $X_1, \ldots, X_N$. The general equation for computing a subject's criterion score (predicted criterion score $\hat{Y}$) takes the form:

$$\hat{Y} = W_1X_1 + W_2X_2 + \cdots + W_KX_K$$

Where $\hat{Y}$ is the predicted criterion score

$W_1$ is a weight for a predictor

$X_1$ is a predictor score

$K$ is the number of predictor variables.

Since there is an error or discrepancy between the predicted score $\hat{Y}$ and the subject's actual criterion score $Y$ (signified by $Y - \hat{Y}$), the equation for the subject's actual criterion score takes the form:

$$Y = \hat{Y} + E = W_1X_1 + W_2X_2 + \cdots + W_KX_K + E.$$
Where $E$ is the error of prediction: $Y - \hat{Y}$. The task of multiple linear regression is to seek a set of weights which, when applied to the respective raw scores of each subject, will yield a set of error values with a minimum prediction error (Veldman, 1967:281-282).

Program LINEAR (Jennings, 1971) on file at The University of Texas Computation Center was the principal program used to perform the statistical tests of the hypotheses.

Summary

Observations by trained observers using the Descriptive Observation Record of Individualized Instruction were compiled on 345 teachers, who also completed a self-report of the Individualization of Instruction Inventory. Subjects consisted of vocational, academic, and special education teachers housed in elementary, junior high, or high schools of twenty independent school districts located within one hundred miles of Austin, Texas. Data were punched onto cards and analyzed on the CDC 6600 computer of The University of Texas at Austin, using the statistical computer program LINEAR developed by Jennings (1971).
CHAPTER IV

FINDINGS OF THE STUDY

The primary purpose of this study was to submit to systematic inquiry the self-reported estimates of individualization of instruction as compared with the observations of the same teachers by trained observers. This purpose was to be served by testing the three hypotheses irrespective of affirmation, disaffirmation, or inconclusiveness of the tests of these hypotheses.

A secondary purpose of the study was to develop a set of factors for converting self-reports of individualization of instruction obtained from the Teacher Questionnaire of the Individualization of Instruction Inventory into scores which would be reasonable predictions of scores recorded by trained observers.

This study was designed to provide information concerning the relationships between self-reported practices of vocational, academic, and special education teachers and observations of such practices as reported by trained observers. Are teachers' self-reported scores of individualization of instruction practices on the Individualization of Instruction Inventory higher than scores
of observers using the same instrument? Do discrepancies between scores derived from teachers' self-reports of individualization of instruction practices on the Individualization of Instruction Inventory and scores from observations made using the same instrument increase as the degree of individualization of instruction, reflected from observations, decreases? Do scores derived from teachers' self-reports of individualization of instruction vary in a systematic way from scores derived from observations made using the same instrument? These are some of the questions the study attempted to answer.

The Sample

Computer analyses were performed on data representing 335 teachers from twenty school districts located within one hundred miles of Austin, Texas. In the same were 200 academic, 86 vocational, and 49 special education teachers. A total of 169 were in ten districts which had offered in-service training in the individualization of instruction; 166 were in ten districts where other topics were emphasized during in-service teacher training. In each set of ten districts, self-reports were collected prior to observation in two districts and subsequent to observation in eight districts.
The Three Studies

Data collected during the course of this study were to be used in three separate but interrelated reports. The focus of this report is a comparison of observed and self-reported individualization of instruction of vocational, academic, and special education teachers. Another study described differences of observed practices of individualization of instruction among teachers exposed to different in-service teacher training activities (Heeney, 1973). In a third study, findings of the first two are to be analyzed in relation to specified demographic factors.

The hypotheses were presented in Chapter I. The literature relating systematically observed classroom practices and self-reports of such practices was reviewed in Chapter II. The design and procedure for executing the study were discussed in Chapter III. The present chapter will describe the results.

Multiple Linear Regression Analysis

Formulas for multiple linear regression were presented in Chapter III. The general equation may be written as
\[ \hat{Y} = W_1X_1 + W_2X_2 + \ldots + W_KX_K. \]

Where \( \hat{Y} \) is the predicted criterion score

- \( W_1 \) is a weight for a predictor
- \( X_1 \) is a predictor score and
- \( K \) is the number of predictor variables.

Considering the discrepancy between the predicted score \( \hat{Y} \) and the subject's actual score \( Y \), the error of prediction, \( Y - \hat{Y} \), the equation for the subject's actual score takes the form:

\[ Y = \hat{Y} + E = W_1X_1 + W_2X_2 + \ldots + W_KX_K + E. \]

Where \( E \) is the error of prediction: \( v - \hat{v} \).

Multiple linear regression analysis may be used to indicate the probability that a criterion variable is or is not statistically related to a set of predictor variables, but cause and effect judgments are subject to interpretation.

McNemar (1969:196) cautions:

Suppose that we have a dependent variable and four independent variables which might be used in the dependent variable. The cause and effect, as opposed to concomitant, relationship among variables is a logical problem which must be faced by the investigator as a logician rather than as a statistician.
Stufflebeam and others (1971:142) testify to the value of multiple regression analysis while echoing McNemar's caution:

An associational technique of utility with a large number of variables is multiple regression analysis, with the concept of contribution to total variance being the inclusion-exclusion criterion in refining the mathematical model. While most techniques are based on use of one dependent variable, canonical analysis is based on a model of multiple independent and dependent variables. Most association techniques are dependent upon logic and time sequence to establish the causal network. Causality, in the sense of cause and effect in the experiment, is never unequivocally established in correlational analysis; however, correlational analysis is a workable tool for prediction.

Veldman (1967:294) describes multiple correlation thus:

Multiple correlation may be considered a special case of the more general correlation model, with multiple predictors on one side and a single criterion on the other. The analytic procedure determines a set of weights for the predictor variables \(X_1\) which yield a composite variable \(\hat{Y}\) that correlates maximally with the criterion variable \(Y\).

Since Veldman referred to multiple correlations as a special case of canonical analysis, clarification of the later term is appropriate.

The goal of canonical analysis is to define the primary independent dimensions which relate one set of
variables to another set of variables. The output of a canonical analysis should suggest answers to questions concerning the number of ways in which the two sets of measures are related, the strengths of the relationships, and the nature of the relationships so defined.

Veldman (1967:282) summarized the chief value of multiple linear regression by saying,

the linear-regression approach to analysis of variance offers the investigator the possibility of making more precise tests of his hypotheses without many of the usual assumptions of traditional analysis of variance procedures.

Garrett (1966) offered four limitations to be considered in using partial and multiple correlations. First, linear regression is necessary. Secondly, the number of cases should be large. Thirdly, the influence of factors is not clear cut; and, lastly, the sampling error may be cumulative.

In this study, multiple linear regression analysis was employed to examine variables relevant to the prediction of the criterion, observed individualization of instruction, in order to determine which variables contributed to predictive efficiency. Also of importance was the possibility that the relationship of one predictor to the criterion depends upon the value of a second
predictor, considered simultaneously with all weighed predictor variables.

Program LINEAR (Jennings, 1971) was the principal program used to perform the statistical tests of the hypotheses. This program is on permanent file at The University of Texas Computation Center. It is a flexible program which permits the manipulation of data. Readers interested in learning more about the versatility of multiple linear regression are referred to Introduction to Linear Models by Ward and Jennings (1973).

