Refining valid, reliable and discriminating student feedback items for use as one component of a total teacher performance evaluation system

Les M. Omotani
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Educational Administration and Supervision Commons, and the Educational Assessment, Evaluation, and Research Commons

Recommended Citation
Omotani, Les M., "Refining valid, reliable and discriminating student feedback items for use as one component of a total teacher performance evaluation system " (1992). Retrospective Theses and Dissertations. 9941.
https://lib.dr.iastate.edu/rtd/9941

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Refining valid, reliable and discriminating student feedback items for use as one component of a total teacher performance evaluation system

Omotani, Les M., Ph.D.
Iowa State University, 1992

Copyright ©1992 by Omotani, Les M. All rights reserved.
Refining valid, reliable and discriminating student feedback items for use as one component of a total teacher performance evaluation system

by

Les M. Omotani

A Dissertation Submitted to the Graduate Faculty in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Department: Professional Studies in Education
Major: Education (Educational Administration)

Approved:
Signature was redacted for privacy.

In Charge of Major Work
Signature was redacted for privacy.

For the Department
Signature was redacted for privacy.

For the Education Major
Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1992
DEDICATION

This dissertation is dedicated to the memory of Robert Carter, a superintendent who truly cared for others, found that which was good in all, and made a difference in the lives of students. Although Bob's own dissertation remains unfinished, his life served as a model of the value of friendship and his death served as a reminder of the precious nature of every day.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER I. THE PROBLEM</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td>6</td>
</tr>
<tr>
<td>The Research Questions</td>
<td>8</td>
</tr>
<tr>
<td>Definition of Terms</td>
<td>9</td>
</tr>
<tr>
<td>Delimitations of the Study</td>
<td>11</td>
</tr>
<tr>
<td>Human Subjects Release</td>
<td>14</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER II. REVIEW OF LITERATURE</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Challenges and Opportunities</td>
<td>15</td>
</tr>
<tr>
<td>Formative or Summative Evaluation</td>
<td>24</td>
</tr>
<tr>
<td>Original SIM Efforts</td>
<td>30</td>
</tr>
<tr>
<td>A Total Systems Approach</td>
<td>35</td>
</tr>
<tr>
<td>Developing a Complete Picture</td>
<td>38</td>
</tr>
<tr>
<td>An Informed Client</td>
<td>39</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER III. METHODS</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Questionnaires</td>
<td>44</td>
</tr>
<tr>
<td>The Sample</td>
<td>48</td>
</tr>
<tr>
<td>Administration of the Questionnaires</td>
<td>50</td>
</tr>
<tr>
<td>Treatment of Data</td>
<td>52</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER IV. FINDINGS</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overview</td>
<td>60</td>
</tr>
<tr>
<td>Student Feedback Questionnaire</td>
<td>61</td>
</tr>
</tbody>
</table>
### CHAPTER V. SUMMARY, CONCLUSIONS, LIMITATIONS, DISCUSSION, AND RECOMMENDATIONS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>101</td>
</tr>
<tr>
<td>Conclusions</td>
<td>103</td>
</tr>
<tr>
<td>Limitations</td>
<td>106</td>
</tr>
<tr>
<td>Discussion</td>
<td>108</td>
</tr>
<tr>
<td>Recommendations for Practice</td>
<td>112</td>
</tr>
<tr>
<td>Recommendations for Further Research</td>
<td>114</td>
</tr>
</tbody>
</table>

### BIBLIOGRAPHY

Page 117

### ACKNOWLEDGMENTS

Page 123

### APPENDIX A. K-2 CAVE CREEK REGULAR STUDENT RATING OF TEACHERS QUESTIONNAIRE

Page 126

### APPENDIX B. K-2 CAVE CREEK FLOATER STUDENT RATING OF TEACHERS QUESTIONNAIRE

Page 128

### APPENDIX C. 3-5 CAVE CREEK REGULAR STUDENT RATING OF TEACHERS QUESTIONNAIRE

Page 131

### APPENDIX D. 3-5 CAVE CREEK FLOATER STUDENT RATING OF TEACHERS QUESTIONNAIRE

Page 134

### APPENDIX E. 6-8 CAVE CREEK REGULAR STUDENT RATING OF TEACHERS QUESTIONNAIRE

Page 137

### APPENDIX F. 6-8 CAVE CREEK FLOATER STUDENT RATING OF TEACHERS QUESTIONNAIRE

Page 140

### APPENDIX G. 9-12 CAVE CREEK REGULAR STUDENT RATING OF TEACHERS QUESTIONNAIRE

Page 143

### APPENDIX H. ORIGINAL PROCEDURES FOR ADMINISTERING THE QUESTIONNAIRES

Page 146

### APPENDIX I. 50-ITEM BUBBLE SHEET STUDENT RESPONSE RATING FORM

Page 148
<table>
<thead>
<tr>
<th>Appendix</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>J</td>
<td>Revised Instruction Sheets for Administering Questionnaires</td>
<td>150</td>
</tr>
<tr>
<td>K</td>
<td>SPSS-X Computer Programs for Processing and Analyzing Data</td>
<td>153</td>
</tr>
<tr>
<td>L</td>
<td>Cave Creek Career Ladder Plan Algorithm</td>
<td>180</td>
</tr>
<tr>
<td>M</td>
<td>Human Subjects Release Letter from Dave Alexander</td>
<td>182</td>
</tr>
<tr>
<td>N</td>
<td>1990 Cave Creek Student Ratings of Teachers Analysis and Summary</td>
<td>184</td>
</tr>
<tr>
<td>O</td>
<td>1991 Cave Creek Student Ratings of Teachers Analysis and Summary</td>
<td>192</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1. A synthesis of the research literature for student ratings of teacher performance 41
Table 2. Questionnaire items developed by Cave Creek Public Schools 46
Table 3. Questionnaire items revised by SIM for Cave Creek Public Schools 47
Table 4. Participating students, teachers, and schools 49
Table 5. Analysis of variance for two groups with 15 subjects per group 55
Table 6. Item discrimination power of student feedback to teachers, 1990 6-8 regular questionnaire 64
Table 7. Item discrimination power of student feedback to teachers, 1990 6-8 floater special-area questionnaire 66
Table 8. Item discrimination power of student feedback to teachers, 1990 9-12 senior high questionnaire 68
Table 9. Item discrimination power of student feedback to teachers, 1991 6-8 regular questionnaire 70
Table 10. Item discrimination power of student feedback to teachers, 1991 9-12 senior high questionnaire 72
Table 11. Discrimination power of items developed by teachers and administrators at the local level, 1990 6-8 regular, 6-8 floater special-area, and 9-12 senior high questionnaire 75
Table 12. Discrimination power of items developed by teachers and administrators at the local level, 1991 6-8 regular and 9-12 senior high questionnaire 76
Table 13. Analysis of student mean score ratings of teachers by grade level 78
Table 14. Means and standard deviations of student mean score ratings by subject 79
<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 15</td>
<td>One-way analysis of variance of student mean score ratings by subject</td>
<td>79</td>
</tr>
<tr>
<td>Table 16</td>
<td>Means and standard deviations of student mean score ratings by questionnaire level</td>
<td>82</td>
</tr>
<tr>
<td>Table 17</td>
<td>One-way analysis of variance of student mean score ratings by questionnaire level</td>
<td>82</td>
</tr>
<tr>
<td>Table 18</td>
<td>Means and standard deviations of student mean score ratings by period of the day</td>
<td>83</td>
</tr>
<tr>
<td>Table 19</td>
<td>One-way analysis of variance of student mean score ratings by questionnaire level</td>
<td>84</td>
</tr>
<tr>
<td>Table 20</td>
<td>Analysis of student mean score ratings of teachers by core and elective nature of courses</td>
<td>85</td>
</tr>
<tr>
<td>Table 21</td>
<td>Analysis of student mean score ratings of teachers by career ladder participation status</td>
<td>86</td>
</tr>
<tr>
<td>Table 22</td>
<td>Analysis of student mean score ratings of teachers by gender</td>
<td>87</td>
</tr>
<tr>
<td>Table 23</td>
<td>Maximum likelihood factor analysis with varimax rotation for the 20 senior high school items</td>
<td>92</td>
</tr>
<tr>
<td>Table 24</td>
<td>Maximum likelihood factor analysis with varimax rotation for the 20 middle school regular items</td>
<td>93</td>
</tr>
<tr>
<td>Table 25</td>
<td>Comparison of items by factor analysis clusters and Judkins (1987) SIM performance areas, 1990 senior high questionnaire</td>
<td>94</td>
</tr>
<tr>
<td>Table 26</td>
<td>Comparison of items by factor analysis clusters and Judkins (1987) SIM performance areas, 1990 middle school regular questionnaire</td>
<td>95</td>
</tr>
<tr>
<td>Table 27</td>
<td>SIM performance area criteria</td>
<td>96</td>
</tr>
</tbody>
</table>
CHAPTER I. THE PROBLEM

Futurist Marvin Cetron (1989) states, "The high school graduating class of 2000 has already begun its odyssey through the American education system.... Its members entered kindergarten in September 1987" (p. 8). Consequently, the important question of whether this class receives the quality education it will need to be successful in the twenty-first century is directly linked to the nation's ability to improve the overall quality and effectiveness of teaching and learning.

Introduction

In an attempt to ensure that an affirmative response is recorded by those who will write the history of the Class of 2000, the issue of teacher performance evaluation must be satisfactorily resolved. It is, however, strongly debated. While some would argue whether the dominant strategies currently being employed actually improve teacher performance, many would agree with Furtwengler's (1991) assertion that "Educational leaders must never lose sight of the purpose of teacher evaluation: to improve the performance of teachers and thus the learning of students" (p. 787).

Schools have demonstrated a long tradition of viewing education as a student's privilege and of believing that success in school is largely dependent upon the initiative and ability of that student. Historically, educators have not viewed students' failure to learn or to receive a quality education as a flaw in the system of schooling. Similarly,
industry has, by tradition, ignored the possibility that declining productivity and profit could be the result of businesses having ignored the needs and wants of its customers.

Recently, however, businesses and corporations have been successful in actively pursuing the goal of achieving total customer satisfaction (Albrecht, 1990). Similarly, effective and excellent schools have rejected traditional attitudes toward students and have begun to assume more responsibility for ensuring that all boys and girls learn. Levine and Lezotte (1990) indicate that most of the research focusing upon such unusually effective schools supports the conclusions that these schools tend to rank high on specific functions commonly known as the effective schools correlates (p. 9). Further, Levine and Lezotte have determined that when educators in these schools are committed to providing students an education designed to achieve total customer/client satisfaction, the quantity and quality of learning increases.

This change from the traditional orientation toward that of the student as a client or customer of the school has served to focus attention upon the student as an important source of information regarding teacher effectiveness. Bailey (1983) states, "Teachers need to identify the different sources of feedback they can use for instructional improvement: 1) students, 2) self, 3) peers, 4) administrators or supervisors, and 5) parents.... Soliciting student opinion represents a potentially large sample or measure of information" (p. 6). Bailey goes on to say that, currently, most of the information concerning the teacher
and the quality of instruction is limited to the perceptions of one or two administrators.

The use of student ratings as one component of a total performance evaluation system, for teachers in the Kindergarten through Grade 12 (K-12) education system, is in its infancy stage of implementation. Feldman, citing Cashin, indicates that in the higher education system over "1,300 citations can be found in the Educational Resources Information Center data base on student evaluation of teacher performance" (Weimer, 1990, p. 5). Feldman also indicates that as few as 25 to 50 of these articles might contain directly relevant data and that for a specific sub-topic there may only be five or fewer relevant articles.

In this current review of literature fewer than 50 research articles directly relating to the K-12 education system were identified. Such a contrast is made even more noticeable when one compares both the quality of research and subsequent implementation of teacher performance findings at the K-12 and higher education levels.

This point notwithstanding, Professor Abrami cautions,

I think it would be unwise to make too much of an issue of the difference between, say, secondary and postsecondary teaching and learning. The similarities are both many and important. In fact, I have said in print a couple of times how postsecondary researchers might learn from researchers at lower levels and vice versa. (P. C. Abrami, personal communication, March 28, 1991)

Ironically, as both levels of the education system face increased scrutiny and accountability, each looks toward the other for insight regarding the process for improving instruction and teacher evaluation. Kenneth Feldman (1990) states,
From one point of view the quality [of the research] is high—hardly a surprising observation since the research has usually been done by professional researchers, who generally know what they are doing.... From another point of view, however, I do want to register certain complaints. More complex, multivariate analyses are needed, with increased emphasis paid to the possible causal connections among variables. (p. 6)

If such a need is evident in higher education then an even greater need for additional analyses is likely to exist at the K-12 level.

The universal and, in particular, the North American goal of improving education is directly linked to the improvement of the quality of teaching performance. Obtaining and using student feedback is one significant action that can contribute to the desired improvement(s). Efforts to improve quality of classroom instruction have resulted in some school districts and their superintendents, specifically David Alexander of Cave Creek Unified School District and Larry Bramblett of New Hampshire Public Schools, to move to the leading edge and incorporate the use of student ratings of teachers as one component of a total performance evaluation system.

According to Bramblett (L. R. Bramblett, personal communication, April 9, 1991), the traditional process of a single administrator conducting a limited number of classroom observations, for purposes of teacher evaluation, is just not enough. Bramblett further contends that the use of student feedback as an additional component of the teacher evaluation process provides both teachers and administrators an opportunity to obtain a comprehensive picture of teacher performance.

An important philosophical and practical educational goal involves securing the cooperation and support of teachers and their unions
regarding issues of student ratings of teacher performance. School districts using and implementing the Iowa State University School Improvement Model's (SIM) Total Performance Evaluation system have been successful in securing high levels of teacher support for the enhanced teacher evaluation process.

Specifically, school districts following the SIM Total Performance Evaluation system ensure that teachers account for approximately half of the total membership on the project's planning group called the stakeholders committee. Stakeholder committee members, including teachers, administrators, parents, and board members, are appointed by the superintendent and as a group serve at the pleasure of the board of education. Members of the stakeholders committee are told that an important committee task is that of "deciding to recommend" (Manatt, 1988) issues of planning, implementation, and evaluation criteria.

Bramblett (1991) believes that teacher support is obtained by providing teachers with a direct opportunity to participate in those decisions that affect them personally. They must also, according to Bramblett, be meaningfully involved in the development of the student feedback process.

When asked to comment on concerns raised as a result of the development and implementation of student ratings as one component of a total performance evaluation system, Superintendent Bramblett (1991) stated that teachers are legitimately nervous regarding whether student feedback questionnaires are as valid or reliable as they could or should be. Such an observation indicates that the increased acceptance and
subsequent use of student ratings as one component of a total teacher performance evaluation system is largely dependent upon the development of valid, reliable, and legally discriminating student ratings of teacher questionnaires (Abrami, 1990; Manatt, 1988).

In conclusion, noted researchers and practitioners share the common concern that America's need to improve the quality of teacher performance will continue to increase in magnitude. Consequently, additional sources of information such as student ratings may inevitably become legitimate components of the total teacher performance evaluation system.

Statement of the Problem

The desire and need to improve classroom instruction is a priority goal for the nation's political and educational leaders. L'Hommedieu (1990) believes that "Researchers who address the question of improving instruction through systematic feedback are exploring an issue of immense practical value" (p. 239). The current study addresses the issue of instructional improvement by attempting to establish the discriminating power of specific student feedback items used as one component of a total performance evaluation system. The analysis of data obtained from all secondary students and teachers in one school system will serve as a desirable extension of the original SIM directed work undertaken by Hidlebaugh in 1973, Judkins in 1987, and the ongoing SIM improvements to specific items and the questionnaires.

The problem for this study will be the identification of teacher performance evaluation rating items for use by students, based on item
discrimination power and grade level of instruction. First, selected items from the original Hidlebaugh and Judkins studies will be administered to all members of the Cave Creek (Arizona) Unified School District No. 93. Those items which continue to show discriminating power will be identified.

Second, district-developed items which possess discriminating power will also be identified. The Menne and Tolsma (1971) methodology of analyzing data will be used to identify those items possessing discriminating power (Hidlebaugh, 1973; Look, 1983; Judkins, 1987).

Third, the appropriateness of student feedback questionnaires as one component of a total teacher performance evaluation system will be assessed by determining the ability of items to discriminate according to criteria previously validated by Hidlebaugh (1973) and Judkins (1987). Items are judged to possess discriminating power when members of the same group provide similar responses and members of a second group provide different responses when the groups have experienced dissimilar conditions or events. In other words, item discrimination power exists when the within-group variance is low in relation to the between-group variance (Menne & Tolsma, 1971).

Fourth, a measure of reliability will be calculated to determine the amount of internal consistency of all items with discrimination power. Fifth, student ratings will be analyzed to show whether differences exist as a result of teacher gender. Sixth, a set of school district norms for middle school and senior high school student ratings of teachers will be established. Finally, for each teacher the relationship between the
variance of student rating scores generated by one class and the variance of student rating scores generated by all classes will be examined and described. Related factors such as grade level, subject, core/elective, career ladder status, and period of the day will also be examined.