The basic full model for the study was

\[ Y = Y_1X_1 + W_2X_2 + \ldots + W_KX_K + E \]

where the values were dependent on the relationship being tested at that time. For example, in the test of equal differences by levels for all subjects, the full model was

\[ Y = W_1X_1 + W_2X_2 + W_3X_3 + E \]

in which \( X_1 \) was "1" if the associated criterion score was related to self-reported scores 76 and below, "0" otherwise; \( X_2 \) was "1" if the associated criterion score was related to teacher self-reported scores 77 through 88, "0" otherwise; and \( X_3 \) was "1" if the associated criterion
scores was related to self-reported scores of 89 and above; and where the following set of restrictions could be imposed:

\[ W_2 - W_1 = C_0 = \text{Constant} \]
\[ W_2 = C_0 + W_1 \]
\[ W_3 - W_1 = 2C_0 \]
\[ W_3 = 2C_0 + W_1 \]

therefore

\[ Y = W_1X_1 + (C_0 + W_1)X_2 + (2C_0 + X_1)X_3 + E \]
\[ = W_1X_1 + C_0X_2 + W_1X_2 + 2C_0X_3 + W_1X_3 + E \]
\[ = W(X_1 + X_2 + X_3) + C_0(X_2 + 2X_3) + E \]
\[ = W_1U + C_0(X_2 + 2X_3) + E \]

where

\[ U = (X_1 + X_2 + X_3). \]

The restricted model was

\[ Y = A_1U + C_0(X_2X_3) + E^2 \]

where values for \( C_0, W_1, W_2, \) and \( W_3 \) were determined from the full model.
Continuing the example of the test for equal differences by levels for all subjects, the following values were determined from the full model:

\[ C_0 = 4.88 \pm 5 \]
\[ W_1 = 22.28 \pm 22 \]
\[ W_2 = 27.20 \pm 27 \]
\[ W_3 = 32.04 \pm 32. \]

The restricted model imposed the restriction that being in one of the three levels makes no significant contribution to prediction. Only one type of restriction was imposed on the basic full model at a time. Each predictor vector was removed from the full model, one at a time, to determine the effect of that predictor in the full model. The results of the example cited were presented in Table 4.6. Statistical tests in this study employed subroutines of the standard computer program LINEAR (Jennings, 1971) utilizing full and restricted models of multiple linear regression analysis.

The restricted model compared to the full model permits computation of F-Test values useful in testing hypotheses. The formula used for the F-Test is:

\[ F = \frac{(ESS_R - ESS_F)/df_1}{ESS_F/df_2}. \]
Where:

$\text{ESS}_R = \text{Error sum of squares of the restricted model.}$
$\text{ESS}_F = \text{Error sum of squares of the full model.}$
$\text{df}_1 = \text{Degrees of freedom---the number of linearly independent predictors in the full model, minus the number of linearly independent predictors in the restricted model.}$
$\text{df}_2 = \text{Degrees of freedom---the number of elements in } Y \text{ (Number of Subjects), minus the number of linearly independent predictors in the full model.}$

The First Hypothesis Tested

The first hypothesis postulated that teachers' self-reported scores of individualization of instruction would be higher than scores recorded by observers. Stated as a directional hypothesis, it was: Teachers' self-reports of individualization of instruction practices on the Individualization of Instruction Inventory will yield higher scores than observations made using the same instrument. For the purpose of applying the F-Test, this directional hypothesis was restated so as to postulate
<table>
<thead>
<tr>
<th>Item</th>
<th>Means</th>
<th>Standard Deviations</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed Intraclass Grouping</td>
<td>13.1370</td>
<td>6.7717</td>
<td>45.8559</td>
</tr>
<tr>
<td>Observed Variety of Materials</td>
<td>11.0350</td>
<td>5.4638</td>
<td>29.8530</td>
</tr>
<tr>
<td>Observed Pupil Autonomy</td>
<td>12.8834</td>
<td>6.0531</td>
<td>36.6395</td>
</tr>
<tr>
<td>Observed Differentiated Assmt.</td>
<td>13.2507</td>
<td>5.0977</td>
<td>25.8847</td>
</tr>
<tr>
<td>Observed Tutoring</td>
<td>8.4956</td>
<td>3.9904</td>
<td>15.9235</td>
</tr>
<tr>
<td>Observed Total Score</td>
<td>58.3615</td>
<td>21.2120</td>
<td>449.9509</td>
</tr>
<tr>
<td>Self Intraclass Grouping</td>
<td>17.7930</td>
<td>4.4781</td>
<td>20.0534</td>
</tr>
<tr>
<td>Self Variety of Materials</td>
<td>17.3207</td>
<td>4.2605</td>
<td>16.3228</td>
</tr>
<tr>
<td>Self Pupil Autonomy</td>
<td>17.9038</td>
<td>3.5034</td>
<td>12.2735</td>
</tr>
<tr>
<td>Self Differentiated Assmt.</td>
<td>16.8047</td>
<td>3.8973</td>
<td>15.1892</td>
</tr>
<tr>
<td>Self Tutoring</td>
<td>12.6647</td>
<td>3.9907</td>
<td>15.9255</td>
</tr>
<tr>
<td>Self Total Score</td>
<td>82.4636</td>
<td>15.3694</td>
<td>236.2195</td>
</tr>
</tbody>
</table>
equality: self-report total scores on the Individualization of Instruction Inventory are equal to observation total scores on the same instrument. The F-Test probability was 0.000, which was significant below the 1 percent level. The predicted mean self-score was 82.47; the predicted mean observation score was 58.36, yielding a difference of 24.11. One could conclude that self-scores were consistently higher by a sizable margin.

Inspection of Table 4-1 reveals that means for self-reported scores for each category of individualization of instruction were higher than the observed score in the same category. Standard deviations and variances were higher for observed scores than for self-reported scores in all categories except self-tutoring. The mean for the total self-reported score was 82.4636 and the standard deviation was 15.3694; while the mean for the total observed score was 58.3615 and the standard deviation was 21.2120.

The relationships or correlations among each of the means displayed in Table 4-1 were tested. These correlations are displayed in Tables 4-2 and 4-3. All correlations were positive, but none was .90 or greater. The fact that none of the scores of the five parts of the
<table>
<thead>
<tr>
<th></th>
<th>Observed Intraclass Grouping</th>
<th>Observed Variety of Materials</th>
<th>Observed Pupil Autonomy</th>
<th>Observed Differentiated Assessment</th>
<th>Observed Tutoring</th>
<th>Observed Total Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self Total Score</td>
<td>.3434</td>
<td>.2969</td>
<td>.2692</td>
<td>.3402</td>
<td>.2490</td>
<td>.3740</td>
</tr>
<tr>
<td>Observed Intraclass Grouping</td>
<td>1.0000</td>
<td>.5431</td>
<td>.7496</td>
<td>.5291</td>
<td>.8289</td>
<td></td>
</tr>
<tr>
<td>Observed Variety of Materials</td>
<td>.5431</td>
<td>1.0000</td>
<td>.5317</td>
<td>.4624</td>
<td>.6428</td>
<td></td>
</tr>
<tr>
<td>Observed Pupil Autonomy</td>
<td>.7496</td>
<td>.5317</td>
<td>1.0000</td>
<td>.3370</td>
<td>.6784</td>
<td></td>
</tr>
<tr>
<td>Observed Differentiated Assessment</td>
<td>.7192</td>
<td>.5058</td>
<td>.6431</td>
<td>1.0000</td>
<td>.5289</td>
<td>.8992</td>
</tr>
<tr>
<td>Observed Tutoring</td>
<td>.5291</td>
<td>.4624</td>
<td>.3370</td>
<td>.5289</td>
<td>1.0000</td>
<td>.7135</td>
</tr>
<tr>
<td>Observed Total Score</td>
<td>.8289</td>
<td>.6428</td>
<td>.6784</td>
<td>.6992</td>
<td>.7135</td>
<td>1.0000</td>
</tr>
<tr>
<td>Self Intraclass Grouping</td>
<td>.4071</td>
<td>.2616</td>
<td>.2985</td>
<td>.3644</td>
<td>.2753</td>
<td>.4071</td>
</tr>
<tr>
<td>Self Variety of Materials</td>
<td>.1556</td>
<td>.2791</td>
<td>.1078</td>
<td>.1604</td>
<td>.1090</td>
<td>.1976</td>
</tr>
<tr>
<td>Self Pupil Autonomy</td>
<td>.1773</td>
<td>.1140</td>
<td>.2516</td>
<td>.1792</td>
<td>.0495</td>
<td>.1898</td>
</tr>
<tr>
<td>Self Differentiated Assmt.</td>
<td>.3439</td>
<td>.2672</td>
<td>.2531</td>
<td>.3570</td>
<td>.2188</td>
<td>.3551</td>
</tr>
<tr>
<td>Self Tutoring</td>
<td>.2060</td>
<td>.1840</td>
<td>.1123</td>
<td>.1933</td>
<td>.2803</td>
<td>.2534</td>
</tr>
</tbody>
</table>
### Table 4-3

**Correlations of Self-Reported Scores with Self-Reported and Observed Scores**

<table>
<thead>
<tr>
<th></th>
<th>Self Intraclass Grouping</th>
<th>Self Variety of Materials</th>
<th>Self Pupil Autonomy</th>
<th>Self Differentiated Assessment</th>
<th>Self Tutoring</th>
<th>Self Total Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self Total Score</td>
<td>.7999</td>
<td>.7814</td>
<td>.7344</td>
<td>.8152</td>
<td>.6742</td>
<td>1.0000</td>
</tr>
<tr>
<td>Observed Intraclass Grouping</td>
<td>.4071</td>
<td>.1556</td>
<td>.1773</td>
<td>.3439</td>
<td>.2060</td>
<td>.3434</td>
</tr>
<tr>
<td>Observed Variety of Materials</td>
<td>.2616</td>
<td>.2791</td>
<td>.1140</td>
<td>.2672</td>
<td>.1840</td>
<td>.2969</td>
</tr>
<tr>
<td>Observed Pupil Autonomy</td>
<td>.2985</td>
<td>.1078</td>
<td>.2516</td>
<td>.2531</td>
<td>.1123</td>
<td>.2692</td>
</tr>
<tr>
<td>Observed Differentiated Assmt.</td>
<td>.3644</td>
<td>.1804</td>
<td>.1792</td>
<td>.3570</td>
<td>.1933</td>
<td>.3402</td>
</tr>
<tr>
<td>Observed Tutoring</td>
<td>.2753</td>
<td>.1090</td>
<td>.0495</td>
<td>.2188</td>
<td>.2803</td>
<td>.2490</td>
</tr>
<tr>
<td>Observed Total Score</td>
<td>.4071</td>
<td>.1976</td>
<td>.1898</td>
<td>.3551</td>
<td>.2534</td>
<td>.3740</td>
</tr>
<tr>
<td>Self Intraclass Grouping</td>
<td>1.0000</td>
<td>.5067</td>
<td>.4791</td>
<td>.5750</td>
<td>.4271</td>
<td>.7989</td>
</tr>
<tr>
<td>Self Variety of Materials</td>
<td>.5067</td>
<td>1.0000</td>
<td>.4860</td>
<td>.5911</td>
<td>.3697</td>
<td>.7814</td>
</tr>
<tr>
<td>Self Pupil Autonomy</td>
<td>.4791</td>
<td>.4660</td>
<td>1.0000</td>
<td>.5416</td>
<td>.3649</td>
<td>.7344</td>
</tr>
<tr>
<td>Self Differentiated Assmt.</td>
<td>.5750</td>
<td>.5911</td>
<td>.5416</td>
<td>1.0000</td>
<td>.4123</td>
<td>.8152</td>
</tr>
<tr>
<td>Self Tutoring</td>
<td>.4271</td>
<td>.3697</td>
<td>.3649</td>
<td>.4123</td>
<td>1.0000</td>
<td>.6742</td>
</tr>
</tbody>
</table>
instrument correlates at the .95 level or greater with the total scores indicated that each of the five parts of the instrument were necessary to obtain a total composite score. Strongest relationships were generally noted between component parts and the total scores. An exception to this was that the correlation between observed pupil autonomy and observed intraclass grouping was .7496, while the correlation between observed pupil autonomy and the observed total score was calculated to be .6784. The strongest relationships included those between (1) observed total score and observed differentiated assignments, .8992; (2) observed total score and observed intraclass grouping, .8289; (3) self-total score and self-differentiated assignment, .8152; and (4) self-total score and self-intraclass grouping, .7989.

Examination of self-reported scores and observation scores revealed some isolated incidences, usually toward the high or low extremes, in which observed scores were higher than self-reported; however, it was demonstrated statistically that self-reported scores tended to be higher than observed scores.

Table 4.4 presents predicted mean scores of vocational, academic and special education teachers. The
<table>
<thead>
<tr>
<th>Type of Teacher</th>
<th>Source</th>
<th>Predicted Mean</th>
<th>Standard Deviation</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocational</td>
<td>Self-reported</td>
<td>85.28</td>
<td>12.33</td>
<td>152.02</td>
</tr>
<tr>
<td></td>
<td>Observed</td>
<td>63.98</td>
<td>20.18</td>
<td>407.23</td>
</tr>
<tr>
<td>Academic</td>
<td>Self-reported</td>
<td>80.89</td>
<td>16.31</td>
<td>266.00</td>
</tr>
<tr>
<td></td>
<td>Observed</td>
<td>55.44</td>
<td>21.35</td>
<td>455.67</td>
</tr>
<tr>
<td>Special Education</td>
<td>Self-reported</td>
<td>82.37</td>
<td>14.95</td>
<td>223.42</td>
</tr>
<tr>
<td></td>
<td>Observed</td>
<td>62.10</td>
<td>19.43</td>
<td>277.55</td>
</tr>
</tbody>
</table>

*Predicted scores are calculated by the use of the actual scores. The predicted score is a score that will produce the smallest error sum of squares. This is done by subtracting the actual score from the predicted score, thereby obtaining an error score. The error score is squared and summed to obtain the error sum of squares.
mean predicted self-reported score of vocational teachers was 86.28, while the mean predicted observed score was 63.98, yielding a difference of 22.30. Special education teachers recorded the lowest difference, 20.27, between a mean predicted self-reported score of 82.37 and a mean predicted observed score of 62.10. Lowest predicted means, 80.89 self-reported and 55.44 observed, were recorded for academic teachers who also were noted to have the greatest difference, 25.45.