The Research Questions

This study will attempt to identify and refine discriminating student feedback items for use as one component in a total teacher performance evaluation system. Student ratings of all grades 6-8 (regular) teachers, grades 6-8 (floater) special-area teachers including art, physical education, band, music, special education, and English as a second language, and grades 9-12 teachers in one school district will be used to determine discriminating power of individual student feedback items. This study can be specifically defined by the following research questions:

1. Do any of the items on the 6-8 regular, 6-8 floater, and 9-12 questionnaires possess discriminating power?
2. Is it possible for teachers at the local level to develop discriminating items?
3. What effect does (a) grade level, (b) subject area taught, (c) middle school level or high school level, (d) period of the day, (e) core versus elective nature of the course, (f) career ladder status, and (g) teacher gender have upon the mean score student rating for teachers?
4. When a factor analysis is conducted on the discriminating items, how many factors can be identified?
Definition of Terms

The following definitions will be used in order to add clarity and provide for more meaningful understanding of this investigation.

1. **Criteria** - a standard, rule, or test that can be used to judge performance based upon the research on effective teaching.

2. **Discriminating Item** - an item which separates high teacher performance from that of average and low performance. An instrument is considered to be most effective when it has a high level of item discrimination.

3. **Evaluation** - making a value judgment regarding the worth, quality, or effectiveness of the classroom teacher's instructional practice.

4. **Floater** - a teacher who offers instruction to students from several grade levels in a special curriculum area or subject (special education, languages, E.S.L., music, art, physical education, and band).

5. **Improvement of Instruction** - a series of steps that leads to an increased level of professional competence in the classroom.

6. **Mean** - the arithmetic average of the total student responses.

7. **Performance Criterion** - a specific teacher behavior that has been validated by the research on teaching.

8. **Rater** - a student who uses a questionnaire or feedback instrument to evaluate teacher performance.

9. **Rating** - an estimate of the degree to which a teacher has performed a given task or behavior.
10. **Reliability** - raters of a particular teacher consistently rate that individual teacher similarly on a specific item.

11. **Stakeholder** - a school community member who has a direct interest in the operation and effectiveness of the school.

12. **Student Feedback** - the process of collecting pupil information for the purpose of instructional improvement.

13. **Student Feedback Instrument** - a form or tool used to collect student opinions regarding teacher performance.

14. **Total Teacher Performance Evaluation System** - a process that incorporates the use of valid, reliable, and legally discriminating criteria for the purpose of combining improvement of instruction and accountability in a single evaluation system. This comprehensive process is intended to assist teachers in the enhancement of instructional performance and is based upon a belief that feedback from each of the individual component groups adds important information about unique aspects of teachers' performance. Desirable sources of information are: (a) supervisor evaluation, (b) self-evaluation, (c) peer evaluation, (d) student achievement scores, (e) student feedback, and (f) parent feedback.

15. **Valid Instrument Items** - items that measure what they are intended to measure.

16. **Variance** - the relationship of scores to a central value, such as the mean. Variance is the sum of squared deviations around the mean. It describes how similar or different, for a given group, the scores are from the mean.
Note: For purposes of editorial style and maintaining reader interest, the Cave Creek (Arizona) Unified School District No. 93 will also be referred to as Cave Creek, Cave Creek Public Schools, Cave Creek School District, or Cave Creek Unified School District. Further, the terms floater and special-area will be used interchangeably.

Delimitations of the Study

This study of the ability of teacher-rating instrument items to discriminate when administered to a total school population inherently possesses the following delimitations:

1. This study was conducted in Cave Creek (Arizona) Unified School District No. 93. Schools in the district are organized into four building units: Cave Creek School (K-2), Black Mountain Elementary School (3-5), Desert Arroyo Middle School (6-8), and Cactus Shadows High School (9-12). In 1990, 18 teachers were rated by 383 (119 grade 6, 134 grade 7, and 130 grade 8) students from Desert Arroyo Middle School, and 27 teachers were rated by 443 (129 grade 9, 113 grade 10, 109 grade 11, and 92 grade 12) students from Cactus Shadows High School.

The data for this study were collected in May of 1990. Some students rated teachers for year-long courses and others rated teachers for courses taken during the second semester of the school year. Student feedback questionnaires were used by all students to rate their teachers.

2. The student ratings questionnaires were administered and the data were collected by the school district staff. It is assumed that the written procedures and guidelines developed by the School Improvement
Model and by the researcher were implemented, in an appropriate manner, by the school district personnel (see Appendix H).

3. An intact stakeholders committee determined the amount of demographic information that was recorded for students and teachers. Since information such as student gender and student achievement scores was not available, it was not possible to examine the relationship of these factors to the available student ratings.

4. Only one school district's grades 6 through 12 students and teachers, grouped by classes, were used as the study sample population. The student raters were those individuals who comprised the teacher's intact class for a specific period of an instructional day. Each student rated several different teachers as a result of the student's instructional course timetable.

5. The data collected were used as an integral component of an existing career ladder merit pay system. Approximately 50 percent of the teachers in the study were on the school district's career ladder plan. Inherent in the career ladder was the motivation that teachers receiving high student ratings would be granted higher salaries than teachers receiving lower student ratings. Career ladder teachers may possess increased motivation and therefore demonstrate differing behaviors toward students than do non-career ladder participants.

6. Issues surrounding validation of the student feedback instrument items were not included in this study. Item validation was established as a part of the Hidlebaugh (1973) and Judkins (1987) research studies. These findings were accepted for the purposes of this investigation. A
simple factor analysis, however, was used to determine whether the items contained in the student ratings of teacher questionnaires continued to "load" on the specific criteria factors identified by Judkins.

7. The student rating questionnaires included 20 discriminating items. Each student response was weighted from 0 to 4 points yielding a possible total-rating score of 80. For each teacher all of the individual student total-rating scores were averaged to calculate a total mean score rating. This total mean score was used for the purpose of showing whether differences among teachers existed according to the variables associated with this study.

8. In 1990 teachers and administrators in the Cave Creek (Arizona) School District developed and administered three separate student feedback questionnaires. Many items are consistent between questionnaires. However, some differences exist due largely to an initial concern of special-area teachers that the student feedback questionnaires designed for regular education teachers were not appropriate for special programs. Consequently, grades 6-8 special-area teachers developed a questionnaire based on locally-developed criteria considered to be specific to the evaluation of special-area teacher performance. In 1991 the stakeholders committee recommended that one questionnaire be used for all grades 6-8 special and regular education teachers and that the original 1990 questionnaire continue to be used for all grades 9-12 teachers.
Human Subjects Release

On October 25, 1990 a letter authorizing this research was written to Professor Manatt by Dr. David Alexander, superintendent for Cave Creek Unified School District No.93. The data set being used as the basis for this investigation is district property. The superintendent provided authorization to use the data set for this investigation (see Appendix M).

The Iowa State University Committee on the Use of Human Subjects in Research reviewed this project and concluded that the rights and welfare of the human subjects were adequately protected, that risks were outweighed by the potential benefits and expected value of the knowledge sought, that confidentiality of data was assured, and that informed consent was obtained by appropriate procedures.
President Bush and the nation's governors have agreed upon six challenge goals for education. Harp (1990) states, "The national goals for the year 2000 include decreasing the dropout rate to 10 percent, raising U.S. students to the top rung of international mathematics and science achievement, preparing all children for the 1st grade and improving adult literacy and continuing education efforts" (p. 29). "The National Governors' Association, chaired by Iowa Governor Terry Branstad, and the Council of Chief State School Officers are marshalling resources to move each state toward achieving the ambitious goals" (Bartusek, 1990, p. 1).

Challenges and Opportunities

Stiggins and Duke (1988) state, "Teacher evaluation is the key to school improvement. As we pursue excellence in education through the promotion of professional development of teachers, we cannot overlook the potential contribution of the teacher evaluation process to that development" (p. xi). In the Iowa Business and Education Round Table Task Force's report (Hornbeck, 1991), the goal of developing "World Class Schools" is presented as a challenging school improvement initiative with positive attention focused upon the single most important unit in a school system--the classroom and its students (p. 4).

The future ability of America to compete, grow, and prosper (Hornbeck, 1991) is directly linked to the development of world class
Across the nation no single focus can be of more concern than the classroom as the center for the delivery of instruction for students. The move to create world class schools assumes a need to concentrate upon the classroom and teacher effectiveness. With classrooms and student performance having become key elements in school improvement initiatives, the nation's focus has also shifted to the importance of teacher performance and subsequent teacher evaluation as key functions which deserve and demand additional attention (Hunter, 1988).

In a response to increased demands for proven methods of teacher evaluation, this study will continue the work initiated by Hidlebaugh (1973), who created a model for developing a teacher performance evaluation system using a multiple-appraiser approach, and the more recent work of Judkins (1987), who studied and identified discriminating items for the student evaluation of teachers and the ongoing efforts of the School Improvement Model (SIM) team of Iowa State University. The SIM team and local school districts continue to collaboratively develop processes using student feedback instruments as one component of a total teacher performance evaluation system.

In 1987 Judkins cited reports such as, "A Nation at Risk: The Imperative for Educational Reform" and "Time for Results, The Governors' 1991 Report on Education," as evidence of the growing need to improve America's schools. Throughout the nation, the continent, and the world, public leaders are searching for ways to develop and assess world class education. In "America 2000," President Bush (U.S. Department of Education, 1991) remarks that "We must challenge not only the methods and
the means that we've used in the past, but also the yardsticks that we've used to measure our progress" (p. 3). Increased use of valid and reliable student ratings of teachers (Feldman, 1990) may be one additional strategy that promotes the attainment of world class schools.

The review process

A review of the literature related to student ratings of teacher performance reveals an extensive body of information associated with higher education and a very limited amount of information related to K-12 education. The literature findings included in this review range in quality from researcher insight based upon experience to carefully formulated judgments which result from rigorous and extensive research studies.

The focus of this review was to locate those sources of information which directly refer to the role of students as raters of teacher performance. A secondary screen was used to search for information directly related to the K-12 school system. Finally, those sources which specifically examined the areas of age, grade level, gender bias, subject taught, and time of day were critically reviewed.

The review process was initiated by examining studies directly related to the current investigation, such as the research of Hidlebaugh (1973) and Judkins (1987). A deliberate attempt was made to locate appropriate, new, and different information from that which had previously been reviewed. Additionally, the present investigator sought and examined refereed or professional association journal reports of noted authorities.
These sources were identified and located by accessing the ERIC System, Dissertation Abstracts, Library Indexes, Educational Administration Abstracts, Handbook of Research on Teaching, and personal interview contacts.

After having examined and reviewed a number of studies, those which were judged to have reported findings based on direct empirical data and those which had a clear and readily understood methodology were assigned the highest levels of credibility and importance. Specifically, the following factors were considered when examining and evaluating research studies: 1) clarity of purpose and stated hypothesis, 2) researcher or sampling bias, 3) sample size, 4) control of important variables, 5) measurement techniques that yield reliable and valid results, 6) statistical tests, and 7) quality of statistical analyses.

In order to place this study into the context of current events, a review of such national and international issues as school improvement, teacher performance, and changing student roles was conducted by referring to popular publications and agency reports. Additionally, citations obtained from initial sources were identified and examined.

Most studies included in this review were conducted in North American school systems. During the search and review process efforts were made to present research which was both supportive and non-supportive of using students as raters of teacher performance. Nevertheless, the majority of research reported tends to support the role of students as raters of teacher performance.
Student as knowledgeable client

Recent demands for school improvement and reform have created an opportunity to consider the interaction of students and their teachers from the perspective of the student being in the role of a knowledgeable client. From this perspective, the purpose of this study was to examine the item discriminating power of three student ratings of teacher questionnaires. One questionnaire was intended for use with regular middle school teachers and students, the second questionnaire was designed to be used with special-area middle school teachers and students, and the third questionnaire was to be used with all senior high school teachers and students.

A non-voluntary sample

A unique feature of this study was the collection of data from all teachers in a school system rather than a select sample of volunteers. Cave Creek Unified School District No. 93 is one of 14 school districts in the state of Arizona to have adopted a pay for performance system for its teachers. It is the only school district in the state to include student ratings as an integral part of the algorithm used, during the 1990 and 1991 school years, to place teachers on the career ladder. Therefore, student ratings were collected and analyzed for all teachers and students in the school district.

The use of a total system sample may be the most important element of this study. In both the 1973 Hidlebaugh study and the 1987 Judkins study the subjects were volunteers. Hidlebaugh originally developed a pool of
discriminating teacher evaluation items for use by principals. Obviously, normal school conditions limit the number of school principals who are likely to observe an individual teacher perform. Therefore, when the research design and data analysis procedures required a minimum of 15 raters (Menne and Tolsma, 1971) to assess the performance of each teacher, Hidlebaugh used students as additional evaluators. In this manner the items were validated by a sufficient number of raters who collectively comprised a multiple evaluator system for teacher performance.

Subsequently, Judkins continued the work initiated by Hidlebaugh and refined a pool of items, based on valid and reliable criteria, which discriminated between levels of teacher performance. The research on teaching was used to validate the criteria. Judkins developed a valid, reliable, and discriminating set of teacher evaluation instrument items suitable for use by secondary and elementary students.

The current investigation and an associated K-5 study (Weber, 1992) continues to refine the work of Hidlebaugh (1973) and Judkins (1987). However, unlike the previous studies of Hidlebaugh and Judkins, this investigation provides the first opportunity to examine the discriminating power of items used with all classes, students, and teachers in a single school district. By using a total school sample rather than a select sample of volunteers, this study will determine whether or not student rating questionnaire items from the original Hidlebaugh and Judkins studies continue to be discriminating.
A matter of reliability and validity

The chances of improving the reliability and discriminating power of each student rating questionnaire as well as the individual items which compose each questionnaire increases as a result of a larger variance in the students’ ratings of teacher performance. In this study, increased reliability and variance were expected over that which was concluded by Judkins, as a result of the total sample including teachers who demonstrated a wider diversity in levels of teaching performance.

Obviously, issues of reliability and validity are important considerations when examining the ability of items to discriminate varying levels of teacher performance. Most research investigations involving student ratings examine these issues. For example, L’Hommedieu, Menges, and Brinko (1990) conducted a meta-analysis designed to integrate studies and to examine variables that might moderate the measured effect of student rating feedback. With a small effect size (.342), indicating a positive effect for feedback but also indicating limited practical significance, the researchers examined the possible threats to validity and reliability in student ratings research.

Focusing upon 28 studies which met the meta-analysis selection criteria, the researchers began by examining the area of internal validity. The inability of instruments to discriminate was cited as a common limitation of most Likert scale instruments. In almost all pre-posttest measures examined, the pre-test score was above the mid-point and left little room for improved ratings. This factor limited the measurement of effects of student ratings feedback.
One specific conclusion of this meta-analysis, of particular importance to the current investigation, is the recommendation that researchers must consider using rating systems with tested reliability and validity estimates. For example, of the 28 studies reviewed only three instruments were used more than once. The researchers (L'Hommedieu, Menges, & Brinko, 1990) conclude that "Tests of the utility of student ratings feedback for instructional improvement are more easily interpreted if the results are not confounded with the reliability and validity threats associated with new rating systems" (p. 234). In the current study, the continued refinement of the discriminating items for student ratings of teacher performance, initially identified by Hidlebaugh and further developed by Judkins, is highly consistent with L'Hommedieu's, Menges' and Brinko's conclusions regarding the need for replication.

The Menne and Tolsma methodology

Establishing discriminating power of items is an essential step in the development of valid and reliable student ratings of teacher questionnaires. The basic design of this study will follow and continue a well-established and credible methodology.

In 1973 Hidlebaugh used the Menne and Tolsma (1971) methodology to identify discriminating items for use in teacher performance evaluation. Look (1983) used this same methodology to identify effective criteria for the evaluation of school principals. In 1987 Judkins also used this methodology to identify discriminating items for use in the student evaluation of teachers. Most recently, Green (1990) applied the same
method of using knowledgeable multiple raters to identify valid, reliable, and discriminating criteria for use in substitute teacher evaluation instruments.

When included as one component of a total teacher performance evaluation system, student ratings can be used to provide feedback to teachers and to show differences in the quality of instruction. Menne (1972) contends that differences in performance quality and effectiveness can be measured when three conditions are present: 1) there must be more than one rater; 2) the raters must closely agree on their ratings; and 3) the ratings must indicate differences between the persons rated.

In discussing the first and second conditions respectively, Menne (1972) concludes,

[First,] there is no check on a single rater that would give evidence that he is even rating the performance aspect intended. [Second,] if all raters indicated that a given teacher rated a score of four out of a possible five points on some performance aspect, such as "well prepared for class" then this consistency of raters indicates something may have been measured. On the other hand, if the ratings of the same teacher varied one to five, then nothing has been measured--the average rating in such a situation would be a misleading statistic. Therefore, there must be a consistency or a low variance between raters. (p. 5)

In order to serve the purpose for which they are intended, student feedback questionnaire items must possess item discrimination power. In other words, the items must yield ratings which describe differences between the teachers being rated. Menne (1972) clarifies and illustrates this point by explaining,

If all students in a class were asked the sex of their teacher, there would be consistency (low, or in this case, zero variance) in the responses of this group of student "raters." Other classes, with the same or different students, should also have a consistent response to the question. But if the teachers are
not of the same sex, there will be a difference in the responses between classes. So teacher ratings must be consistent and also must indicate differences between the performance of different teachers.

The present study will use the methodology described by Menne and Tolsma to identify discriminating items for use in student ratings of teacher performance. Because student ratings are one measure of teacher performance (Menne, 1972), an intended result of this study was to determine whether the conditions necessary for meaningful measurement are being met.