The Second Hypothesis Tested

The second hypothesis theorized that the self-reported scores of teachers who were observed to be individualizing instruction more would be closer estimates of the degree of individualization of instruction than self-reported scores of teachers who were observed to be individualizing instruction less. The hypothesis statement was: Discrepancies between scores derived from teachers' self-reports of individualization of instruction practices on the Individualization of Instruction Inventory and scores from observations made using the same instrument will increase as the degree of individualization of instruction, reflected from observations, decreases.
In order to test the hypothesis, it was assumed that the mean of the observation scores (58.3615), as reflected in Table 4-1, was a reasonable point of division into upper and lower observation scores. For the purpose of applying the F-Test, the second hypothesis was restated so as to presume equality: the differences between self- and observation scores, with observation scores above the mean, are equal to the differences between self- and observation scores below the mean. The F-Test probability was 0.000, which was significant below the 1 percent level. The below-the-mean difference was 36.28 and the above-the-mean difference was 10.65, yielding a discrepancy of 25.63 between the two groups. One concludes that the lower observation scores have a greater difference than the higher scores, thus confirming the second hypothesis.

A reproduction of the computer print-out is included in the Appendix on which self-reported scores on the Individualization of Instruction Inventory were plotted along the horizontal axis and observed scores were plotted along the vertical axis with the lowest possible score (twenty-five) at the upper left. It was included so that future researchers would be able to verify findings of this study and also to provide additional evidence of the
confirmation of hypotheses one and two. Since self-reported scores were generally higher than observed scores, it was practical to subtract observed scores from self-reported scores for purposes of data analysis. A graphic representation of self-scores minus observed scores plotted against observed scores is also presented in the Appendix. Visual inspection of this plot helped to confirm that lower observation scores have a greater difference than the higher scores.

The Third Hypothesis Tested

The third hypothesis dealt with (1) a systematic relationship between self-reports and observations and (2) use of self-reported scores of individualization of instruction to assess general practices of a population of teachers. The third hypothesis was stated thus: Scores derived from teachers' self-reports of individualization of instruction practices will vary in a systematic way from scores derived from observations made using the same instrument, so as to permit the development of a set of factors for converting self-reports of individualization of instruction obtained from teacher self-reports into scores which would be reasonable predictors of scores.
recorded by trained observers. In order to test this hypothesis, a number of statistical tests, all employing multiple linear regression, were employed. These tests included tests of equal difference, self-scores minus observed scores plotted against self-scores, tests for linearity, tests for equality of levels, and comparisons of F-Tests for subgroups. The essence of all these statistical tests, each of which will be discussed subsequently, was that self-reports of the total sample were divided into three levels which may be used to predict similar levels of scores reported by trained observers.

Dividing lines for the three levels were found to be one-half standard deviation above and below the mean of self-reported scores. Scores derived from teachers' self-reports of individualization of instruction into scores which were reasonable predictors of scores recorded by trained observers was developed. These factors included the following three levels: (1) for scores 76 and below--subtract 22 points; (2) for scores 77 through 88--subtract 27 points; (3) for scores 89 and above--subtract 32 points.

At this point, the author cautions the reader that this relationship and these factors applied only to the total sample tested in this particular study. This relationship
was not confirmed for various subgroups, nor could generalizations be made from only one study.

Tests of Equal Difference by Levels. Table 4.5 shows the results of the test of equal difference by levels for all subjects. Three discrete variables which are shown along the horizontal axis included: (1) 76 and below—self-reported scores more than one-half standard deviation below the mean; (2) 77 through 88—self-reported scores between one-half standard deviation below the mean and one-half standard deviation above the mean; and (3) 89 and above—self-reported scores more than one-half standard deviation above the mean. The difference between self-reported scores and observed scores is plotted on the vertical axis. Graphic presentation of both the full model predicted values and restricted model showed that self-reports of all respondents in the sample tested could be divided into three groups with an equal difference of approximately five points separating the mean self-scores minus observed scores. The approximate self-score minus observed-score level for scores 76 and below was 22; for scores 77 through 88, it was 27; and for scores 89 and above, it was 32. The probability of equal difference by levels for all subjects was .9801. This high probability
TABLE 4.5

TEST OF EQUAL DIFFERENCE BY LEVELS--TOTAL

<table>
<thead>
<tr>
<th>Level</th>
<th>76 and Below</th>
<th>77-88</th>
<th>89 and Above</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>22.28</td>
<td>27.20</td>
<td>32.04</td>
</tr>
<tr>
<td>Restricted</td>
<td>22.29</td>
<td>27.17</td>
<td>32.05</td>
</tr>
</tbody>
</table>

Self-Report Levels

- Full Model Predicted Values
- Restricted Model Predicted Values

Probability .9801
of equal differences by levels applied only to the total
sample. Table 4.6 indicates that the probability of equal
differences by levels for vocational teachers was only
.1544. Special education had the highest probability
(.8968) of equal differences by levels, as shown in Table
4.7. A low probability (.2163) of equal differences by
levels for academic teachers is displayed in Table 4.8.
The constant error for all subjects (Table 4.5) was ap-
proximately 5. It is interesting to note that while the
constant error for academic teachers (Table 4.8) was ap-
proximately , the constant error for vocational teachers
(Table 4.6) was 6.75, or 1.75 above 5; and the constant-
error for special education teachers was 3.24, or 1.76
below 5. Table 4.9 indicates that the probability of equal
difference by level for teachers in districts which had
offered in-service in the individualization of instruction
was .1386 and the constant error was 5.46. The proba-
bility of equal difference by levels for teachers in dis-
tricts which had not offered in-service in individualiza-
tion of instruction is shown in Table 4.10 to be .0998,
while the constant error was 4.37.

Test for Equality by Level. Vocational and
academic teachers were probably not equal (probability--
TABLE 4.6

TEST OF EQUAL DIFFERENCE BY LEVEL--VOCATIONAL TEACHERS

<table>
<thead>
<tr>
<th>Self Minus</th>
<th>Observed</th>
<th>Full</th>
<th>Restricted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below</td>
<td>76-88</td>
<td>89</td>
<td></td>
</tr>
<tr>
<td>Full</td>
<td>19.79</td>
<td>19.96</td>
<td>31.85</td>
</tr>
<tr>
<td>Restricted</td>
<td>17.12</td>
<td>23.87</td>
<td>30.62</td>
</tr>
</tbody>
</table>

Self-Report Levels

- Full Value Predicted Values
- Restricted Model Predicted Values

Probability .1544
Constant Error 6.75
TABLE 4.7

TEST OF EQUAL DIFFERENCE BY LEVEL--
SPECIAL EDUCATION TEACHERS

<table>
<thead>
<tr>
<th>Level</th>
<th>76 and Below</th>
<th>77-88</th>
<th>89 and Above</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>24.17</td>
<td>26.76</td>
<td>30.55</td>
</tr>
<tr>
<td>Restricted</td>
<td>23.90</td>
<td>27.14</td>
<td>30.39</td>
</tr>
</tbody>
</table>

Self-Report Levels

- Full Model Predicted Values
- Restricted Model Predicted Values

Probability .8968

Constant Error = 3.24
TABLE 4.8

TEST OF EQUAL DIFFERENCE BY LEVEL--ACADEMIC TEACHERS

<table>
<thead>
<tr>
<th>Level</th>
<th>76 and Below</th>
<th>77-88</th>
<th>89 and Above</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>22.63</td>
<td>30.47</td>
<td>32.57</td>
</tr>
<tr>
<td>Restricted</td>
<td>23.49</td>
<td>28.46</td>
<td>33.43</td>
</tr>
</tbody>
</table>

Self-Report Levels

- Full Model Predicted Values
- Restricted Model Predicted Values

Probability = 0.2163

Constant Error = 4.97
TABLE 4.9

TEST OF EQUAL DIFFERENCE BY LEVEL--IN-SERVICE 1 AND 2

<table>
<thead>
<tr>
<th>Level</th>
<th>76 and Below</th>
<th>77-88</th>
<th>Above</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>22.82</td>
<td>24.17</td>
<td>33.16</td>
</tr>
<tr>
<td>Restricted</td>
<td>21.30</td>
<td>26.76</td>
<td>32.22</td>
</tr>
</tbody>
</table>

Self-Report Levels

- Full Model Predicted Values
- Restricted Model Predicted Values

Probability .1586
Constant Error = 5.46
TABLE 4.10

TEST OF EQUAL DIFFERENCE BY LEVEL--IN-SERVICE 3

<table>
<thead>
<tr>
<th>Level</th>
<th>76 and Below</th>
<th>77-88</th>
<th>89 and Above</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>21.84</td>
<td>30.42</td>
<td>30.68</td>
</tr>
<tr>
<td>Restricted</td>
<td>23.14</td>
<td>27.51</td>
<td>31.89</td>
</tr>
</tbody>
</table>

Self Report Levels

- Full Model Predicted Values
- Restricted Model Predicted Values

Probability = 0.0998

Constant Error = 4.37
by the three levels tested: (1) 76 and below, (2) 77 through 88, and (3) 89 and above, as is shown in Table 4.11. The probability of equality of levels for vocational and special education teachers is shown in Table 4.12 to be .4944. Special education and academic teachers are shown in Table 4.13 to have the highest probability (.7531) of equality by levels. A low probability (.1666) of equality of levels for teachers in schools which offered in-service for individualizing instruction and teachers in districts which did not offer in-service for individualizing instruction is shown in Table 4.14.

**Test for Interaction.** All tests for interactions failed to provide evidence of interaction between subgroups of subjects. The probability of interaction between vocational and academic teachers is shown in Table 4.15 to be .1146. The probability of interaction between vocational and special education teachers, revealed in Table 4.16, was .4648. Special education teachers and academic teachers are shown in Table 4.17 to have a probability of interaction of .6975. Examination of Table 4.18 reveals low probability (.0856) of interaction between teachers in districts which had offered programs of in-service in individualization of instruction and teachers in districts where such in-service was not offered.
TABLE 4.11

TEST FOR EQUALITY BY LEVELS—VOCATIONAL TEACHERS
AND ACADEMIC TEACHERS

<table>
<thead>
<tr>
<th>Level</th>
<th>Vocational</th>
<th>76 and Below</th>
<th>77-88</th>
<th>89 and Above</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>19.78</td>
<td>19.97</td>
<td>31.85</td>
</tr>
<tr>
<td>Academic</td>
<td>22.63</td>
<td>30.46</td>
<td>32.58</td>
<td></td>
</tr>
</tbody>
</table>

Self-Report Levels

- Full Model Predicted Values—Vocational
- Full Model Predicted Values—Academic
- Restricted Model Predicted Values for Vocational and Academic

Probability .0341
TABLE 4.12

TEST FOR EQUALITY OF LEVELS—VOCATIONAL TEACHERS AND
SPECIAL EDUCATION TEACHERS

<table>
<thead>
<tr>
<th>Level</th>
<th>Vocational</th>
<th>Special Education</th>
</tr>
</thead>
<tbody>
<tr>
<td>76 and Below</td>
<td>19.79</td>
<td>24.17</td>
</tr>
<tr>
<td>77-88</td>
<td>19.96</td>
<td>26.76</td>
</tr>
<tr>
<td>Above</td>
<td>31.85</td>
<td>30.55</td>
</tr>
</tbody>
</table>

Self-Report Levels

- Restricted Model Predicted Values for Vocational and Special Education
- Full Model Predicted Values—Vocational
- Full Model Predicted Values—Special Education

Probability .4944
### Table 4.13

**Test for Equality of Levels--Special Education Teachers and Academic Teachers**

<table>
<thead>
<tr>
<th>Level</th>
<th>76 and Below</th>
<th>77-88</th>
<th>89 and Above</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic</td>
<td>22.63</td>
<td>30.46</td>
<td>32.57</td>
</tr>
<tr>
<td>Special Education</td>
<td>24.17</td>
<td>26.77</td>
<td>30.55</td>
</tr>
</tbody>
</table>

**Graph**

- Full Model Predicted Values--Academic
- Full Model Predicted Values--Special Education
- Restricted Model Predicted Values for Special Education and Academic

Probability .7531
TABLE 4.14

TEST FOR EQUALITY OF LEVEL--IN-SERVICE 1 AND 2
AND IN-SERVICE 3

<table>
<thead>
<tr>
<th>Level</th>
<th>76 and Below</th>
<th>77-88</th>
<th>89 and Above</th>
</tr>
</thead>
<tbody>
<tr>
<td>In-Service 1 &amp; 2</td>
<td>22.81</td>
<td>24.17</td>
<td>33.18</td>
</tr>
<tr>
<td>In-Service 3</td>
<td>21.85</td>
<td>30.42</td>
<td>30.68</td>
</tr>
</tbody>
</table>

Self-Report Levels

- Full Model Predicted Values--In-Service 1 & 2
- Full Model Predicted Values--In-Service 3
- Restricted Model Predicted Value--In-Service 1 & 2 and In-Service 3

Probability .1666
TABLE 4.15

TEST FOR INTERACTION—VOCATIONAL TEACHERS AND ACADEMIC TEACHERS

<table>
<thead>
<tr>
<th>Self Minus Observation</th>
<th>0</th>
<th>10</th>
<th>20</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>76</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>77-88</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>89</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Full Model Predicted Values—Vocational
Full Model Predicted Values—Academic
Predicted Values Assuming no Interaction

Probability .1146
TABLE 4.16

TEST FOR INTERACTION--VOCATIONAL TEACHERS AND SPECIAL EDUCATION TEACHERS

<table>
<thead>
<tr>
<th>Self-Report Levels</th>
<th>Full Model Predicted Values--Vocational</th>
<th>Full Model Predicted Values--Special Education</th>
<th>Predicted Values Assuming no Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>76 and Below</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>77-88 and Below</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>89 and Above</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Probability .4648
TABLE 4.17

TEST FOR INTERACTION--SPECIAL EDUCATION TEACHERS AND ACADEMIC TEACHERS

<table>
<thead>
<tr>
<th>Self-Minus Observation</th>
<th>0</th>
<th>10</th>
<th>20</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>76 and Below</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>77-88</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>89 and Above</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Self-Report Levels
- Full Model Predicted Values--Academic
- Full Model Predicted Values--Special Education
- Predicted Values Assuming no Interaction

Probability .6973
TABLE 4.18

TEST FOR INTERACTION--IN-SERVICE 1 AND 2
AND IN-SERVICE 3

Self-Minus Observation

<table>
<thead>
<tr>
<th>76 and Below</th>
<th>77</th>
<th>78</th>
<th>79 and Above</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Self-Report Levels

- Full Model Predicted Values--In-Service 1 & 2
- Full Model Predicted Values--In-Service 3
- Predicted Values Assuming no Interaction

Probability 0.0856
Self-Scores Minus Observed Scores Plotted Against Self-Scores. Self-scores minus observed scores plotted against self-scores are displayed in the Appendix. These reproductions of computer plotting all plot self-reported scores on the vertical axis and the absolute difference between self-reports and observation scores on the horizontal axis.