Formative or Summative Evaluation

McGreal (1983) argues that student ratings should not be used for purposes of summative teacher evaluation. He states,

While attitudes regarding the value of student ratings vary, the average elementary and secondary teacher is uncomfortable with the concept. Teachers generally lack faith in the students' ability to accurately rate their performance. In many respects their fears are justified. There is not a great deal of support for the accuracy of student ratings, and the support that does exist is not strong enough to justify using student ratings in any summative sense. (p. 134)

The practice of including students as judges of college and university faculty performance has continued to grow in popularity since its initial development at Harvard University in 1924 (Ghorpade & Lackritz, 1991). Nevertheless, fundamental questions regarding the validity and reliability of student evaluation of faculty continue to be asked.

Specifically, Ghorpade and Lackritz (1991) suggest caution regarding serious consequences which may occur when student ratings of teachers are
used to influence decisions regarding retention, tenure, or promotion of minority faculty. Reports of increased racial and gender prejudice, tension, and violence are cited as reasons to "scrutinize student ratings of faculty for the presence of bias" (p. 63).

Ghorpade and Lackritz (1991) designed a study to assess student ratings for the presence of bias. From the outset a qualification was made, i.e., that the purpose of the investigation is not to develop a definitive case of discrimination against student ratings but rather to scrutinize student ratings of faculty from an Equal Employment Opportunity perspective.

The study utilized data from an existing published source as well as data collected through a student-developed questionnaire. Faculty teaching undergraduate classes with enrollments of 20 or more students had the option of participating. While the study design appropriately attended to a desirable number of raters, the self-selection process of faculty and students volunteering to participate was a weakness in methodological design. Nevertheless, 20,242 students (67 percent of the total enrollment in the classes) returned the questionnaire. By using a factor analysis and a subsequent analysis of variance procedure, Ghorpade and Lackritz made the following conclusions which are relevant to the current investigation: 1) Female faculty attained significantly higher mean student rating scores than males (N=412) on Factor 1 [seven general assessment questions] (3.35 to 3.26**); and 2) female faculty received significantly higher mean student rating scores than males on Factor 10 [overall rating as an instructor] (3.36 to 3.27*). In other words, it was
concluded that a gender bias existed among the student ratings of faculty. An unanswered question remains as to whether the ratings can be explained by actual differences in teaching effectiveness and performance.

While Ghorpade and Lackritz found gender bias in student ratings, Aleamoni (1987a) presents a somewhat contradictory view of male/female bias. Additionally, he cites little relationship between student ratings and several other critical variables which are a part of the present investigation.

The majority of the research I have looked at indicates little or no relationship between such variables as class size, gender of the student or gender of the instructor, the time of day that the class is offered, the major or non-major status of the student or the term or semester that the course is offered and the way in which students rate a course or instructor.

Aleamoni concludes that course level and age or grade level of students are reported as affecting, or more appropriately being associated with, student ratings of teachers.

Paradoxically, as critics and supporters of students as raters continue their debate, one recent study suggests that students may not value involvement in the process (Weimer, 1991). In a study conducted at the University of Washington, a team led by Robert D. Abott used a 2x2x2 experimental design to test student satisfaction. After having collected students' opinions regarding their instructors, the researchers concluded that students were more satisfied when the instructor, following the process, shared comments and reactions concerning the evaluation. Students were least supportive of the traditional means of collecting student opinion when rating forms were used at the end of the course and when no response was given by the instructor.
In summary, concerns exist that would limit the role of student ratings solely to support summative evaluation of the teacher. Nevertheless, a further review of the literature will illustrate that not all researchers and practitioners share the concerns regarding the accuracy of student ratings, the presence of gender bias, or the limited value students attach to the rating process. The research literature is supportive of the conclusion that the seriousness of such concerns is reduced when applied to the college and university level. Murray (1987) concludes, "Evidence from five different sources, namely logical argument, personal observation, faculty surveys, field experiments and longitudinal comparisons supports the conclusion that student instructional ratings have had a positive impact on quality of teaching in higher education" (p. 13).

In addition to the application of logical argument and personal observation, systematic research findings are used to show that student ratings positively influence teaching performance. In a 1980 survey study carried out by Outcult, 67 percent of 4,468 respondents indicated that student ratings had helped to improve their teaching and 78 percent reported they had actually made changes in their teaching as a result of student ratings.

Similar findings were reported by Murray (1987) from a University of Western Ontario, Canada study where 54 percent of the faculty indicated that global student ratings provided useful feedback and 78 percent said that student ratings of specific teaching behaviors were valuable for feedback purposes. Unfortunately, weaknesses, such as unknown sample size
and unsubstantiated inferential judgments, inherent in the cited field experiments and longitudinal studies serve to mitigate against this study's findings. Nevertheless, the findings of additional studies support Murray's conclusions (Abrami, 1990; Aleamoni, 1987b; Braskamp, Brandenbury, & Dry, 1984; Manatt, 1988; and Stiggins & Duke, 1988).

A valuable source of feedback

In a similar fashion, Scriven (1990) adds,

It should be mentioned that student ratings, if gathered in a suitable secure way, using a suitably designed form, can provide a useful basis for rating teachers. They are extensively used for this purpose at the tertiary level, and some tertiary institutions have been using them for 40 years. Their use is still rare at the secondary, let alone at the primary level, but they should work there if students and teachers are prepared for their use. In general, they are worth a great deal more than the usual reports from occasional visitors, whose reports suffer from samples that are inadequate in size and not representative, measurement artifacts, style bias, and failures of empathy and are usually vulnerable to personal bias. (p. 91)

Aleamoni (1987b) has developed a useful listing of myths which he judges to be untrue regarding student ratings of teacher performance. The following selected myths are related to questions to be examined during the current study:

- Students cannot make consistent judgments about the instructor and instruction because of their immaturity, lack of experience, and capriciousness.

- Student ratings are both unreliable and invalid.

- The time and day the course is offered affects [is associated with] student ratings.

- Students cannot be meaningfully used to improve instruction.
• The level of course affects student ratings.
• Whether students take a course as a required or elective affects student ratings.

An appropriate response to such myths is dependent upon the findings of well-designed and rigorous research studies. Specifically, the ability of students to provide ratings of varying levels of teacher performance is dependent upon the development and identification of valid, reliable, and discriminating items.

A non-traditional variable

Shepherd and Trank, in a 1989 study, attempted to show that student construct system development is as deserving of research attention as other input variables such as class size, gender, level of course, subject, and required or elective nature of the course. The study involved 431 students enrolled in 28 sections of "Rhetoric" at a large Midwestern university.

Student descriptors of "good and bad" teachers were used to provide measures of construct system development. The research team employed factor analyses and three pairs of hierarchical regression analyses to test the research hypotheses. The hypothesis which tested relational effectiveness as a function of construct differentiation yielded a significant interaction (t=-2.18, df=425, p=.030). When testing construct differentiation and task effectiveness in predicting relational
effectiveness, a second significant interaction was found \( t = -2.62, \) \( df = 425, p = .030 \).

The researchers concluded that the findings confirmed the influence of the tested student input variable, construct differentiation, on teacher evaluation. Although this study was limited in that it focused upon first-year students enrolled in one course and did not account for varying degrees of student performance or ability, it does indicate the possible need to consider the impact of student ratings resulting from how students perceive task and relational dimensions in the classroom.

In examining the student rating of teacher performance research, cited thus far in this review, the issue of whether student rating scores are sufficiently credible for use in summative evaluation remains unresolved. Therefore, in designing a suitable form or questionnaire for student ratings of teacher performance, the developers should be focused upon the identification of instrument items which possess discrimination power. By using the Menne and Tolsma (1971) method for identifying discriminating items and also by drawing upon many of the valid, reliable, and discriminating items previously identified by Hidlebaugh and Judkins, this study will attempt to provide information useful in answering the question of whether student feedback is best used in a summative or formative manner.

Original SIM Efforts

The current investigation is the third phase of an ongoing SIM effort to refine the role of student ratings as one component of a total teacher
performance evaluation system. In the Hidlebaugh study, 69 elementary and secondary teachers were rated by 1,140 students, 207 peer teachers, and 20 administrators in Naperville (Illinois) Community District.

A total of 94 of the original 360 items were found to be appropriate and to discriminate between teachers when students, peer teachers, and administrators rated selected teachers. According to Hidlebaugh (1973), "The bulk of the empirical evidence indicates that student ratings are probably the best single indicator of a teacher's true performance" (p. 96).

Judkins (1987) laments, "One valuable source of information on teacher performance, students, is rarely used at the secondary and elementary level, although the advantages of student ratings have been thoroughly researched at the college level" (p. 2). After reviewing this research on teacher performance and student ratings, Judkins decided that a need remained to develop and identify items that would effectively discriminate teacher performance. He developed a pool of items, based upon current knowledge and research, for use by secondary and elementary school students.

All students and teachers who participated in the Judkins (1987) study were volunteers from school systems in Springfield, Missouri and El Dorado Hills, California. Data were collected in February and March 1986 from 3,560 students and 188 teachers. Actual data used in the statistical analyses were from 2,919 students and 132 teachers. Judkins reported the use of volunteers as a limitation of the 1987 study. The current
investigation addresses this concern by using a total school system's population.

In the Judkins study, discriminating items were developed for three performance areas generally classified by SIM as Productive Teaching Techniques, Organized, Structured Classroom Management, and Positive Interpersonal Relations. The factor analysis conducted in this current study will show how the discriminating items "load" or cluster in each of these areas.

The number of discriminating items identified by Judkins for the four levels of questionnaires ranged from 18 to 52: 25 in grades K-2, 18 in grades 3-6, 34 in grades 7-8, and 57 in grades 9-12. Cronbach alpha reliability coefficients were calculated to determine the internal consistency of all items with a discrimination value equal to or exceeding 13 percent. The coefficients were .859 for grades 9-12, .839 for grades 7-8, .650 for grades 3-6, and .596 for grades K-2.

In the current study, items will continue to be identified as discriminating according to the criteria and procedures employed by Judkins. Cronbach alpha reliability coefficients will also be calculated and reported to determine the degree of internal consistency of items and to facilitate cross-study comparisons.

In reflecting upon the findings of the 1987 study, Judkins concluded, "Students from Kindergarten through the twelfth grade are capable of providing student feedback to teachers that discriminate among teachers" (p. 107). The original pool of discriminating items has provided the
Students hold a unique perspective

Both the initial 1973 Hidlebaugh study, which designed a pool of teacher evaluation questionnaire items for use by principals but also included students as raters, and the similar but more recent 1987 Judkins study have made a valuable contribution to the field of student ratings and teacher performance evaluation. With a similar focus on students as one component of a total performance evaluation system, many researchers and practitioners believe that students can provide reliable evaluative information about classroom instruction (Doyle, 1975; McKeachie, 1980; Manatt, 1987; Millman, 1990).

In fact, Stiggins and Duke (1988) conclude that students may be in a better position than anyone else to make judgments regarding a teacher's effectiveness in the classroom:

Students can provide useful data on other aspects of instruction besides their own achievement. For example, no one is in a better position to comment on the clarity of teacher directions than the students for whom the directions are intended. Students are the only observers who are in class on a regular basis. As a result they are in a unique position to comment on important dimensions of their learning environment. As long as they are not asked to comment on aspects of teaching for which they have no expertise, students constitute a rich source of data for professional development. (p. 139)

Stiggins and Duke summarize their views by stating that as a result of their unique perspective for observing teachers, students should be included as a valuable information source in a total teacher performance evaluation system. They further contend that it is time to stop debating
whether students should be given the opportunity to provide feedback regarding teacher performance. Additionally, researchers and practitioners need to focus their time and energy toward developing a highly effective process for facilitating the collection, interpretation, and appropriate use of student ratings of teachers.

A non-voluntary sample

Two limitations cited by Judkins in his 1987 study were, "(5) only teachers and students in selected volunteer schools in Springfield, Missouri, and El Dorado, California were involved in the study; and (7) to obtain the approval of the Human Subjects Research Committee to conduct this study, subjects were permitted to not return the feedback questionnaires if they chose not to participate in the data analysis" (p. 7). In responding to this limitation, the present study will use a data sample that represents the student ratings for all secondary students and teachers for one complete school system.

The opportunity to conduct this investigation is a result of two highly desirable circumstances existing at the same time. First, the work of Hidlebaugh in 1973 and Judkins in 1987 have resulted in the existence of a well-established collection of criterion-referenced student feedback items which can be examined in the present study. A second vital factor was the willingness of one school district to involve all students as raters of teachers and to rate all career ladder and non-career ladder teachers using the same process and the same student feedback questionnaires.
The desirability of further assessing the discriminating power of the questionnaire items with a total non-volunteer sample was identified by Judkins as a purposeful and worthy research pursuit. Consequently, this study is intended to add additional information and perspective to the existing work on Teacher Performance Evaluation conducted by the School Improvement Model.

A Total Systems Approach

Manatt (1987) contends that teacher performance evaluation is a complicated but necessary process. School improvement, he continues, is contingent upon changing how teachers perform, and in order for educators to begin changing the way that teachers perform, the profession must view school improvement as a process, not a product.

Core to this process are four key questions which must be asked whenever performance evaluation is being discussed: "1) What are our criteria? 2) How high are our standards? 3) How should we monitor and report progress? 4) How shall we help the evaluatee improve after we have identified a profile of strengths and weaknesses?" (Manatt, 1988, p. 86). These simple, yet profound, questions serve to sharply and appropriately focus the process of a total systems approach to teacher performance evaluation.

According to Manatt, the component parts of a process approach to teacher performance evaluation will include, formative pre-observation conferences, scheduled and unscheduled classroom observations, post-observation conferences, portfolio data, a summative report, a summative
conference, a professional growth plan, a student achievement report, and a student feedback summative report. Each component part of the total performance evaluation system is a simple yet necessary source of data for effective teacher evaluation, in that the systems approach seeks to build as solid a foundation as possible from which to make accurate summative judgments. Manatt further states that many teachers support the use of multiple sources of teacher evaluation data and often will identify traditional single-indicator approaches as serious problems in performance assessment.

In explaining the advantages of multiple sources of data for teacher performance evaluation, Manatt (1988) advocates shoring up the traditional single data gathering procedures by adding feedback from students. In fact, "Student ratings of teachers were found to be more discriminating than any other source of teacher data once proper instruments were developed" (Manatt, 1987, p. 10). By employing a total systems approach to teacher performance evaluation, the School Improvement Model addresses the limitations of each single measure by utilizing the data obtained from other sources.

More is better

The practice of using multiple sources of data for teacher evaluation has been widely accepted at the university level. In writing about faculty evaluation for higher education, Braskamp et al. (1984) state, "Information collected from a number of sources and by a variety of methods, each reflecting a diversity of criteria, is the ideal for
obtaining a fair and credible assessment of faculty teaching competence.... Sources include self, alumni, students, records and colleagues" (p. 33).

Current research, then, appears to have established that multiple sources of data including student ratings, add to the credibility of a total performance picture. Student raters, however, are often judged according to the reliability and validity of their ratings. In the current study, validity of the rating criteria has been established by using the research on teaching.

Reliability refers to the extent that the ratings yield information which is free from biases due to characteristics of the students, courses, and time of administration. Reliability consists of measuring agreement among student ratings for a teacher within a given class, and the consistency of the rating questionnaire to yield similar scores for a specific teacher and class at different times. The current investigation will examine the effect of these factors on student ratings of teacher performance.

In conclusion, reliability of student ratings is based upon four generalizations from the research literature:

1. Student agreement on global ratings is sufficiently high if the class has over fifteen students.
2. Students are consistent in their global ratings of the same instructor at different times in the course.
3. An instructor's overall teaching performance in a course can be generalized from ratings from five or more classes taught by the instructor in which at least fifteen students were enrolled in each class.
4. The same instructor teaching different sections of the same course receives similar global ratings from each section (Braskamp et al., 1984, p. 56).
Developing a Complete Picture

In order to properly produce a commercial color brochure, a printer must ensure that each of the base or primary colors is present. If any one of these colors is omitted, the results are distorted and less than satisfactory. It can be argued that the same basic premise must be considered in teacher performance evaluation.

Duke and Stiggins (1986) identify two goals for teacher evaluation systems. One is the support of personnel management decisions. A second and equally important goal for teacher evaluation is the improvement of instruction by promoting the professional development of teachers.

In commenting upon the student as a potential evaluator of teacher performance, Duke and Stiggins (1986) conclude,

Student evaluation of teacher performance might be suspect in a termination hearing. Participants could regard students as easily influenced, biased, or unqualified to judge minimum competence. There may, however, be no more valid source of information on and criticism of learning environments than students who live and work in those environments. When their views are sought in careful, thoughtful ways in evaluation systems designed to promote teachers' continued growth, students can provide insights no one else can. (p. 31)

As the evaluation of teacher performance becomes a more important component in the improvement of educational programs being offered to students, the need to improve the quality of data gathered from all sources will also increase. Student feedback is one desirable and important source of information regarding teacher performance (Aleamoni, 1987a; Costa, 1988; Doyle, 1975; Manatt, 1987; McGreal, 1988; Popham, 1988).
Although McGreal (1988) argues against the use of student ratings in summative teacher evaluation, he also concludes, "The use of student evaluation of teachers has long been a recommended source of data about teaching. Like most other sources discussed in this commonality there would seem to be considerable logic in asking students to provide evaluative judgments about their teachers" (p. 133). Soldin (1980), in a discussion of performance evaluation of professors, states, "It is manifestly true that the only direct, daily observers of a professor's classroom teaching performance are the students in the classroom. Students are thus a potentially valuable source of student judgmental information in the overall evaluation of the professor" (p. 36). He further addresses the potential for students' involvement in a changing educational system and states that as educators prepare to transform America's schools to meet the challenges of the twenty-first century, students should be provided with a legitimate role in the improvement of teacher performance.