Tests for Linearity. Of the tests for linearity which are presented in Tables 4.19 through 4.24, one concludes that linearity did not exist for the total sample (Table 4.19), for academic teachers (Table 4.22), for in-service 1-2 (Table 4.23), or for in-service 3 (Table 4.24). The low probability of linearity of scores for vocational teachers (.0616) shown in Table 4.20 failed to substantiate a conclusion of linearity. The .7228 probability of linearity for special education teachers displayed in Table 4.21 indicates a need for further investigation into the possibility of linearity. In each of these tables, 4.19 through 4.24, observed scores are plotted on the vertical axis while self-reported scores are plotted on the horizontal axis.

Comparison of F-Tests for Subgroups. This study dealt with comparing teacher self-reports of individualization
<table>
<thead>
<tr>
<th></th>
<th>Observed</th>
<th>Self-Reported</th>
<th>Predicted Values Full Model</th>
<th>Predicted Values Restricted Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>125</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TABLE 4.19**

**TEST FOR LINEARITY--TOTAL SAMPLE**

![Graph showing observed and predicted values with a line indicating probability of linearity 0.0]
### TABLE 4.20

**TEST FOR LINEARITY -- VOCATIONAL TEACHERS**

<table>
<thead>
<tr>
<th>Self-Reported</th>
<th>Predicted Values Full Model</th>
<th>Predicted Values Restricted Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed</td>
<td>25</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>125</td>
</tr>
</tbody>
</table>

*Probability of Linearity: 0.0616*
TABLE 4.21

TEST FOR LINEARITY--SPECIAL EDUCATION TEACHERS

<table>
<thead>
<tr>
<th>Observed</th>
<th>Probability of Linearity .7228</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td></td>
</tr>
<tr>
<td>75</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
</tr>
<tr>
<td>125</td>
<td></td>
</tr>
</tbody>
</table>

Self-Reported

- Predicted Values Full Model
- Predicted Values Restricted Model
Table 4.22

Test for Linearity--Academic Teachers
TABLE 4.23

TEST FOR LINEARITY--IN-SERVICE 1 AND 2

Observed

Self-Reported

Predicted Values Full Model

Predicted Values Restricted Model

Probability of Linearity .0000

Probability of Linearity .0000
TABLE 4.24

TEST FOR LINEARITY--IN-SERVICE 3

<table>
<thead>
<tr>
<th>Observed</th>
<th>Predicted Values Full Model</th>
<th>Predicted Values Restricted Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Probability of Linearity 0.0
of instruction with reports of trained observers. In his study, Heeney (1973) reported results of F-Tests based on observer reports. Table 4.25 summarizes the F-Test probability for twelve tests, displaying the F-Test probability calculated from all scores, from self-reports only, and from observer reports. F-Tests, between vocational and academic teachers, are significant at the .01 level for all reports, self-reports, and observer reports. F-Tests between academic teachers who had in-service and academic teachers with no in-service was significant at the .01 level for all reports and self-reports, but was slightly more than .05 for observer reports. The F-Test between districts 15 and 16 and districts 11, 12, 13, 14, 17, 18, 19, and 20 was significant at the .01 level for self-reports. Districts 15-16 offered no in-service in individualizing instruction and teachers completed self-reports prior to observation. In the other districts (11, 12, 13, 14, 17, 18, 19, and 20), where teachers completed self-reports after classroom observations were made, scores were consistently higher than in districts 15 and 16. As also indicated in Table 4.32, F-Tests based on observer reports were significant below the .05 level for comparisons between special education teachers and academic
<table>
<thead>
<tr>
<th>Description</th>
<th>F-Test Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test between Vocational Teachers who had in-service (method 1 and 2 combined) and in-service 3 (no in-service)</td>
<td>0.3459 0.1897 0.4345</td>
</tr>
<tr>
<td>Test between Special Education Teachers who had in-service (1 and 2) and in-service 3 (no in-service)</td>
<td>0.7559 0.3787 0.2005</td>
</tr>
<tr>
<td>Test between Academic Teachers who had in-service (1 and 2) and in-service 3 (no in-service)</td>
<td>0.0070 0.0046 0.0512</td>
</tr>
<tr>
<td>Test between Vocational and Special Education Teachers</td>
<td>0.2434 0.1056 0.5901</td>
</tr>
<tr>
<td>Test between Vocational Teachers and Academic Teachers</td>
<td>0.0003 0.0062 0.0012</td>
</tr>
<tr>
<td>Test between Special Education Teachers and Academic Teachers</td>
<td>0.0867 0.5655 0.0412</td>
</tr>
<tr>
<td>Test between in-service 1 and in-service 2</td>
<td>0.9051 0.1468 0.3914</td>
</tr>
<tr>
<td>Test between in-service (1 and 2 combined) and in-service 3 (no in-service)</td>
<td>0.0271 0.0902 0.0519</td>
</tr>
<tr>
<td>Test between in-service 1 and in-service 3 (no in-service)</td>
<td>0.0856 0.6537 0.0359</td>
</tr>
<tr>
<td>Test between in-service 2 and in-service 3</td>
<td>0.0529 0.0317 0.2262</td>
</tr>
<tr>
<td>Test between districts 1 and 5 and districts 2, 3, 4, 6, 7, 8, 9, and 10</td>
<td>0.6964 0.0569 0.2061</td>
</tr>
<tr>
<td>Test between districts 15, 16 and districts 11, 12, 13, 14, 17, 18, 19 and 20</td>
<td>0.0757 0.0044 0.4182</td>
</tr>
</tbody>
</table>
teachers; and also between in-service 1 (laboratory approach) and in-service 3 (no in-service) based on self-reports. Pooling all reports yielded a significant difference, below .05, for teachers in districts which had offered any type of in-service for individualizing instruction compared with teachers in districts where such in-service had not been offered.

Based on the information contained in Table 4.25, one cannot conclude that use of either self-reports exclusively or observer reports exclusively would surface all significant relationships between variables tested.

Summary of the Chapter

Findings of the study were presented in this chapter. Following a review of the sample, the three studies, and multiple linear regression, were presentations of tests of each of the hypotheses. A summary of the findings and the implications for future research will be presented in Chapter V.
CHAPTER V

SUMMARY AND CONCLUSIONS

This study sought to test three hypotheses concerning relationships between self-reports of individualization of instruction by a sample of 335 vocational, academic, and special education teachers and reports completed by trained observers. The Individualization of Instruction Inventory (Harris and McIntyre, 1971) was the instrument employed in this study. Aspects of individualization of instruction included in the instrument were intraclass grouping, variety of materials, pupil autonomy, differentiated assignments, and tutoring.

The three hypotheses tested by this study were:

1. Teachers' self-reports of individualization of instruction practices on the Individualization of Instruction Inventory will yield higher scores than observations made using the same instrument.

2. Discrepancies between scores derived from teachers' self-reports of individualization of instruction practices on the Individualization of Instruction
Inventory and scores from observations made using the same instrument will increase as the degree of individualization of instruction, reflected from observations, decreases.

3. Scores derived from teachers' self-reports of individualization of instruction practices will vary in a systematic way from scores derived from observations made using the same instrument, so as to permit development of a set of factors for converting self-reports of individualization of instruction obtained from teacher self-reports into scores which would be reasonable predictors of scores recorded by trained observers.