An Informed Client

Almost all recent advocates of teacher performance evaluation systems suggest that one source of information can appropriately come from students' evaluations of their teachers (Braskamp et al., 1984; Herman, 1973; Manatt, 1987; McGreal, 1988; Popham, 1988; Stevens, 1987). But as Harris (1986) correctly observes, "The use of data on teaching from students has had only limited acceptance despite obvious potential"
(p. 103). In the views of many educators, this lack of acceptance is unfortunate. It is the students, after all, who have the most extensive opportunities of all potential observers to view the teacher's performance and also see a variety of teachers and performance levels over time.

Table 1 is a synthesis of the student feedback research presented as a part of this current study.

Conclusion

Clearly, the studies conducted in higher education demonstrate that most concerns regarding the inability of students to accurately assess teacher performance are overstated. Nevertheless, continued research focusing upon the ability of students in public schools to provide useful feedback regarding teacher performance is needed if students are to be a legitimate source of data in a total teacher performance evaluation system.
Table 1. A synthesis of the research literature for student ratings of teacher performance

<table>
<thead>
<tr>
<th>Source researcher</th>
<th>Students as raters</th>
<th>Level grade/age</th>
<th>Concerns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hiddlebaugh</td>
<td>Student ratings are probably best single indicator of teacher’s true performance</td>
<td>Secondary 7-12</td>
<td>Valid and reliable</td>
</tr>
<tr>
<td>Judkins</td>
<td>Students can provide feedback that discriminates among teachers</td>
<td>K-12</td>
<td>College University</td>
</tr>
<tr>
<td>McGreal®</td>
<td>Should not be used for summative evaluation but supports formative</td>
<td>K-12</td>
<td>College University</td>
</tr>
<tr>
<td>Murray</td>
<td>Positive impact on quality of teaching</td>
<td>College University</td>
<td></td>
</tr>
<tr>
<td>Doyle</td>
<td>Students can provide useful information</td>
<td>K-12</td>
<td>College University</td>
</tr>
<tr>
<td>Feldman</td>
<td>Student ratings are multi-dimensional, reliable, uncontaminated, valid, and useful</td>
<td>College University</td>
<td></td>
</tr>
<tr>
<td>Manatt</td>
<td>Student ratings more discriminating than any other source of data</td>
<td>K-12</td>
<td></td>
</tr>
</tbody>
</table>

*Non-supportive position regarding use of student ratings.*
<table>
<thead>
<tr>
<th>Source researcher</th>
<th>Students as raters</th>
<th>Level grade/age</th>
<th>Concerns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stiggins Duke</td>
<td>Provides useful data on the other aspects of instruction other than achievement: multiple sources</td>
<td>K-12</td>
<td>Ratings may be suspect in a termination hearing</td>
</tr>
<tr>
<td>Soldin</td>
<td>Students are potentially valuable source of judgmental information in the overall evaluation of a professor</td>
<td>College University</td>
<td>Support and coaching also required</td>
</tr>
<tr>
<td>Aleamoni</td>
<td>Students can make consistent judgments one of multiple sources</td>
<td>College University</td>
<td>Course level grade level affect ratings</td>
</tr>
<tr>
<td>Abrami</td>
<td>K-12 and higher education more similar than different. Students capable of rating teacher performance</td>
<td>College University</td>
<td>Meta-analysis reveals need for more rigorous investigations to control for validity</td>
</tr>
<tr>
<td>Popham</td>
<td>Judgment based teacher evaluations: students are one of five data sources</td>
<td>K-12 College University</td>
<td>Students are positively disposed toward some subjects</td>
</tr>
<tr>
<td>Herman</td>
<td>Student as recipient of education is best able to evaluate giver (teacher) daily contact</td>
<td>K-12</td>
<td>Student may rate strict teachers low, immature, peer pressure</td>
</tr>
<tr>
<td>Harris</td>
<td>Reliable, objective, and valid opportunity to frequently observe</td>
<td>K-12</td>
<td>Only limited acceptance due to traditional attitudes</td>
</tr>
<tr>
<td>Source researcher</td>
<td>Students as raters</td>
<td>Level grade/age</td>
<td>Concerns</td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------------</td>
<td>----------------</td>
<td>----------</td>
</tr>
<tr>
<td>Ghorpade$^a$</td>
<td>Potential bias raises questions re: validity and reliability of students' ratings</td>
<td>College</td>
<td>Bias in favor of females versus males. Bias in favor of white males versus minority males</td>
</tr>
<tr>
<td>Lackritz$^a$</td>
<td></td>
<td>University</td>
<td></td>
</tr>
<tr>
<td>L'Hommédieu</td>
<td>Ratings are effective for evaluating and improving instruction</td>
<td>College</td>
<td>Additional study should replicate using existing instruments</td>
</tr>
<tr>
<td>Menges</td>
<td></td>
<td>University</td>
<td></td>
</tr>
<tr>
<td>Brinko</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scriven</td>
<td>Useful basis for rating teachers</td>
<td>K-12</td>
<td>Must be gathered in a secure way and in a suitably designed form</td>
</tr>
<tr>
<td>Braskamp</td>
<td>Multiple sources are desirable and students are a legitimate source</td>
<td>K-12</td>
<td>Students are not in a position to judge relevance and recency of the course content and knowledge or scholarship of the instructor</td>
</tr>
<tr>
<td></td>
<td></td>
<td>College</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>University</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER III. METHODS

This study identified, refined, and tested three sets of student feedback items comprising separate questionnaires for use as one component in a total teacher performance evaluation system. The primary purpose of this investigation was to examine and describe the discriminating power of student feedback items and the effect of 1) grade and level, 2) subject area taught, 3) time of day, 4) gender of teacher, 5) core versus elective nature of the course, and 6) career ladder status, on the mean score student rating of teachers. All students, from one school system, in grades 6-12 rated their teachers' performance by completing one of three student feedback questionnaires. The questionnaires, subjects who participated, data collection procedures, and statistical analyses are reviewed in this chapter.

The Questionnaires

Seven student feedback questionnaires, used to assess students' ratings of teacher performance (Appendices A, B, C, D, E, F, and G), were developed in 1990 by teachers and administrators in the Cave Creek (Arizona) Unified School District No. 93. This current investigation, and an associated K-5 study (Weber, 1992), refined the work of Hidlebaugh (1973) and Judkins (1987) by re-examining the discriminating power of items previously validated with volunteers.
Developing the questionnaires

The Cave Creek Career Ladder Plan Stakeholders Committee affirmed the inclusion of student ratings of teacher performance as one component of the career ladder algorithm. Consequently, a need existed to develop valid, reliable, and legally discriminating student feedback questionnaires. As a result of involving the School Improvement Model (SIM) team, and specifically Professor Richard P. Manatt, in the restructuring of the Cave Creek Career Ladder Plan, the stakeholders committee obtained a high quantity and quality of services and products. Specifically, SIM provided the Cave Creek Public School system direct access to sample questionnaires and the entire pool of discriminating items previously developed and identified by Judkins (1987).

Each of the three Cave Creek student feedback questionnaires consisted of 20 items. Most of the items comprising the student feedback instruments were selected from the pool of valid, reliable, and legally discriminating items identified by Judkins. However, in response to local teacher concerns, the stakeholders committee also supported the local development of additional student feedback items. Specifically, Table 2 lists the two middle school regular, five middle school floater, and two high school questionnaire items which were developed by local teachers and administrators. These items had not been previously tested for validity, reliability, or discriminating power.

Table 3 contains those items which the Cave Creek Career Ladder Plan Stakeholders Committee identified as being negatively worded; that is, items which contained the words not, never, or implied a negative action
<table>
<thead>
<tr>
<th>Questionnaire level</th>
<th>Item number</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Middle school</td>
<td></td>
<td>My teacher likes it when we ask questions.</td>
</tr>
<tr>
<td>regular</td>
<td>8</td>
<td>My teacher likes it when we ask questions.</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>My teacher starts lessons explaining what we are going to do and why we are going to do it.</td>
</tr>
<tr>
<td>Middle school</td>
<td></td>
<td>My teacher likes it when we ask questions.</td>
</tr>
<tr>
<td>floater</td>
<td>8</td>
<td>My teacher likes it when we ask questions.</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>My teacher explains the rules for classroom behavior very clearly.</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>My teacher starts lessons explaining what we are going to do and why we are going to do it.</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>My teacher is easy to understand when talking.</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>My teacher’s tests are fair.</td>
</tr>
<tr>
<td>Senior high school</td>
<td>16</td>
<td>My teacher likes it when we ask questions.</td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>My teacher is available to me during class-time and other times during the school day.</td>
</tr>
</tbody>
</table>

on the part of students or teachers. These items were subsequently revised by SIM to read as positive descriptors of teacher and student behaviors. Student ratings of teacher data for both 1990 and 1991 were examined and tested, using the Menne and Tolsma (1971) methodology, to identify those items which continued to possess discrimination power. The 1991 data were added to the current study’s 1990 data because it became available at an opportune time. Further, an opportunity was provided to
Table 3. Questionnaire items revised by SIM for Cave Creek Public Schools

<table>
<thead>
<tr>
<th>Questionnaire level</th>
<th>Middle school Regular Items</th>
<th>Middle school Floater Items</th>
<th>Senior high Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>3,5,6,7,9, 11,15</td>
<td>3,5,6,7,11, 17,19,20</td>
<td>3,4,6,7,8,11, 12,13,14,17</td>
<td></td>
</tr>
</tbody>
</table>

examine what differences occurred in 1991 when all grades 6 through 8 students completed the middle school regular questionnaire.

Deciding to recommend

In 1991 the stakeholders committee accepted and supported a SIM recommendation that student ratings of teacher performance, for both regular and special-area teachers, be obtained by using the four regular student feedback questionnaires. The SIM recommendation to discontinue the use of the three floater questionnaires was based upon the small numbers of raters in several special-area classes and upon preliminary findings of the current investigation. The initial findings indicated that the floater items, especially at the primary and the elementary levels, did not necessarily possess sufficient discrimination power or reliability to warrant continued use of the questionnaires.

Although it would have been preferable to delay the decision until further analysis of findings was completed, practical considerations associated with the ongoing refinement of the career ladder plan necessitated the early recommendation. Consequently, three feedback
questionnaires in the current study, and four in an associated investigation (Weber, 1992) administered in 1990, were analyzed for item discrimination power.

The remaining two regular student feedback questionnaires, and two in an associated investigation (Weber, in progress) administered in 1991, were examined and tested using the Menne and Tolsma (1971) methodology to identify those items which continued to possess discrimination power. Cronbach’s coefficient alpha method was also used to calculate a reliability coefficient. No additional statistical tests were used to analyze the 1991 data. Therefore, as originally planned, only the 1990 data were used to study the effects which extraneous influences (to teacher performance), such as grade and level, subject area taught, time of day, gender of teacher, core versus elective, and career ladder status, have on the students’ mean score ratings of teachers.

The Sample

Cave Creek Unified School District No. 93 is located just north of Scottsdale, Arizona. All secondary students from Cave Creek (Arizona) Unified School District No. 93 completed the student feedback questionnaires in May of 1990. Further, all grades 6-12 Cave Creek Public School teachers participated in the study by administering the questionnaires to their classes and were consequently rated by their students. This non-voluntary sample, consisting of the total population of a single school system, is an important element of this study. All
previous studies regarding student ratings of teacher performance have relied upon volunteers as subjects.

Table 4 describes the two secondary schools, the grade levels, the number of teachers, and the number of students who participated in the study. Although the actual number of raters for each teacher varied from class to class, all teachers were rated by at least 15 students. Specifically, 45 teachers and 826 students served as the sources of data for this investigation.

Table 4. Participating students, teachers, and schools

<table>
<thead>
<tr>
<th>School name</th>
<th>Grade level</th>
<th>Number of teachers</th>
<th>Student enrollment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Desert</td>
<td>6</td>
<td>5</td>
<td>119</td>
</tr>
<tr>
<td>Arroyo</td>
<td>7</td>
<td>5</td>
<td>134</td>
</tr>
<tr>
<td>Middle</td>
<td>8</td>
<td>5</td>
<td>130</td>
</tr>
<tr>
<td>School</td>
<td>6,7,8</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>6,7,8</td>
<td>18</td>
<td>383</td>
</tr>
<tr>
<td>Cactus</td>
<td>9</td>
<td></td>
<td>129</td>
</tr>
<tr>
<td>Shadows</td>
<td>10</td>
<td></td>
<td>113</td>
</tr>
<tr>
<td>High</td>
<td>11</td>
<td></td>
<td>109</td>
</tr>
<tr>
<td>School</td>
<td>12</td>
<td>9,10,11,12</td>
<td>27</td>
</tr>
<tr>
<td>TOTAL</td>
<td>9,10,11,12</td>
<td>27</td>
<td>443</td>
</tr>
</tbody>
</table>
Administration of the Questionnaires

The Cave Creek Stakeholders Committee and the SIM team developed uniform and specific procedures for administering the questionnaires (Appendix H). Local teachers administered the questionnaire according to the agreed upon procedures. Exchanging classes was a recommended procedure for all grade levels. Secondary students recorded their rating of teacher performance by marking an electronic scanning (bubble sheet) response form (Appendix I). Questionnaires were completed and scanning response forms were collected and grouped by class and teacher. Forms were labeled and delivered to SIM for processing and analysis.

Quality control

The completed student response sheets were received for this study, and a parallel study by Weber, and were prepared for electronic scanning. Each of the forms was checked for proper coding and for the number of raters noted. Forms were organized by teacher and school level category, and all necessary information was recorded. All completed student ratings forms were scanned by the Iowa State University (ISU) Test and Evaluation Services department.

After scanning was successfully completed, the data for the current investigation were transferred to the ISU Computation Center’s mainframe computer. The ISU adapted version of Stanford University’s WYLBUR computer program was used to process the data and establish data files for use with the SPSS-X statistical software package. The ISU Test and Evaluation Services department facilitated the analysis of all
questionnaire items using the Menne and Tolsma (1971) statistical methodology and computer program. In addition, Cronbach's coefficient alpha method was used to calculate a reliability coefficient and to estimate the internal consistency of the discriminating items for each questionnaire.

In order to facilitate and enhance the quality of possible future research studies, revised procedures for administering and completing the student feedback questionnaires were developed on March 7, 1991. The revised procedures and developed forms ensured that desirable and necessary demographic data would be obtained and recorded during the 1991 administration of the student feedback questionnaires. Specifically, the revised instruction sheets (Appendix J) were intended to collect additional information that 1) was more accurate and complete than the initial 1990 teacher demographic data, 2) identified student raters by gender, 3) eliminated the inclusion of extraneous information, and 4) clarified the grade and subject area of the teacher being rated.

Further, the procedures for grouping, labeling, shipping, receiving, and scanning future student ratings of teacher performance data were improved and described.

Creating data management tools

In response to the identified needs of the Cave Creek School District and the needs of this study, several statistical programs using the SPSS-X statistical software package were created. The programs were designed for
the purpose of tabulating data (Appendix K), recoding scores to numerically "fit" the Cave Creek Career Ladder Plan algorithm (Appendix L), and facilitating further statistical manipulation, analysis, and display of the data. Initially, students' ratings of teachers were categorized and analyzed by teacher, questionnaire item, and period of the school day. Specifically, a mean score and measure of standard deviation for each category was calculated in order to appropriately describe how teachers were rated by their students. These respective measures of central tendency and variability provide essential information concerning the value of student ratings as one component of a total teacher performance evaluation system.

Additional computer-managed statistical programs were written, using the SPSS-X software package, to test the current investigation's stated research questions. In order to facilitate additional examination of the data and possible replication of the current investigation, all study data and computer managed statistical programs have been stored, using the ISU WYLBUR computer program, on magnetic cartridge tape files.

Treatment of Data

Hildebaugh's (1973), Look's (1983), Judkin's (1987), Ferrare's (1990), and Green's (1990) application of the Menne and Tolsma (1971) methodology for determining item discrimination in instruments using group responses was used in this study to identify which of the 60 items contained in the three questionnaires possessed item discrimination power. This methodology is based on the premise that a percentage of the total
sum of squares which is due to groups or between-groups can be appropriately used as an index of item discrimination.

According to Menne and Tolsma (1971), whether an item contained in a student feedback questionnaire designed to measure group responses of teacher performance is a discriminating one can be inferred from the pattern of between-group and within-group variances. Consequently, Hidlebaugh, Look, Judkins, Ferrare, and Green each believed that this procedure was advantageous because it was pragmatic to have used the percentage of the total sum of squares due to between-groups as an appropriate discrimination index.

Specifically, Hidlebaugh (1973) observes that the ratio of between-to within-group mean squares, under the analysis of variance assumptions, varies as the $F$ statistic and is also influenced by sample size. Further, Hidlebaugh concludes that the percentage of the total sum of squares due to between groups is independent of sample size and asserts, therefore, that the Menne and Tolsma methodology is an advantageous procedure for determining item discrimination power.

The Menne and Tolsma methodology

By definition, for an item to be identified as possessing discriminating power, a certain percentage of the total sum of squares must be due to the variance between teachers (Judkins, 1987). The minimum percentage was established for this investigation based on the assumption of a minimum of 15 raters for each teacher. A theoretical minimum was required because Menne and Tolsma (1971) concluded that, "If an item is a
discriminating one in a situation involving a few small groups, then it will also be capable of discriminating among more numerous and/or larger groups. The reverse, of course, is not true." It is possible for an item to be discriminating in a situation where there are several large groups but not discriminating in a situation involving a few small groups. Ferrare (1990) observes that using more than 15 students as a required minimum number of raters could reduce the effect of the results of this study in smaller school districts. Conversely, using a smaller minimum number would increase the difficulty of finding discriminating items and it would not be representative of public or nonpublic class sizes.

Table 5 illustrates the rationale for establishing 13 percent as a minimum percentage for identifying discriminating items at the .05 level of significance. Hidlebaugh (1973), Look (1983), Judkins (1987), Ferrare (1990), and Green (1990) utilized this method in previous research studies. Ferrare (1990) concludes that 13 percent is a between-group minimum percentage of the total sums of squares sufficient to discriminate at the .05 level of significance. Further, a between-group minimum percentage of the total sums of squares sufficient to discriminate at the .01 level of significance is 22 percent. Both minimum percentages assume that the item distinguishes between two teachers each rated by at least 15 students. Many more teachers and student raters were included in the data analyzed for this study. Based upon the assumptions of Menne and Tolsma (1971), it is reasonable to conclude that items selected using the 13 percent criteria will be discriminating items.
Table 5. Analysis of variance for two groups with 15 subjects per group

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>2-1-1</td>
<td>x (13%)</td>
<td>x (13%)</td>
<td>4.20*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>100-x/28</td>
<td>(87/28)</td>
</tr>
<tr>
<td>Within groups</td>
<td>2(15-1)=28</td>
<td>100-x (87%)</td>
<td>100-x (87/28)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>29</td>
<td>100%</td>
<td>100%</td>
<td>28</td>
</tr>
</tbody>
</table>

Therefore:

\[
\frac{x}{100-x} = 4.20
\]

\[
x = 4.20 \left( \frac{100 - x}{28} \right)
\]

\[
28x = (4.20)(100-x)
\]

\[
28x = (420 - 4.20x)
\]

\[
(28 + 4.20)x = 420
\]

\[
32.2x = 4.20
\]

\[
x = 13.04
\]

\[
100 - x = 86.96
\]

*The critical F value with 1 and 28 degrees of freedom at the .05 level is 4.20.

Look (1983) provides a caution regarding the application of the Menne and Tolsma (1971) methodology. For the purposes of the current study, discriminating items were identified based on the analysis of data for all teachers rated by at least 15 student raters. Look believes that the large number of raters ensures that the items identified as discriminating are representative of items which measure significant differences.
Nevertheless, these same items may not discriminate between teachers' performances when rated by fewer student raters. School districts wanting to use the existing discriminating items and student feedback questionnaires should continue to analyze the data based upon the number of raters completing the questionnaire for the designated teachers.

Cronbach's coefficient alpha method was used to calculate a reliability coefficient and to estimate the internal consistency of all discriminating items at the .05 level of significance for each questionnaire. Cronbach's coefficient alpha is a general form of the Kuder-Richardson (K-R 20) formula that assesses inter-item consistency or homogeneity of the items and is used when items are not scored dichotomously. Since student participants were asked to rate teacher performance on each item using a five-point, Lickert-type scale, coefficient alpha is the appropriate method for computing reliability (Borg and Gall, 1989).

**Examining the effect of specific variables**

A one-way classification analysis of variance (ANOVA) was used to show whether the group means for each level, middle school regular, middle school floater, and senior high school were not equal. The ANOVA one-way classification is the method for testing the null hypothesis,

\[ H_0: \mu_1 = \mu_2 = \mu_3 \ldots = \mu_k. \]

Hinkle, Wiersma, and Jurs (1988) state, "Using one-way ANOVA, the equality of all population means can be tested simultaneously while
maintaining the preestablished Type I error rate." The Scheffé multiple range test was calculated to determine which group means differed significantly.

A second one-way ANOVA was used to test whether a statistically significant difference existed among the group means for each school subject or course. A planned set of orthogonal contrasts was used to test the statistical significance of group means for the categories of math/science, social studies/English, fine arts, and others.

Hinkle, Wiersma, and Jurs (1988) state,

By definition, two contrasts are said to be orthogonal if knowledge of the outcome of one contrast in no way helps to predict the outcome of the second contrast. The hypotheses reflected by these contrasts are independent, and thus the experimentwise error rate is maintained at alpha.

Math/science was labeled group 1, social studies/English was group 2, fine arts was group 3, and others was group 4. The planned orthogonal contrasts were calculated as follows:

Contrast 1 = 1 1 -1 -1
Contrast 2 = 1 -1 0 0
Contrast 3 = 0 0 -1 1.

A third one-way ANOVA was used to show whether the period or time of day for the class being rated affected the group mean score ratings of teachers. Post hoc non-orthogonal contrasts were used to test for statistically significant differences between morning and afternoon class periods. In addition, four t-tests for groups were used to test whether group mean scores differed significantly by Grade (group 1 = grades 6 and 7 and group 2 = grades 11 and 12), Core (group 1)/Elective (group 2),
Teacher Gender (group 1 = males and group 2 = females), or Career Ladder Status (group 1 = on career ladder and group 2 = off career ladder).

Borg and Gall (1989) state,

The t distribution is used to determine the level of statistical significance of an observed difference between sample means. Generally, educational researchers will reject the null hypothesis if the t value reaches a significance level of .05.

Borg and Gall conclude that a significance test based on the calculation of a t or z value is appropriate when the researcher wishes to determine whether the difference between two sample means reflects population differences. In the current investigation t-tests are calculated to determine whether the mean score ratings of two samples, dichotomous groups of students or teachers, differ significantly from each other.

Finally, a factor analysis was conducted on the discriminating items to determine how many factors could be identified among the items comprising the Cave Creek student feedback questionnaires. Norusis (1988) describes factor analysis as a statistical technique used to identify a relatively small number of factors that can be used to represent relationships among sets of many interrelated variables.

Kim and Mueller (1978) suggest that factor analysis need not be confined to exploration. Since most of the original Cave Creek items were selected from the work of Judkins (1987), the researcher expected that three factors would be identified and that specific items would also be associated with one dimension or another. The approach taken in this study was to employ a confirmatory factor analysis.
A summary of the steps

All data were carefully processed, coded, and electronically scanned. Data sets were then transferred to the ISU Computer Services mainframe computer and analyzed using specific computer software and statistical programs. Frequency listings of all data were produced and examined to check for errors in processing and scanning. A detailed set of procedures for administering and processing the student feedback questionnaire response forms was developed. And finally, all data have been stored on magnetic cartridge tape files for purposes of validation and future study.
CHAPTER IV. FINDINGS

The basic problem for this study was the refinement of teacher performance rating items for use by students, based upon item discrimination power and grade level of instruction. To accomplish this task, 826 students completed one of three secondary level questionnaires in rating their teacher(s). A total of 18 middle school and 27 senior high school teachers were rated by students.

Overview

Desirable and legal discrimination

The subjects completed a five-point, Likert-type scale to rate the performance of their teacher(s) on each of 20 items. Items which discriminated at the .05 level of significance were identified using the Menne and Tolsma (1971) method for determining item discrimination power. Subsequently, Cronbach's coefficient alpha method was used to calculate a reliability coefficient and to estimate the internal consistency of all discriminating items, at the .05 level of significance, for each questionnaire.

Effect of extraneous variables

One-way analysis of variance and t-tests for groups were the inferential statistical tests used to determine whether the differences between sample means reflect population differences (Borg and Gall, 1989). Specifically, these inferential techniques were used to determine whether two or more sample means were significantly different from one another.
In using one-way ANOVA it is not possible to directly determine which of the three or more sample means differ significantly from one another. Therefore, whenever significant differences were identified, the Scheffé multiple-range post hoc was used.

**Expecting three performance categories**

Judkins' (1987) study developed items for three performance areas generally classified by SIM as 1) Productive Teaching Techniques, 2) Organized, Structured Classroom Management, and 3) Positive Interpersonal Relations. A factor analysis was conducted on the data for each of the two regular questionnaires to show how the discriminating items "load" or cluster into categories. Specifically, the factor analysis was used to describe whether the same number of factors were identified in the present study and whether the items contained in each factor were similarly grouped according to Judkins' original performance areas.

In this chapter, each of the research questions presented in Chapter I has been restated, and a null hypothesis will be stated for each statistical test conducted. Further, the results of the statistical tests are also displayed in table form. The questionnaires used to collect the data for this study can be found in Appendices E, F, and G.

**Student Feedback Questionnaire**

Three questionnaires, Middle School Regular, Middle School Floater, and Senior High School, were used to collect the data for this 1990 study.
In addition, data from two questionnaires (the original Middle School Regular and the Senior High School) administered in 1991, were used to further examine the item discrimination power of the students' ratings of teacher performance items.

**Item discrimination**

**Research Question 1** Research Question 1 asked whether any of the items on the 6-8 regular, 6-8 floater, and 9-12 questionnaires possess discriminating power. In order to have determined whether the discrimination power of the items differed, Hidlebaugh's (1973) and Judkins' (1987) adaptation of the Menne and Tolsma (1971) methodology for determining the discrimination index for items in instruments using group responses was applied to all 20 items on each questionnaire.

The statistical analysis indicated which questionnaire items had a sum of squares between-group variance equal to or exceeding the previously established 13 percent of the total sums of squares variance criterion for discrimination at the .05 level of significance. Further, items which had a sum of squares between-groups variance equal to or exceeding 22 percent of the total sums of squares variance were considered to discriminate at the .01 level of significance.

The null hypothesis for Research Question 1 stated that there would be no significant difference in the discriminating power of the items used in the student feedback questionnaires. The analysis indicated that a total of 56 of 60 items had a sum of squares between-groups variance equal
to or exceeding 13 percent of the total sums of squares variance and therefore discriminated at the .05 level of significance.

As shown in Table 6, the 1990 6-8 regular questionnaire contained eight of 20 items that had a between-groups variance equal to or exceeding 13 percent and 12 of 20 items that had a between-groups variance equal to or exceeding 22 percent. Further, Table 7 illustrates that the 1990 6-8 floater special-area questionnaire contained 12 of 20 items that had a between-groups variance equal to or exceeding 13 percent, four of 20 items that had a between-groups variance equal to or exceeding 22 percent, and four items that had a between-groups variance of less than 13 percent. In addition, Table 8 describes that the 1990 9-12 senior high questionnaire contained 20 of 20 items that had a between-groups variance equal to or exceeding 22 percent.

In comparison, Table 9 illustrates that the 1991 6-8 regular questionnaire contained 14 of 20 items that had a between-groups variance equal to or exceeding 13 percent and six of 20 items that had a between-groups variance equal to or exceeding 22 percent. Further, Table 10 shows that the 1991 9-12 senior high questionnaire contained 15 of 20 items that had a between-groups variance equal to or exceeding 13 percent and five of 20 items that had a between-groups variance equal to or exceeding 22 percent.

For the 1990 data the Cronbach alpha reliability coefficients calculated to determine the internal consistency of all items with a discrimination value equal to or exceeding 13 percent were 1) 94 percent
### Table 6. Item discrimination power of student feedback to teachers, 1990 6-8 regular questionnaire

<table>
<thead>
<tr>
<th>Item no.</th>
<th>Item</th>
<th>Item discrimination in percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>My teacher makes class work interesting.</td>
<td>31**</td>
</tr>
<tr>
<td>2</td>
<td>My teacher is fair with all.</td>
<td>29**</td>
</tr>
<tr>
<td>3</td>
<td>My teacher maintains discipline in our classroom.</td>
<td>19*</td>
</tr>
<tr>
<td>4</td>
<td>My teacher is well prepared for our class.</td>
<td>23**</td>
</tr>
<tr>
<td>5</td>
<td>My teacher gives assignments related to the subject we are studying.</td>
<td>28**</td>
</tr>
<tr>
<td>6</td>
<td>We discuss and summarize each lesson just studied.</td>
<td>26**</td>
</tr>
<tr>
<td>7</td>
<td>Our discussions focus on the topic of the lesson.</td>
<td>26**</td>
</tr>
<tr>
<td>8</td>
<td>My teacher likes it when we ask questions.</td>
<td>18*</td>
</tr>
<tr>
<td>9</td>
<td>I have more time to do my work than I need.</td>
<td>16*</td>
</tr>
<tr>
<td>10</td>
<td>My teacher starts lessons explaining what we are going to do and why we are going to do it.</td>
<td>19*</td>
</tr>
<tr>
<td>11</td>
<td>My teacher asks us questions in class to see if we understand what is being taught.</td>
<td>23**</td>
</tr>
<tr>
<td>12</td>
<td>My teacher explains new ideas in a way that is easy to understand.</td>
<td>27**</td>
</tr>
<tr>
<td>13</td>
<td>My teacher looks at our work, as we are doing it, to see if we understand the lesson.</td>
<td>19*</td>
</tr>
<tr>
<td>14</td>
<td>My teacher knows more about the subject than other teachers I have had.</td>
<td>25**</td>
</tr>
</tbody>
</table>

*Indicates items that discriminate at the .05 level of significance.

**Indicates items that discriminate at the .01 level of significance.
<table>
<thead>
<tr>
<th>Item no.</th>
<th>Item</th>
<th>Item discrimination in percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>My teacher has work for me to do if I finish an assignment before the class is over.</td>
<td>15*</td>
</tr>
<tr>
<td>16</td>
<td>My teacher often makes materials and worksheets for us to use.</td>
<td>21*</td>
</tr>
<tr>
<td>17</td>
<td>My teacher gives tests and quizzes.</td>
<td>36**</td>
</tr>
<tr>
<td>18</td>
<td>My teacher returns tests and assignments quickly.</td>
<td>24**</td>
</tr>
<tr>
<td>19</td>
<td>My teacher uses a variety of classroom activities and resources.</td>
<td>22**</td>
</tr>
<tr>
<td>20</td>
<td>My teacher gives enough time to do our work.</td>
<td>19*</td>
</tr>
</tbody>
</table>
Table 7. Item discrimination power of student feedback to teachers, 1990 6-8 special-area questionnaire

<table>
<thead>
<tr>
<th>Item no.</th>
<th>Item</th>
<th>Item discrimination in percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>My teacher makes class work interesting.</td>
<td>26**</td>
</tr>
<tr>
<td>2</td>
<td>My teacher is fair with all.</td>
<td>16*</td>
</tr>
<tr>
<td>3</td>
<td>My teacher maintains discipline in our classroom.</td>
<td>13*</td>
</tr>
<tr>
<td>4</td>
<td>My teacher is well prepared for class.</td>
<td>21*</td>
</tr>
<tr>
<td>5</td>
<td>My teacher gives assignments related to the subject we are studying.</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>We discuss and summarize each lesson just studied.</td>
<td>19*</td>
</tr>
<tr>
<td>7</td>
<td>Our discussions focus on the topic of the lesson.</td>
<td>22**</td>
</tr>
<tr>
<td>8</td>
<td>My teacher likes it when we ask questions.</td>
<td>18*</td>
</tr>
<tr>
<td>9</td>
<td>My teacher explains the rules for classroom behavior very clearly.</td>
<td>20*</td>
</tr>
<tr>
<td>10</td>
<td>My teacher starts lessons explaining what we are going to do and why we are going to do it.</td>
<td>21*</td>
</tr>
<tr>
<td>11</td>
<td>My teacher asks us questions in class to see if we understand what is being taught.</td>
<td>17*</td>
</tr>
<tr>
<td>12</td>
<td>My teacher explains new ideas in a way that is easy to understand.</td>
<td>18*</td>
</tr>
<tr>
<td>13</td>
<td>My teacher looks at our work, as we are doing it, to see if we understand the lesson.</td>
<td>18*</td>
</tr>
<tr>
<td>14</td>
<td>We do the same thing every day in class.</td>
<td>8</td>
</tr>
</tbody>
</table>

*Indicates items that discriminate at the .05 level of significance.