The sample of 335 teachers consisted of 86 vocational, 200 academic, and 49 special education teachers located in 20 school districts within 100 miles of Austin, Texas. One hundred sixty-nine teachers were from 10 districts which had offered programs of in-service aimed toward helping teachers to improve practices of individualization of instruction. The remaining 166 teachers were from districts where individualization of instruction had not been a prime topic during in-service sessions. Teachers from two districts in each set completed the
self-report prior to being observed. Teachers from the
remaining eight districts in each set completed self-
reports after being observed. Observations were com-
pleted during the spring of 1973 by a team of trained
observers from the Office of School Surveys and Studies
of The University of Texas at Austin as part of a project
funded by the Division of Occupational Research and De-
velopment of the Texas Education Agency. Interobserver
reliability was reported to have been .82 (Heeney, 1973).
Data gathered by observers were also used in a study to
test the effect of in-service education on teacher per-
formance. Heeney (1973) reported that in-service train-
ing improved teacher performance as evidenced by reports
by trained observers using the Individualization of In-
struction Inventory. All data used in these studies were
made available for use in a third study, still in the
conceptual stages at the time of this writing.

Data were analyzed using multiple linear re-
gression analysis (Ward and Jennings, 1973) by means of
the computer program LINEAR (Jennings, 1971) on file at
the University of Texas Computation Center.

The major findings of the study were as fol-
lows:
1. There was a general tendency for scores of teachers' self-reports of individualization of instruction practices on the Individualization of Instruction Inventory to be significantly higher than observations made using the same instrument.

2. There was a significant difference between self-reports of teachers whose observed behaviors were below the mean and self-reports of teachers whose observed behaviors were above the mean.

3. Discrepancies between scores derived from teachers' self-reports of individualization of instruction practices on the Individualization of Instruction Inventory and scores from observations made using the same instrument increased as the degree of individualization of instruction, reflected from observations, decreased.

4. Total scores derived from teachers' self-reports of individualization of instruction practices varied in a systematic way from scores derived from observations made using the same instrument. A set of factors was developed for converting self-reports of individualization of instruction obtained from teachers' self-reports into scores which would be reasonable
predictors of scores recorded by trained observers. These factors are shown in Table 5-1. The
ness of these factors was tested by gene random sample of 49 teachers from the 335 in the
study. Twenty-four of the 49 fell into the same high, medium, or low level of individualization on actual
observer-reported scores as on predicted observation scores. In 12 cases, predicted levels of individuali-
ization of instruction were higher than observed levels, while in 13 cases observed levels were higher than
predicted levels. This indicates that conversion factors are useful for actuarial statistics, not for
individual statistics. For example, knowledge that
one teacher had a self-reported score on the Individ-
ualization of Instruction Inventory of 67 while another
teachers' self-reported score was 97 would not be
sufficient to make judgments concerning practices of
individualization by the two teachers. On the other
hand, if the mean self-reported score for teachers in
Region A was 67 and 97 for teachers in Region B, then
two conclusions would be justified: (1) teachers in
Region B practice individualization of instruction to
a greater extent than teachers in Region A, and (2)
TABLE 5.1
FACTORS FOR ESTIMATING OBSERVED SCORES FROM SELF-REPORTS ON THE INDIVIDUALIZATION OF INSTRUCTION INVENTORY

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>Range of Self-Reports</th>
<th>Factors to Subtract from Self-Reported Score to Predict Observation Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-reported scores more than one-half standard deviation below the mean</td>
<td>Scores 76-below</td>
<td>22</td>
</tr>
<tr>
<td>Self-reported scores between one-half standard deviation below mean and one-half standard deviation above mean</td>
<td>Scores 77-88</td>
<td>(Difference = 5)</td>
</tr>
<tr>
<td>Self-reported scores more than one-half standard deviation above the mean</td>
<td>Scores 89-above</td>
<td>32</td>
</tr>
</tbody>
</table>
there is a greater need for in-service education in individualizing instruction in Region A than in Region B. Let us carry the example a step further. If teachers in Region A were involved in in-service education on the basis of assessed subsequently the self-report form of the Individualization of Instruction Inventory administered, higher self-reported posttest scores might indicate improved teacher practices but also could be associated with greater sophistication in understanding terminology employed in the instrument.

5. Factors for predicting observed levels of individualization of instruction from self-reported scores met tests of statistical significance for the sample population. While factors for predicting observed levels of individualization of instruction from self-reported scores of subgroups did not meet the test of statistical significance, probabilities were too high to be rejected statistically.

6. In districts where individualization of instruction had not been a prime concern of in-service, and where observation preceded self-reporting, teachers had significantly higher self-reported scores on the
Individualization of Instruction Inventory than teachers from the same set of districts in which observation followed self-reporting. In districts where teachers had experienced in-service education in individualization of instruction, the difference between self-reports of teachers who self-reports prior to observation when compared to teachers who completed self-reports after being observed was not significant. Teachers who are individualizing instruction tend to know that they are doing so; on the other hand, teachers who are not individualizing instruction do not seem to know whether or not they are. When self-reporting follows observation, the self-report is tied to a specific class session; therefore, it would appear to be somewhat threatening. A self-report which precedes observation and is not tied to a specific session may elicit middle-of-the-road—"sometimes I do and sometimes I don't"—responses. One implication of this is that researchers, when asking for information which may call for a self-indictment, should use as non-threatening methods as possible. Another possible implication is that in-service education aimed at individualization of instruction
helped teachers to have perceptions of their own performances which were affected less by whether observation or self-reporting came first, thereby reducing differences in self-reported scores.

**Limitations**

This study was subject to the usual constraints which confront researchers in their attempts to publish findings of significant impact on the educational community—namely, time, resources, and geography.

Factors developed for estimating observed scores from self-reports on the Individualization of Instruction Inventory applied only to the population involved in this particular study.

Practical considerations limited this study to one observation and one self-report per teacher. Thus, reasonable caution should be exercised in drawing conclusions and generalizations from the data.

**General Conclusions Within the Limitations of the Study**

Previous research (Harris, 1968; Goody, 1967; Murray, 1971; and Heeney, 1973) has demonstrated that
trained observers are able to describe teacher practices of five aspects of what is called individualization of instruction through use of the Individualization of Instruction Inventory (Harris and McIntyre, 1971).

This study has presented evidence that self-reported perceptions of individualization of instruction by a sample population which included vocational, academic, and special education teachers, using the Teacher Questionnaire form of the Individualization of Instruction Inventory, bear sufficient relationships to scores recorded by trained observers who used the Individualization of Instruction Inventory so as to permit describing teacher practices of these five aspects of individualization of instruction as being high, medium, or low. Dividing points fell one-half standard deviation below the mean and one-half standard deviation above the mean for self-reported perceptions on the Teacher Questionnaire of the Individualization of Instruction Inventory. These dividing points, as shown in Table 5-1, set the limits of the three groups as being high—scores 89 and above; medium—scores 77-88; and low—scores 76 and below. Subtracting 32, 27, and 22, respectively, from the mean of each group yielded reasonable estimates of the means of observer-recorded scores.
Teachers' self-reported scores on the Individualization of Instruction Inventory generally were higher than observer reported scores. Teachers who were observed to be individualizing instruction more tended to have more realistic perceptions of their practices than do teachers who were observed to be individualizing instruction less. A serendipitous observation of this researcher was that of the teachers visited, those who were from districts which had offered in-service education in individualization of instruction seemed more enthusiastic as evidenced by invitations to view materials and student work, as well as positive comments concerning the teaching-learning environment in which they operated.

Implications for Future Research

Four categories of suggestions for further study are offered. First, future researchers might consider replicating the study using another sample population. Such a study would help to determine if, in fact, generalizations could be made from the findings of this study.

Secondly, Section 16.310 of the Texas Education Code provides that teachers in Texas' public schools experience 10 days of in-service education and preparation
for the beginning and ending of schools. Educational decision makers perceive individualization of instruction as a priority area for in-service training (Nutt, 1973).