**Indicates items that discriminate at the .01 level of significance.
Table 7. Continued

<table>
<thead>
<tr>
<th>Item no.</th>
<th>Item</th>
<th>Item discrimination in percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>My teacher is easy to understand when talking.</td>
<td>19*</td>
</tr>
<tr>
<td>16</td>
<td>My teacher's tests are fair.</td>
<td>24**</td>
</tr>
<tr>
<td>17</td>
<td>My teacher gives tests and quizzes.</td>
<td>11</td>
</tr>
<tr>
<td>18</td>
<td>My teacher returns tests and assignments quickly.</td>
<td>11</td>
</tr>
<tr>
<td>19</td>
<td>My teacher uses a variety of classroom activities and resources.</td>
<td>23**</td>
</tr>
<tr>
<td>20</td>
<td>My teacher expects me to do the best work I can.</td>
<td>21*</td>
</tr>
</tbody>
</table>
Table 8. Item discrimination power of student feedback to teachers, 1990 9-12 senior high questionnaire

<table>
<thead>
<tr>
<th>Item no.</th>
<th>Item</th>
<th>Item discrimination in percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>My teacher makes class work interesting.</td>
<td>41**</td>
</tr>
<tr>
<td>2</td>
<td>My teacher asks questions to see if we understand what has been taught.</td>
<td>33**</td>
</tr>
<tr>
<td>3</td>
<td>My teacher gives assignments related to the subject we are studying.</td>
<td>37**</td>
</tr>
<tr>
<td>4</td>
<td>We discuss and summarize each lesson we have just studied.</td>
<td>31**</td>
</tr>
<tr>
<td>5</td>
<td>My teacher tells us how we can use what we have already learned to learn new things.</td>
<td>31**</td>
</tr>
<tr>
<td>6</td>
<td>My teacher maintains discipline in our classroom.</td>
<td>40**</td>
</tr>
<tr>
<td>7</td>
<td>My teacher returns tests and assignments quickly.</td>
<td>32**</td>
</tr>
<tr>
<td>8</td>
<td>My teacher gives me feedback about my performance.</td>
<td>26**</td>
</tr>
<tr>
<td>9</td>
<td>My teacher knows more about the subject than other teachers I have had.</td>
<td>33**</td>
</tr>
<tr>
<td>10</td>
<td>My homework helps me to learn the subject being taught.</td>
<td>31**</td>
</tr>
<tr>
<td>11</td>
<td>My teacher makes materials and worksheets for us to use.</td>
<td>36**</td>
</tr>
<tr>
<td>12</td>
<td>My teacher uses a variety of classroom activities and resources.</td>
<td>36**</td>
</tr>
<tr>
<td>13</td>
<td>The films and videotapes we watch help us learn about the subject we are studying.</td>
<td>36**</td>
</tr>
<tr>
<td>14</td>
<td>My teacher tells the class about library/media materials that will help us learn about the subject we are studying, when appropriate.</td>
<td>36**</td>
</tr>
</tbody>
</table>

**Indicates items that discriminate at the .01 level of significance.
Table 8. Continued

<table>
<thead>
<tr>
<th>Item no.</th>
<th>Item</th>
<th>Item discrimination in percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>My teacher is well organized.</td>
<td>37**</td>
</tr>
<tr>
<td>16</td>
<td>My teacher likes it when we ask questions.</td>
<td>34**</td>
</tr>
<tr>
<td>17</td>
<td>We work in different groups depending upon the activity in which we are involved.</td>
<td>35**</td>
</tr>
<tr>
<td>18</td>
<td>My teacher encourages us to look at problems in new ways and find new ways to solve problems.</td>
<td>33**</td>
</tr>
<tr>
<td>19</td>
<td>My teacher is available to help me during class time and other times during the school day.</td>
<td>37**</td>
</tr>
<tr>
<td>20</td>
<td>My teacher looks at our work, as we are doing it, to see if we understand the lesson.</td>
<td>32**</td>
</tr>
</tbody>
</table>
Table 9. Item discrimination power of student feedback to teachers, 1991 6-8 regular questionnaire

<table>
<thead>
<tr>
<th>Item no.</th>
<th>Item</th>
<th>Item discrimination in percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>My teacher makes class work interesting.</td>
<td>25**</td>
</tr>
<tr>
<td>2</td>
<td>My teacher is fair with all.</td>
<td>19*</td>
</tr>
<tr>
<td>3</td>
<td>My teacher maintains discipline in our classroom.</td>
<td>19*</td>
</tr>
<tr>
<td>4</td>
<td>My teacher is well prepared for our class.</td>
<td>18*</td>
</tr>
<tr>
<td>5</td>
<td>My teacher gives assignments related to the subject we are studying.</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>We discuss and summarize each lesson just studied.</td>
<td>19*</td>
</tr>
<tr>
<td>7</td>
<td>Our discussions focus on the topic of the lesson.</td>
<td>17*</td>
</tr>
<tr>
<td>8</td>
<td>My teacher likes it when we ask questions.</td>
<td>15*</td>
</tr>
<tr>
<td>9</td>
<td>I have more time to do my work than I need.</td>
<td>19*</td>
</tr>
<tr>
<td>10</td>
<td>My teacher starts lessons explaining what we are going to do and why we are going to do it.</td>
<td>19*</td>
</tr>
<tr>
<td>11</td>
<td>My teacher asks us questions in class to see if we understand what is being taught.</td>
<td>18*</td>
</tr>
<tr>
<td>12</td>
<td>My teacher explains new ideas in a way that is easy to understand.</td>
<td>20*</td>
</tr>
<tr>
<td>13</td>
<td>My teacher looks at our work, as we are doing it, to see if we understand the lesson.</td>
<td>24**</td>
</tr>
<tr>
<td>14</td>
<td>My teacher knows more about the subject than other teachers I have had.</td>
<td>18*</td>
</tr>
</tbody>
</table>

*Indicates items that discriminate at the .05 level of significance.

**Indicates items that discriminate at the .01 level of significance.
<table>
<thead>
<tr>
<th>Item no.</th>
<th>Item</th>
<th>Item discrimination in percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>My teacher has work for me to do if I finish an assignment before the class is over.</td>
<td>21*</td>
</tr>
<tr>
<td>16</td>
<td>My teacher often makes materials and worksheets for us to use.</td>
<td>24**</td>
</tr>
<tr>
<td>17</td>
<td>My teacher gives tests and quizzes.</td>
<td>34**</td>
</tr>
<tr>
<td>18</td>
<td>My teacher returns tests and assignments quickly.</td>
<td>23**</td>
</tr>
<tr>
<td>19</td>
<td>My teacher uses a variety of classroom activities and resources.</td>
<td>21*</td>
</tr>
<tr>
<td>20</td>
<td>My teacher gives enough time to do our work.</td>
<td>17*</td>
</tr>
</tbody>
</table>
Table 10. Item discrimination power of student feedback to teachers, 1991 9-12 senior high questionnaire

<table>
<thead>
<tr>
<th>Item no.</th>
<th>Item</th>
<th>Item discrimination in percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>My teacher makes class work interesting.</td>
<td>27**</td>
</tr>
<tr>
<td>2</td>
<td>My teacher asks questions to see if we understand what has been taught.</td>
<td>16*</td>
</tr>
<tr>
<td>3</td>
<td>My teacher gives assignments related to the subject we are studying.</td>
<td>14*</td>
</tr>
<tr>
<td>4</td>
<td>We discuss and summarize each lesson we have just studied.</td>
<td>13*</td>
</tr>
<tr>
<td>5</td>
<td>My teacher tells us how we can use what we have already learned to learn new things.</td>
<td>19*</td>
</tr>
<tr>
<td>6</td>
<td>My teacher maintains discipline in our classroom.</td>
<td>14*</td>
</tr>
<tr>
<td>7</td>
<td>My teacher returns tests and assignments quickly.</td>
<td>18*</td>
</tr>
<tr>
<td>8</td>
<td>My teacher gives me feedback about my performance.</td>
<td>13*</td>
</tr>
<tr>
<td>9</td>
<td>My teacher knows more about the subject than other teachers I have had.</td>
<td>21*</td>
</tr>
<tr>
<td>10</td>
<td>My homework helps me to learn the subject being taught.</td>
<td>15*</td>
</tr>
<tr>
<td>11</td>
<td>My teacher makes materials and worksheets for us to use.</td>
<td>13*</td>
</tr>
<tr>
<td>12</td>
<td>My teacher uses a variety of classroom activities and resources.</td>
<td>23**</td>
</tr>
<tr>
<td>13</td>
<td>The films and videotapes we watch help us learn about the subject we are studying.</td>
<td>32**</td>
</tr>
</tbody>
</table>

*Indicates items that discriminate at the .05 level of significance.

**Indicates items that discriminate at the .01 level of significance.
Table 10. Continued

<table>
<thead>
<tr>
<th>Item no.</th>
<th>Item</th>
<th>Item discrimination in percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>My teacher tells the class about library/media materials that will help us learn about the subject we are studying, when appropriate.</td>
<td>24**</td>
</tr>
<tr>
<td>15</td>
<td>My teacher is well organized.</td>
<td>15*</td>
</tr>
<tr>
<td>16</td>
<td>My teacher likes it when we ask questions.</td>
<td>17*</td>
</tr>
<tr>
<td>17</td>
<td>We work in different groups depending upon the activity in which we are involved.</td>
<td>22**</td>
</tr>
<tr>
<td>18</td>
<td>My teacher encourages us to look at problems in new ways and find new ways to solve problems.</td>
<td>18*</td>
</tr>
<tr>
<td>19</td>
<td>My teacher is available to help me during class time and other times during the school day.</td>
<td>17*</td>
</tr>
<tr>
<td>20</td>
<td>My teacher looks at our work, as we are doing it, to see if we understand the lesson.</td>
<td>19*</td>
</tr>
</tbody>
</table>
for the 6-8 regular discriminating items, 2) 95 percent for the 6-8 floater special-area discriminating items, and 3) 97 percent for the 9-12 senior high discriminating items. The discrimination value for each 1990 item is shown in Tables 6, 7 and 8.

Cronbach alpha reliability coefficients for discriminating items for the 1991 data were 1) 95 percent for the 6-8 regular discriminating items and 2) 95 percent for the 9-12 senior high discriminating items. The discrimination value for each 1991 item is shown in Tables 9 and 10. In conclusion, this study tried to show that there would be significant differences in the discriminating power of the items used in the student feedback questionnaires and was able to do so.

Research Question 2

Research Question 2 asked whether it was possible for teachers at the local level to develop discriminating items. All items developed by secondary grades 9-12 teachers and administrators of the Cave Creek Public Schools are listed in Table 11 and Table 12.

The null hypothesis for Research Question 2 stated that there would be no significant difference in the discriminating power of the items developed at the local level.

Specifically, Table 11 describes the six items from the 1990 data that discriminated at the .05 level of significance and three items that discriminated at the .01 level of significance. Further, discriminating items from the 1991 data are identified in Table 12. Specifically, Table 12 describes the four items from the 1991 data that discriminated at the .05 level of significance.
Table 11. Discrimination power of items developed by teachers and administrators at the local level, 1990 6-8 regular, 6-8 special-area, and 9-12 senior high questionnaire

<table>
<thead>
<tr>
<th>Item no.</th>
<th>Item</th>
<th>Item discrimination in percent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>1990 6-8 regular</strong></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>My teacher likes it when we ask questions.</td>
<td>18*</td>
</tr>
<tr>
<td>10</td>
<td>My teacher starts lessons explaining what we are going to do and why we are going to do it.</td>
<td>19*</td>
</tr>
<tr>
<td></td>
<td><strong>1990 6-8 floater special-area</strong></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>My teacher likes it when we ask questions.</td>
<td>18*</td>
</tr>
<tr>
<td>9</td>
<td>My teacher explains the rules for classroom behavior very clearly.</td>
<td>20*</td>
</tr>
<tr>
<td>10</td>
<td>My teacher starts lessons explaining what we are going to do and why we are going to do it.</td>
<td>21*</td>
</tr>
<tr>
<td>15</td>
<td>My teacher is easy to understand when talking.</td>
<td>19*</td>
</tr>
<tr>
<td>16</td>
<td>My teacher's tests are fair.</td>
<td>24**</td>
</tr>
<tr>
<td></td>
<td><strong>1990 9-12 senior high</strong></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>My teacher likes it when we ask questions.</td>
<td>34**</td>
</tr>
<tr>
<td>19</td>
<td>My teacher is available to help me during class time and other times during the school day.</td>
<td>37**</td>
</tr>
</tbody>
</table>

*Indicates items that discriminate at the .05 level of significance.

**Indicates items that discriminate at the .01 level of significance.
Table 12. Discrimination power of items developed by teachers and administrators at the local level, 1991 6-8 regular and 9-12 senior high questionnaire

<table>
<thead>
<tr>
<th>Item no.</th>
<th>Item</th>
<th>1991 6-8 regular</th>
<th>1991 9-12 senior high</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>My teacher likes it when we ask questions.</td>
<td>15*</td>
<td>17*</td>
</tr>
<tr>
<td>10</td>
<td>My teacher starts lessons explaining what we are going to do and why we are going to do it.</td>
<td>19*</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>My teacher likes it when we ask questions.</td>
<td>17*</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>My teacher is available to help me during class time and other times during the school day.</td>
<td>17*</td>
<td></td>
</tr>
</tbody>
</table>

*Indicates items that discriminate at the .05 level of significance.

Nine items were developed in 1990 and 1991 by local teachers and administrators for the student feedback questionnaires. Thus, 100 percent of the locally developed items had a sum of squares between-groups variance equal to or exceeding 13 percent of the total sums of squares variance and consequently discriminated or measured differences between teachers. In conclusion, this study tried to show that there would be a difference in the discriminating power of items developed at the local level and was able to do so.
Effect of extraneous variables

Research Question 3 asked whether 1) grade level, 2) subject area taught, 3) middle school level or high school level, 4) period of day, 5) core versus elective nature of the course, 6) career ladder status, and 7) teacher gender, have an effect upon the student mean score ratings for teachers.

Research Null Hypothesis 3a stated that there would be no significant difference in the student mean score ratings for teachers between group 1, grades 6 and 7, and group 2, grades 11 and 12. In order to determine if there were significant differences between rater groups, a t-test for independent samples was utilized to examine the relationship of the mean score ratings for the two groups.

For this study, 61 mean score ratings for combined grades 6 and 7 were compared with 33 mean score ratings for combined grades 11 and 12. The null hypothesis that there would be no significant difference in the mean score student ratings for teachers between group 1, grades 6 and 7, and group 2, grades 11 and 12, was rejected ($t=-2.63$, 2-tail prob.$=.010$). The results of the analysis are shown in Table 13. The mean score rating for grades 6 and 7 is 54.18, whereas the mean score rating for grades 11 and 12 was nearly 5 points higher at 59.36.

Research Null Hypothesis 3b stated that there would be no significant difference in the student mean score ratings based on subject area taught. In order to determine if there were significant differences between rater groups, a one-way analysis of variance (ANOVA) was utilized to explore the
Table 13. Analysis of student mean score ratings of teachers by grade level

<table>
<thead>
<tr>
<th>Grade group</th>
<th>Number</th>
<th>Mean</th>
<th>S.D.</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 and 7</td>
<td>61</td>
<td>54.18</td>
<td>8.38</td>
<td>-2.63**</td>
</tr>
<tr>
<td>11 and 12</td>
<td>33</td>
<td>59.36</td>
<td>10.38</td>
<td></td>
</tr>
</tbody>
</table>

**Significant at the .01 level.

relationship among the subjects when grouped by category. Table 14 describes the means and standard deviations for the four groups. A planned set of orthogonal contrasts was used to test the statistical significance of group means for the categories of math/science (group 1), social studies/English (group 2), fine arts (group 3), and other subjects (group 4).

The one-way ANOVA found a significant difference in the mean scores for the four groups. Table 15 reveals an F value of 4.06 with an F probability of .008, which indicates a level of discrimination at the .01 level of significance. Therefore, Research Null Hypothesis 3b: There is no significant difference in the student mean score ratings based on subject area taught, was rejected. Because there was a significant difference among the levels of the independent variable subject, a multiple range test would normally be used to identify which levels produced the difference.

However, for this study three a priori orthogonal comparisons were specified. The contrast t value for group 1 and group 2 compared with
Table 14. Means and standard deviations of student mean score ratings by subject

<table>
<thead>
<tr>
<th>Subject</th>
<th>N</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Math/science</td>
<td>61</td>
<td>54.98</td>
<td>10.74</td>
</tr>
<tr>
<td>English/social studies</td>
<td>55</td>
<td>54.20</td>
<td>9.73</td>
</tr>
<tr>
<td>Fine arts</td>
<td>41</td>
<td>47.24</td>
<td>13.36</td>
</tr>
<tr>
<td>Other subjects</td>
<td>64</td>
<td>52.14</td>
<td>13.02</td>
</tr>
</tbody>
</table>

Table 15. One-way analysis of variance of student mean score ratings by subject

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of squares</th>
<th>Mean squares</th>
<th>F ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>3</td>
<td>1675.98</td>
<td>558.66</td>
<td>4.06**</td>
</tr>
<tr>
<td>Within groups</td>
<td>217</td>
<td>29855.08</td>
<td>137.58</td>
<td></td>
</tr>
</tbody>
</table>

Contrast groups

| Math/science and English/social studies with Fine arts and Other subjects | 3.06** |
| Math/science with English/social studies | 0.36 |
| Fine arts with Other subjects | 2.09* |

*Significant at the .05 level.
**Significant at the .01 level.
group 3 and group 4 was 3.06 with a t probability of .003 indicating a significant difference at the .01 level. Further, the contrast t value for group 3 compared with group 4 was 2.09 with a t probability of .038 indicating a significant difference at the .05 level. When group 1 was compared with group 2, no significant difference was identified.

It can be concluded that student mean score ratings of teacher performance in math, science, English, and social studies are significantly different than the student mean score ratings in fine arts classes and in other subjects. Further, no differences are found between the student mean score ratings in the core subjects of math and science, and English and social studies. However, significant differences did exist between the student mean score ratings in fine arts classes and the other remaining subject areas. Specifically, the mean score ratings in fine arts classes were lower than the mean score ratings in the other remaining subjects.