A rational and cost-effective way to determine the general need for individualization of instruction in-service teacher training by a population of teachers would be a needs assessment study in which the Teacher Questionnaire of the Individualization of Instruction Inventory would be administered to a sample of teachers in the geographic area of concern.

Thirdly, teaching involves a complex set of activities. Heeney (1973) has suggested that in-service experiences may affect observed teacher practices of individualization of instruction. Further research is needed to describe other factors which may affect teacher practices of individualization of instruction. Among the factors which investigators may wish to consider are teacher attitudes, demographic factors, leadership styles of building principals, and availability of resources. Information from such studies would help to provide educational decision makers with some of the factors which may be associated with teacher behavior.

Finally, studies are recommended which concern teacher-pupil interaction--what student differences may
be associated with differences in teacher practices of individualization of instruction?
APPENDIX A, CONTAINING THE DESCRIPTIVE RECORD FOR INDIVIDUALIZATION OF INSTRUCTION (INSTRUMENT 1F-1, REVISED) AND THE INDIVIDUALIZATION OF INSTRUCTION INVENTORY (INSTRUMENT I1, REVISED) BY BETTY COODY AND BEN M. HARRIS, HAS BEEN REMOVED BECAUSE IT WAS COPYRIGHTED.
APPENDIX B

LETTERS

123

119
TO THE SUPERINTENDENT ADDRESSED:

The Office of School Surveys at the University of Texas is under contract to the Texas Education Agency to perform a study of in-service education. The purpose of this study is to determine if in-service education has had an effect upon the practices of vocational, special education, and academic teachers' performance in the classroom. The focus of the study will be on the individualization of instruction.

Collection of data will be by an observation visit to a sample of classrooms and a self-report by the same sample of teachers. Teachers will be asked to voluntarily participate in the study by letter.

We are asking for your help. Ten districts who have had in-service education in individualization of instruction and ten districts who have not had in-service education in individualization of instruction will be selected to participate in the study. Please complete the enclosed form and return it at your earliest convenience.

If you have any questions or if we can be of any assistance to you, please let us know.

Sincerely,

William C. Heeney
Project Director

Carl Ashbaugh, Ph.D.
Director
Office of School Surveys and Studies
WCH:CA:ers
Enclosure
IN-SERVICE EDUCATION QUESTIONNAIRE

1. Has your school district conducted an in-service education program of at least four clock hours in the individualization of instruction since July 1970?
   ____ Yes
   ____ No

2. Would you be willing for your school district to participate in this study of in-service education?
   ____ Yes
   ____ No

3. If yes, who can serve as your district's contact person for this study?
   Name ____________________________
   Title ____________________________

   District ____________________________
   Signature ____________________________

Please return to: Mr. William C. Heeney
Office of School Surveys
Education Annex
The University of Texas at Austin
Austin, Texas 78712
TO THE PRINCIPAL ADDRESSED:

The Office of School Surveys and Studies at the University of Texas is under contract to the Texas Education Agency to perform a study of in-service education. The purpose of this study is to determine if in-service education has had an effect upon the practices of vocational, special education, and academic teachers' performance in the classroom.

Your superintendent has agreed for your district to participate in this study of in-service education.

Collection of data will be by an observation visit to a sample of classrooms and a self-report by the same sample of teachers. Attached for your information is a copy of the letter which is being sent to this sample of teachers on your campus.

If you have any questions or if we can be of any assistance to you, please let us know.

Sincerely,

William C. Heeney
Project Director

Carl R. Ashbaugh, Ph.D.
Director
Office of School Surveys and Studies

Enclosure
TO THE TEACHER ADDRESSED:

The Office of School Surveys and Studies at The University of Texas is under contract to the Texas Education Agency to perform a study of in-service education. The purpose of this study is to determine if in-service education has had an effect upon the practices of vocational, special education, and academic teachers' performance in the classroom.

The superintendent of schools in your district has given us permission to include your district in this study of in-service education.

We need your help. We would like to make an observation visit to your classroom at a mutually agreed upon time. This visit will be for at least thirty minutes and not longer than one hour. In addition to the one classroom observation we will need you to complete a self-report form the first week in May.

Please complete and return the enclosed form at your earliest convenience.

If you have any questions or if we can be of any assistance to you, please let us know.

Sincerely,

William C. Heeney
Project Director

Carl R. Ashbaugh, Ph.D.
Director
Office of School Surveys and Studies

Enclosure
OBSERVATION QUESTIONNAIRE

1. Are you willing to participate in this study of in-service education?
   Yes
   No

2. If yes, as a preliminary part of working out a schedule of observations please list four days and hours in March and April convenient to you for this observation to take place.

   March
   1. ____________________________
   2. ____________________________
   3. ____________________________
   4. ____________________________

   April
   1. ____________________________
   2. ____________________________
   3. ____________________________
   4. ____________________________

You will receive notice before this observation is made in your classroom.

School___________________________________________
District___________________________________________
Name_____________________________________________

Please return to: Mr. William C. Heeney
Office of School Surveys
Education Annex
The University of Texas at Austin
Austin, Texas 78712
TO THE TEACHER ADDRESSED:

There are twenty-five items on this inventory. Circle only one number on each item that most appropriately describes your classroom. Your name and date should appear at the top of the form.

Please complete and return within five days from the date that you receive this inventory.

The members of the research team wish to express their appreciation to you for taking part in this project. We wish you a very pleasant summer.

Sincerely,

William C. Heeney
Project Director

WCH:dh

Enclosures
To the Teacher addressed:

Recently you participated in the first part of a study of individualization of instruction.

Your participation in the second part of this study is needed. Please complete and return the Individualization of Instruction Inventory right now while you are thinking about it. It contains 25 items, is in a yellow cover and was mailed to you in a large envelope.

The members of the research team wish to thank you for participating in this study. Have a happy summer.

Thank you for completing and returning the inventory while it is fresh on your mind.

Sincerely,

William C. Heeney

P.S. If you have already mailed this self-report, accept our thanks.

WCH:cl
APPENDIX

REPRODUCTIONS OF COMPUTER-LOTTED SCORES

131

127
SELF SCORES:LOTTE:D AGAINST OBSERVED SCORES

Self

128
VOCATIONAL SELF-MEAS. OBSERVED SCORES PLOTTED AGAINST SELF SCORES

Self-Observed
ACADEMIC SELF MINUS OBSERVATION SCORES PLOTTED AGAINST SELF SCORES

Self-Observation vs. Self-Score
IN-SERVICE 1 and 2 SELF MINUS OBSERVED SCORES PLOTTED AGAINST SELF SCORES

Self-Observed
IN-SERVICE 3 SELF MINUS OBSERVATION SCORES PLOTTED AGAINST SELF SCORES

Self-Observation
BIBLIOGRAPHY
BIBLIOGRAPHY


Harris, Ben M. *A Research Study of the Effects of Demonstration Teaching upon Experienced and Inexperienced Teachers.* Austin, Texas: The University of Texas, 1966.


Harris, Ben M. and Kenneth E. McIntyre. "Descriptive Observation Record for Individualization of Instruction," adapted from *Individualization of Instruction Inventory* by Betty Coody and Ben M. Harris. Austin, Texas: The University of Texas, 1971.


Jennings, Earl. "LINEAR." The University of Texas, 1971. Computer program on permanent file at the University of Texas at Austin Computation Center.


Texas. *Texas Education Code (1971), Sec. 11.02(a), 11.24 (a), 11.26(a), 11.52(b), 16.310.*