Research Null Hypothesis 3c stated there would be no significant difference in the student mean score ratings for teachers among the three levels of the student feedback questionnaires. In order to determine if there were significant differences between rater groups, a one-way analysis of variance (ANOVA) was utilized to examine the relationship among the mean scores for each of the three questionnaires. Table 16 describes the means and standard deviations for the three groups. A planned set of orthogonal contrasts was used to test the statistical significance of group means for the middle school floater questionnaires.
(group 1), the middle school regular questionnaires (group 2) and the senior high questionnaires (group 3).

The one-way analysis of variance (ANOVA) found no significant differences in the student mean score ratings of the three groups or levels of questionnaires. Table 17 shows an F ratio of .262 with an F probability of .7699. Therefore, Research Null Hypothesis 3c: There is no significant difference in the student mean score ratings for teachers among three levels of student feedback questionnaires, was not rejected. It can be concluded that the grade level of each questionnaire did not significantly affect the students' mean score ratings of teacher performance.

Because the Bartlett-Box F equaled 5.180 with a probability of .006, the separate variance estimate for the t value was used for both orthogonal contrasts. The first contrast (group 1 and group 2 compared with group 3) t value of -0.79 had a t probability of 0.430 and the second contrast (group 1 compared with group 2) t value of 0.20 had a t probability of 0.843. Therefore, neither orthogonal contrast revealed significant differences among the student mean score ratings for teachers when compared by level. In other words, the student mean score rating for all middle school teachers was similar to the student mean score rating for all high school teachers.

Research Null Hypothesis 3d stated there would be no significant difference in the student mean score ratings for teachers when compared according to the period of the day in which the ratings were completed.
Table 16. Means and standard deviations of student mean score ratings by questionnaire level

<table>
<thead>
<tr>
<th>Questionnaire level</th>
<th>N</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Middle school floater</td>
<td>21</td>
<td>53.67</td>
<td>8.333</td>
</tr>
<tr>
<td>Middle school regular</td>
<td>78</td>
<td>53.24</td>
<td>9.722</td>
</tr>
<tr>
<td>Senior high school</td>
<td>147</td>
<td>52.27</td>
<td>12.723</td>
</tr>
</tbody>
</table>

Table 17. One-way analysis of variance of student mean score ratings by questionnaire level

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of squares</th>
<th>Mean squares</th>
<th>F ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>2</td>
<td>69.59</td>
<td>34.796</td>
<td>.262</td>
</tr>
<tr>
<td>Within groups</td>
<td>217</td>
<td>32300.15</td>
<td>132.92</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Contrast groups</th>
<th>Contrast t values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Middle school floater and</td>
<td></td>
</tr>
<tr>
<td>Middle school regular with</td>
<td></td>
</tr>
<tr>
<td>Senior high school</td>
<td>-0.79</td>
</tr>
<tr>
<td>Middle school floater with</td>
<td></td>
</tr>
<tr>
<td>Middle school regular</td>
<td>0.20</td>
</tr>
</tbody>
</table>

Bartlett-Box F = 5.180, p = .006
In order to determine if there were significant differences between rater groups, a one-way analysis of variance (ANOVA) was utilized to examine the relationship among the periods. Table 18 describes the means and standard deviations for the three groups.

The one-way analysis of variance (ANOVA) found no significant differences in the mean score ratings of the various periods of the day. Table 19 shows an F ratio of .660 with an F probability of .7057. Therefore, Research Null Hypothesis 3d: There is no significant difference in the student mean score ratings for teachers based on period of day, was retained. The use of post hoc non-orthogonal contrasts to compare morning class periods with afternoon class periods may appear to be somewhat superfluous since the research null hypothesis was retained. Nevertheless, because some researchers have suggested that contrasts should be calculated regardless of the outcome of the original statistical test, such tests were used. These calculated contrasts showed no

Table 18. Means and standard deviations of student mean score ratings by period of the day

<table>
<thead>
<tr>
<th>Period of the day</th>
<th>N</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>37</td>
<td>51.27</td>
<td>10.38</td>
</tr>
<tr>
<td>2</td>
<td>34</td>
<td>52.47</td>
<td>11.94</td>
</tr>
<tr>
<td>3</td>
<td>35</td>
<td>54.37</td>
<td>11.65</td>
</tr>
<tr>
<td>4</td>
<td>34</td>
<td>51.24</td>
<td>11.08</td>
</tr>
<tr>
<td>5</td>
<td>33</td>
<td>55.58</td>
<td>10.89</td>
</tr>
<tr>
<td>6</td>
<td>32</td>
<td>52.91</td>
<td>11.98</td>
</tr>
<tr>
<td>7</td>
<td>28</td>
<td>52.11</td>
<td>11.81</td>
</tr>
<tr>
<td>8</td>
<td>13</td>
<td>50.15</td>
<td>14.54</td>
</tr>
</tbody>
</table>
Table 19. One-way analysis of variance of student mean score ratings by period of day

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of squares</th>
<th>Mean squares</th>
<th>F ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>7</td>
<td>616.53</td>
<td>88.08</td>
<td>.660</td>
</tr>
<tr>
<td>Within groups</td>
<td>238</td>
<td>32369.74</td>
<td>133.42</td>
<td></td>
</tr>
</tbody>
</table>

significant differences among the mean scores for morning and afternoon student ratings of teachers. Therefore, it can be concluded that period of the day does not have a significant effect upon student mean score ratings for teacher performance.

Research Null Hypothesis 3e stated that there would be no significant difference in the student mean score ratings for teachers between group 1, core, and group 2, elective. In order to determine if there were significant differences between rater groups, a t-test for independent samples was utilized to examine the relationship of the mean score ratings for the two groups.

For this study 141 mean score ratings for core subjects of math, science, English, and social studies were compared with 105 mean score ratings for all other elective subjects such as physical education, art, music, home economics, etc. Specifically, the mean score rating for core subjects was 54.54, whereas the mean score rating for elective subjects was nearly 4 points lower at 50.23. Further, there was a significant difference in the variance for the two groups (F=1.93, probability=0.001). A larger diversity was found among the elective subjects than was
identified for the core subjects. Due to this difference the separate variance estimate was used. The null hypothesis that there would be no significant difference in the mean score student ratings for teachers between group 1, core, and group 2, elective, was rejected ($t=2.82$, 2-tail probability= 0.005). The results of the analysis are shown in Table 20. It can be concluded that a significant difference does exist, with student mean score ratings for core subjects being higher than those for elective subjects.

Table 20. Analysis of student mean score ratings of teachers by core and elective nature of courses

<table>
<thead>
<tr>
<th>Group</th>
<th>Number</th>
<th>Mean</th>
<th>S.D.</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core</td>
<td>141</td>
<td>54.54</td>
<td>9.57</td>
<td>2.82**</td>
</tr>
<tr>
<td>Elective</td>
<td>105</td>
<td>50.23</td>
<td>13.31</td>
<td></td>
</tr>
</tbody>
</table>

**Significant at the .01 level.

Research Null Hypothesis 3f stated that there would be no significant difference in the student mean score ratings for teachers between group 1, career ladder participant, and group 2, non-career ladder teacher. In order to examine whether there was a significant difference between rater groups, a $t$-test for independent samples was utilized to examine the relationship of the student mean score ratings for the two groups.

For this study, 131 mean score ratings for career ladder participants were compared with 115 mean score ratings for non-career ladder teachers.
Specifically, the mean score rating for career ladder participants was 54.98 and the mean score rating for non-participating teachers was lower at 50.10. An F value of 1.13 with a 2-tailed probability of 0.498 confirms the assumption that the population from which the study samples were taken were homogeneous with respect to the student mean score ratings. Therefore, using the pooled variance estimate the null hypothesis that there would be no significant difference in the mean score student ratings for teachers between group 1, career ladder participants, and group 2, non-career ladder teachers, was rejected (t=3.39, 2-tailed probability = 0.001). The results of the analysis are shown in Table 21. It can be concluded that a significant difference does exist (p<.01) with student mean score ratings for career ladder participants being higher than those for non-career ladder teachers.

Research Null Hypothesis 3g stated that there would be no significant difference in the student mean score ratings for teachers between group 1, male teachers, and group 2, female teachers. In order to determine if there were significant differences between rater groups, a t-test for

Table 21. Analysis of student mean score ratings of teachers by career ladder participation status

<table>
<thead>
<tr>
<th>Group</th>
<th>Number</th>
<th>Mean</th>
<th>S.D.</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Career ladder participant</td>
<td>131</td>
<td>54.98</td>
<td>10.93</td>
<td>3.39**</td>
</tr>
<tr>
<td>Non-career ladder teacher</td>
<td>115</td>
<td>50.10</td>
<td>11.62</td>
<td></td>
</tr>
</tbody>
</table>

**Significant at the .01 level.
independent samples was utilized to examine the relationship of the mean score ratings for the two gender groups.

For this study 105 mean score ratings for male teachers were compared with 141 mean score ratings for female teachers. Specifically, the mean score rating for male teachers was 48.87, whereas the mean score rating for female teachers was nearly seven points higher at 55.55. Further, a significant difference in the variance for the two groups was found ($F=1.88$, probability=$0.001$) indicating a greater diversity among the male teachers. Due to this difference the separate variance estimate was used. The null hypothesis that there would be no significant difference in the mean score student ratings for teachers between group 1, male teachers, and group 2, female teachers, was rejected ($t=-4.50$, 2-tail probability=$0.001$). The results of the analysis are shown in Table 22. It can be concluded that a significant difference does exist at the .01 level of significance, with student mean score ratings for female teachers being higher than those for male teachers.

Table 22. Analysis of student mean score ratings of teachers by gender

<table>
<thead>
<tr>
<th>Teacher gender</th>
<th>Number</th>
<th>Mean</th>
<th>S.D.</th>
<th>$t$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>105</td>
<td>48.87</td>
<td>12.90</td>
<td>-4.50**</td>
</tr>
<tr>
<td>Females</td>
<td>141</td>
<td>55.55</td>
<td>9.40</td>
<td></td>
</tr>
</tbody>
</table>

**Significant at the .01 level.
Factor analysis

Research Question 4  Research Question 4 asked how many factors would be identified when a factor analysis was conducted on the discriminating items. Research Null Hypothesis 4 stated there would be no factors identified which could be used to represent relationships among sets of the discriminating items.

Judkins' (1987) study developed items for three performance areas generally classified by SIM as 1) Productive Teaching Techniques, 2) Organized, Structured Classroom Management, and 3) Positive Interpersonal Relations. To test Research Null Hypothesis 4, a factor analysis was conducted on the data for each of the two regular questionnaires to show how the discriminating items "load" or cluster into categories.

Although there was an expectation that three factors would be identified by utilizing a factor analysis, the additional discriminating items developed by Cave Creek teachers and administrators were not initially classified into the SIM performance areas. Therefore, an exploratory form of factor analysis was initially used in this study to identify the number of factors present in these data. Subsequently, the discriminating items developed by Judkins were analyzed to confirm whether they continued to load on the original factors or performance area clusters.

Kim and Mueller (1978) acknowledge that the division between exploratory and confirmatory factor analysis methods is not always clear-cut. They suggest that it is possible for a researcher to be aware of two
or three factors but also to be unable to anticipate exactly what variables, in this case which discriminating items and performance areas, will represent each.

Borg and Gall (1989) support the use of factor analysis by the researcher as a helpful method which provides an empirical base for reducing the number of variables to a few factors by combining variables that are moderately or highly correlated with each other. Further, they define the set of variables that is combined to form a factor as a mathematical expression of the common element that cuts across the combined variables.

In this study the variables that were specified for analysis were the discriminating items for the 1990 middle school regular and the senior high questionnaires. The goal of the factor analysis was to identify the not-directly observable factors by statistically examining a set of observable variables. The data were analyzed in raw score form and the criterion for extracting the number of factors was to establish a minimum eigenvalue of 1. Specifically, this criterion identifies and includes only those factors that account for variances greater than 1.

The goal of the next step, factor extraction, was to determine the factors. Therefore, the correlation matrix for all variables was computed and the Principal Axis Factoring (PAF) method of factor extraction was used to obtain estimates of the common factors. A varimax rotation, a method of orthogonal rotation which simplifies the factor structure by maximizing the variance of a column of pattern matrix, was used to transform the initial matrix into one that was easier to interpret.
Norusis (1988) concludes that rotation does not affect the goodness of fit of a factor solution. Even though the factor matrix was changed, the commonalities, the proportion of variance explained by the common factors, and the percentage of variance explained did not change. Commonalities with a value of 0 indicate that the factors explain none of the variance while commonalities with a value of 1 indicate that all of the variance is explained by the common factors. Norusis (1988) states, "The variance that is not explained by the common factors is attributed to the unique factor and is called the uniqueness of the variable." At the same time, it was important to note that the percentage of variance accounted for by each of the factors did change.

The rotation redistributed the explained variance for the individual factors. Consequently, the rotational method employed described the items associated with the factors in largely similar but somewhat different relationships. Norusis (1988) affirms that the varimax method, used in this study, is the most commonly used method in attempting to minimize the number of variables that have high loadings on a factor. This procedure results in the enhanced interpretability of the factors.

To identify the factors, it was necessary to group the variables that had high loadings upon the same factors. The rotation was also plotted to provide a secondary opportunity to examine the success of the rotation. Further, a sorting of the factor matrix pattern was performed so that the variables with high loadings on the same factor appear together. No loadings less than 0.5 in absolute value were displayed.
Research Null Hypothesis 4: No factors will be identified which could be used to represent relationships among sets of the discriminating items, was rejected. Factor analysis of the two data sets identified two clusters for the senior high data and three clusters for the middle school data.

Table 23 presents a maximum likelihood analysis of the correlations among the 20 senior high questionnaire items. Specifically, two factors were identified for these data. Specifically, items 3, 2, 6, 9, 19, 4, 10, 15, 20, 16, 11, 1, 8, 5, and 7 were found to load on factor 1 (Productive Teaching Techniques) and the remaining items 12, 14, 13, 17, and 18 were found to load on factor 2 (Organized, Structured Classroom Management).

Table 24 presents a maximum likelihood analysis of the correlations among the 20 middle school regular questionnaire items. Specifically, three factors were identified for these data. Specifically, items 1, 12, 2, 19, 10, 14, 8, 13, and 15 were found to load on factor 1 (Productive Teaching Techniques), items 5, 7, 17, 6, 4, 11, 3, 16, and 18 were found to load on factor 2 (Organized, Structured Classroom Management), and the remaining items 20 and 9 were found to load on factor 3 (Positive Interpersonal Relations).

Tables 25 and 26 display the SIM performance areas and the specific criterion associated with each item. Further, Table 27 identifies the specific criteria and performance areas used by SIM to evaluate teacher performance. By considering the information contained in these three tables, it is possible to make several observations. First, although
Table 23. Maximum likelihood factor analysis with varimax rotation for the 20 senior high school items

<table>
<thead>
<tr>
<th>Number</th>
<th>Item</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Factor 1</strong></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>My teacher gives assignments related to the subject we are studying.</td>
<td>.740</td>
</tr>
<tr>
<td>2</td>
<td>My teacher asks questions to see if we understand what has been taught.</td>
<td>.731</td>
</tr>
<tr>
<td>6</td>
<td>My teacher maintains discipline in our classroom.</td>
<td>.679</td>
</tr>
<tr>
<td>9</td>
<td>My teacher knows more about the subject than other teachers I have had.</td>
<td>.675</td>
</tr>
<tr>
<td>19</td>
<td>My teacher is available to help me during class time and other times during the school day.</td>
<td>.673</td>
</tr>
<tr>
<td>4</td>
<td>We discuss and summarize each lesson we have just studied.</td>
<td>.654</td>
</tr>
<tr>
<td>10</td>
<td>My homework helps me to learn the subject being taught.</td>
<td>.647</td>
</tr>
<tr>
<td>15</td>
<td>My teacher is well organized.</td>
<td>.647</td>
</tr>
<tr>
<td>20</td>
<td>My teacher looks at our work, as we are doing it, to see if we understand the lesson.</td>
<td>.640</td>
</tr>
<tr>
<td>16</td>
<td>My teacher likes it when we ask questions.</td>
<td>.638</td>
</tr>
<tr>
<td>11</td>
<td>My teacher makes materials and worksheets for us to use.</td>
<td>.632</td>
</tr>
<tr>
<td>1</td>
<td>My teacher makes class work interesting.</td>
<td>.572</td>
</tr>
<tr>
<td>8</td>
<td>My teacher gives feedback about my performance.</td>
<td>.564</td>
</tr>
<tr>
<td>5</td>
<td>My teacher tells us how we can use what we have already learned to learn new things.</td>
<td>.559</td>
</tr>
<tr>
<td>7</td>
<td>My teacher returns tests and assignments quickly.</td>
<td>.525</td>
</tr>
<tr>
<td></td>
<td><strong>Factor 2</strong></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>My teacher uses a variety of classroom activities and resources.</td>
<td>.647</td>
</tr>
<tr>
<td>14</td>
<td>My teacher tells the class about library/media materials that will help us learn about the subject we are studying, when appropriate.</td>
<td>.633</td>
</tr>
<tr>
<td>13</td>
<td>The films or videotapes we watch help us learn about the subject we are studying.</td>
<td>.617</td>
</tr>
<tr>
<td>17</td>
<td>We work in different groups depending upon the activity in which we are involved.</td>
<td>.599</td>
</tr>
<tr>
<td>18</td>
<td>My teacher encourages us to look at problems in new ways and find new ways to solve problems.</td>
<td>.543</td>
</tr>
</tbody>
</table>
Table 24. Maximum likelihood factor analysis with varimax rotation for the 20 middle school regular items

<table>
<thead>
<tr>
<th>Number</th>
<th>Item</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>My teacher makes classwork interesting.</td>
<td>.703</td>
</tr>
<tr>
<td>12</td>
<td>My teacher explains new ideas in a way that is easy to understand.</td>
<td>.642</td>
</tr>
<tr>
<td>2</td>
<td>My teacher is fair with all.</td>
<td>.591</td>
</tr>
<tr>
<td>19</td>
<td>My teacher uses a variety of classroom activities and resources.</td>
<td>.584</td>
</tr>
<tr>
<td>10</td>
<td>My teacher starts lessons explaining what we are going to do and why we are going to do it.</td>
<td>.523</td>
</tr>
<tr>
<td>14</td>
<td>My teacher knows more about the subject than other teachers I have had.</td>
<td>.501</td>
</tr>
<tr>
<td>8</td>
<td>My teacher likes it when we ask questions.</td>
<td>.483</td>
</tr>
<tr>
<td>13</td>
<td>My teacher looks at our work, as we are doing it, to see if we understand the lesson.</td>
<td>.384</td>
</tr>
<tr>
<td>15</td>
<td>My teacher has work for me to do if I finish an assignment before the class is over.</td>
<td>.243</td>
</tr>
<tr>
<td>5</td>
<td>My teacher gives assignments related to the subject we are studying.</td>
<td>.680</td>
</tr>
<tr>
<td>7</td>
<td>Our discussions focus on the topic of the lesson.</td>
<td>.622</td>
</tr>
<tr>
<td>17</td>
<td>My teacher gives tests and quizzes.</td>
<td>.584</td>
</tr>
<tr>
<td>6</td>
<td>We discuss and summarize each lesson just studied.</td>
<td>.573</td>
</tr>
<tr>
<td>4</td>
<td>My teacher is well prepared for our class.</td>
<td>.562</td>
</tr>
<tr>
<td>11</td>
<td>My teacher asks us questions to see if we understand what has been taught.</td>
<td>.539</td>
</tr>
<tr>
<td>3</td>
<td>My teacher maintains discipline in our classroom.</td>
<td>.486</td>
</tr>
<tr>
<td>16</td>
<td>My teacher often makes materials and worksheets for us to use.</td>
<td>.483</td>
</tr>
<tr>
<td>18</td>
<td>My teacher returns tests and assignments quickly.</td>
<td>.403</td>
</tr>
<tr>
<td>20</td>
<td>My teacher gives enough time to do our work.</td>
<td>.734</td>
</tr>
<tr>
<td>9</td>
<td>I have more time to do my work than I need.</td>
<td>.591</td>
</tr>
</tbody>
</table>

**Factor 1**

**Factor 2**

**Factor 3**
Table 25. Comparison of items by factor analysis clusters and Judkins (1987) SIM performance areas, 1990 senior high questionnaire

<table>
<thead>
<tr>
<th>Number</th>
<th>Item</th>
<th>SIM criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Factor 1</strong></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>My teacher gives assignments related to the subject we are studying.</td>
<td>PT-C2</td>
</tr>
<tr>
<td>2</td>
<td>My teacher asks questions to see if we understand what has been taught.</td>
<td>PT-C2</td>
</tr>
<tr>
<td>6</td>
<td>My teacher maintains discipline in our classroom.</td>
<td>OSC-C14</td>
</tr>
<tr>
<td>9</td>
<td>My teacher knows more about the subject than other teachers I have had.</td>
<td>PT-C7</td>
</tr>
<tr>
<td>19</td>
<td>My teacher is available to help me during class time and other times during the school day.</td>
<td>PT-C9</td>
</tr>
<tr>
<td>4</td>
<td>We discuss and summarize each lesson we have just studied.</td>
<td>PT-C2</td>
</tr>
<tr>
<td>10</td>
<td>My homework helps me to learn the subject being taught.</td>
<td>PT-C7</td>
</tr>
<tr>
<td>15</td>
<td>My teacher is well organized.</td>
<td>OSC-C12</td>
</tr>
<tr>
<td>20</td>
<td>My teacher looks at our work, as we are doing it, to see if we understand the lesson.</td>
<td>PT-C2</td>
</tr>
<tr>
<td>16</td>
<td>My teacher likes it when we ask questions.</td>
<td>PT-C4</td>
</tr>
<tr>
<td>11</td>
<td>My teacher makes materials and worksheets for us to use.</td>
<td>PT-C9</td>
</tr>
<tr>
<td>1</td>
<td>My teacher makes class work interesting.</td>
<td>PT-C3</td>
</tr>
<tr>
<td>8</td>
<td>My teacher gives feedback about my performance.</td>
<td>PT-C5</td>
</tr>
<tr>
<td>5</td>
<td>My teacher tells us how we can use what we have already learned to learn new things.</td>
<td>PT-C2</td>
</tr>
<tr>
<td>7</td>
<td>My teacher returns tests and assignments quickly.</td>
<td>PT-C5</td>
</tr>
<tr>
<td></td>
<td><strong>Factor 2</strong></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>My teacher uses a variety of classroom activities and resources.</td>
<td>PT-C1</td>
</tr>
<tr>
<td>14</td>
<td>My teacher tells the class about library/media materials that will help us learn about the subject we are studying, when appropriate.</td>
<td>OSC-C12</td>
</tr>
<tr>
<td>13</td>
<td>The films or videotapes we watch help us learn about the subject we are studying.</td>
<td>PT-C9</td>
</tr>
<tr>
<td>17</td>
<td>We work in different groups depending upon the activity in which we are involved.</td>
<td>PT-C15</td>
</tr>
<tr>
<td>18</td>
<td>My teacher encourages us to look at problems in new ways and find new ways to solve problems.</td>
<td>PT-C3</td>
</tr>
</tbody>
</table>

*PT=Productive Teaching Techniques, OSC=Organized, Structured Classroom Management, PI=Positive Interpersonal Relations.*
Table 26. Comparison of items by factor analysis clusters and Judkins (1987) SIM performance areas, 1990 middle school regular questionnaire

<table>
<thead>
<tr>
<th>Number</th>
<th>Item</th>
<th>SIM criteriaa</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Factor 1</strong></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>My teacher makes classwork interesting.</td>
<td>PT-C3</td>
</tr>
<tr>
<td>12</td>
<td>My teacher explains new ideas in a way that is easy to understand.</td>
<td>PT-C4</td>
</tr>
<tr>
<td>2</td>
<td>My teacher is fair with all.</td>
<td>OSC-C14</td>
</tr>
<tr>
<td>19</td>
<td>My teacher uses a variety of classroom activities and resources.</td>
<td>PT-C1</td>
</tr>
<tr>
<td>10</td>
<td>My teacher starts lessons explaining what we are going to do and why we are going to do it.</td>
<td>PT-C2</td>
</tr>
<tr>
<td>14</td>
<td>My teacher knows more about the subject than other teachers I have had.</td>
<td>PT-C7</td>
</tr>
<tr>
<td>8</td>
<td>My teacher likes it when we ask questions.</td>
<td>PT-C4</td>
</tr>
<tr>
<td>13</td>
<td>My teacher looks at our work, as we are doing it, to see if we understand the lesson.</td>
<td>PT-C2</td>
</tr>
<tr>
<td>15</td>
<td>My teacher has work for me to do if I finish an assignment before the class is over.</td>
<td>PT-C9</td>
</tr>
<tr>
<td></td>
<td><strong>Factor 2</strong></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>My teacher gives assignments related to the subject we are studying.</td>
<td>PT-C2</td>
</tr>
<tr>
<td>7</td>
<td>Our discussions focus on the topic of the lesson.</td>
<td>PT-C10</td>
</tr>
<tr>
<td>17</td>
<td>My teacher gives tests and quizzes.</td>
<td>PT-C6</td>
</tr>
<tr>
<td>6</td>
<td>We discuss and summarize each lesson just studied.</td>
<td>PT-C2</td>
</tr>
<tr>
<td>4</td>
<td>My teacher is well prepared for our class.</td>
<td>OSC-C13</td>
</tr>
<tr>
<td>11</td>
<td>My teacher asks us questions to see if we understand what has been taught.</td>
<td>PT-C2</td>
</tr>
<tr>
<td>3</td>
<td>My teacher maintains discipline in our classroom.</td>
<td>OSC-C14</td>
</tr>
<tr>
<td>16</td>
<td>My teacher often makes materials and worksheets for us to use.</td>
<td>PT-C9</td>
</tr>
<tr>
<td>18</td>
<td>My teacher returns tests and assignments quickly.</td>
<td>PT-C5</td>
</tr>
<tr>
<td></td>
<td><strong>Factor 3</strong></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>My teacher gives enough time to do our work.</td>
<td>PT-C1</td>
</tr>
<tr>
<td>9</td>
<td>I have more time to do my work than I need.</td>
<td>PT-C1</td>
</tr>
</tbody>
</table>

aPT=Productive Teaching Techniques, OSC=Organized, Structured Classroom Management, PI=Positive Interpersonal Relations.
Table 27. SIM performance area criteria

<table>
<thead>
<tr>
<th>Criterion number</th>
<th>Criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The teacher demonstrates effective planning skills.</td>
</tr>
<tr>
<td>2</td>
<td>The teacher implements the lesson plan.</td>
</tr>
<tr>
<td>3</td>
<td>The teacher motivates students.</td>
</tr>
<tr>
<td>4</td>
<td>The teacher communicates effectively with students.</td>
</tr>
<tr>
<td>5</td>
<td>The teacher provides students with specific evaluative feedback.</td>
</tr>
<tr>
<td>6</td>
<td>The teacher prepares appropriate evaluation activities.</td>
</tr>
<tr>
<td>7</td>
<td>The teacher displays a thorough knowledge of curriculum and subject matter.</td>
</tr>
<tr>
<td>8</td>
<td>The teacher selects learning content congruent with the prescribed curriculum.</td>
</tr>
<tr>
<td>9</td>
<td>The teacher provides opportunities for individual differences.</td>
</tr>
<tr>
<td>10</td>
<td>The teacher ensures student time on task.</td>
</tr>
<tr>
<td>11</td>
<td>The teacher sets high expectations for student achievement.</td>
</tr>
<tr>
<td>12</td>
<td>The teacher plans for and makes effective use of time, materials and resources.</td>
</tr>
<tr>
<td>13</td>
<td>The teacher demonstrates evidence of personal organization.</td>
</tr>
<tr>
<td>14</td>
<td>The teacher sets high standards for student behavior.</td>
</tr>
<tr>
<td>15</td>
<td>The teacher organizes for effective instruction.</td>
</tr>
<tr>
<td>16</td>
<td>The teacher demonstrates effective interpersonal relationships with others.</td>
</tr>
<tr>
<td>17</td>
<td>The teacher demonstrates awareness of the needs of students.</td>
</tr>
<tr>
<td>18</td>
<td>The teacher promotes positive self-concept.</td>
</tr>
<tr>
<td>19</td>
<td>The teacher demonstrates sensitivity in relating to students.</td>
</tr>
<tr>
<td>20</td>
<td>The teacher promotes self-discipline and responsibilities.</td>
</tr>
</tbody>
</table>

Judkins (1987) developed items for three performance areas, only two factors were identified for the senior high school questionnaire items. Second, this finding is not totally surprising since neither Table 25 nor Table 26 show that no questionnaire items were selected by the Cave Creek Stakeholders Committee for the third performance area of Positive Interpersonal Relations. Third, it seems reasonable to associate Factor 1 with Productive Teaching Techniques and Factor 2 with Organized, Structured Classroom Management. And finally, items 6 and 15 would have been expected to be found in Factor 2 but were identified as belonging in Factor 1.

The information contained in these three tables makes it is possible to discuss the findings of the factor analysis with some degree of prior knowledge and expectation. First, Judkins (1987) developed items for three performance areas and three factors were identified for the middle school regular questionnaire items. Second, Table 26 shows that, as was the case with the senior high questionnaire, no discriminating items were selected by the Cave Creek Stakeholders Committee for the third performance area of Positive Interpersonal Relations. Third, it seems reasonable to associate Factor 1 with Productive Teaching Techniques and Factor 2 with Organized, Structured Classroom Management. Even though both items 9 and 20 are identified as being associated with Productive Teaching Technique criterion 1: The teacher demonstrates effective planning skills, Factor 3 appeared to focus upon the time given for a student to complete assigned work and can therefore be associated with Positive Interpersonal Relations. And finally, item 2 would have been
expected to be found in Factor 2 but was identified as belonging in Factor 1.

Even though the factor solutions conducted in this study did not show the expected match or fit with Judkins' (1987) performance areas or the SIM criteria, the results are still important. While variables load on more than one factor, none of the variables tended to have large loadings on more than one factor. The factor analysis showed that it was possible to identify a relatively small number of factors that could be used to represent the relationships among sets of interrelated variables. Specifically, this approach to factor analysis is one of the first attempts to examine the relationship of discriminating items that are selected for use in feedback questionnaires to the original performance criteria for which they were developed to measure. That it is possible to identify specific clusters on the basis of underlying common characteristics is a significant finding.
CHAPTER V. SUMMARY, CONCLUSIONS, LIMITATIONS, DISCUSSION, AND RECOMMENDATIONS

This study was conducted in 1990 and 1991 with the purpose of identifying, refining, and testing three sets of student feedback items comprising separate questionnaires for use as one component in a total teacher performance evaluation system. To accomplish this task, 826 secondary school students rated 18 middle school teacher and 27 senior high school teachers. This non-voluntary sample, consisting of the total population of a single school system, is an important element of this study because all previous and related research studies relied upon volunteers as subjects. Consequently, the analysis of this study's data has served as a desirable extension of the original School Improvement Model (SIM) directed work undertaken by Hidlebaugh in 1973, Judkins in 1987, and the ongoing SIM improvements.

A Career Ladder Stakeholders Committee in the Cave Creek (Arizona) Unified School District No. 93 developed three questionnaires each consisting of 20 items and requiring students to rate teacher performance using a five-point Likert-type response mode. Most of the items selected by the committee had been previously developed by Judkins (1987) and had been identified as being valid, reliable, and legally discriminating. Some of the items, however, were developed by local teachers and administrators.

The items which were shown to have discriminated between teachers' performance were identified according to criteria previously validated by Hidlebaugh (1973) and Judkins (1987). Specifically, the Menne and
Tolsma (1971) methodology of analyzing data was used to identify those items possessing discriminating power. A sum of squares between-groups difference equal to or exceeding 13 percent of the variance for total sums of squares was the criterion established at the .05 level of significance, and 22 percent was the criterion established at the .01 level of significance. Cronbach's coefficient alpha method was also used to calculate a reliability coefficient and to estimate the internal consistency, at the .05 level of significance, of all discriminating items on each questionnaire (Appendix N and Appendix O).

Analysis of variance and t-tests for groups were the inferential statistical tests used to determine whether the differences between sample means reflect population differences. Planned orthogonal contrasts and the Scheffé multiple range test were used to locate the means which differed significantly after a significant F-ratio had been identified in the ANOVA. Specifically, these inferential techniques were used to determine the effect of specified extraneous variables on student ratings of teacher performance. Further, because Judkins (1987) originally developed items for three performance areas and since the items developed by the Cave Creek School District's teachers and administrators could also be classified according to the SIM criteria of 1) Productive Teaching Techniques, 2) Organized, Structured Classroom Management, and 3) Positive Interpersonal Relations, a factor analysis was conducted on the data for each of the two regular questionnaires. Specifically, the factor analysis was used to describe whether the same number of factors was identified in
the present study and whether the items contained in each factor were similarly grouped according to Judkins' original performance areas.

Summary

An analysis of the data indicated:

1. Twenty out of 20 items on the 1990 and 1991 6-8 middle school regular and 9-12 senior high questionnaires discriminated or measured significant differences between the 45 teachers rated by students. All items met the 13 percent criteria and many items exceeded the 22 percent criteria.

2. The 1990 6-8 floater special-area questionnaire contained 12 of 20 items that had a between-groups variance equal to or exceeding the 13 percent criteria, and four of 20 items equalled or exceeded the 22 percent criteria. Item 5, "My teacher gives assignments related to the subject we are studying"; item 14, "We do the same thing everyday in class"; item 17, "My teacher gives tests and quizzes"; and item 18, "My teacher returns tests and assignments quickly" were the only items failing to meet the 13 percent criteria.

3. Nine items were developed in 1990 and 1991 by local teachers and administrators for the student feedback questionnaires. One hundred percent of the locally-developed items had a sum of squares between-group variance equal to or exceeding 13 percent criteria and consequently discriminated or measured differences between teachers' performance.

4. The Cronbach alpha reliability coefficient computed for each rater group indicated that high internal consistency existed among the
rating of items for each of the questionnaires. Specifically, coefficients for the 1990 data were: 0.935 for the middle school regular, 0.950 for the middle school floater special-area, and 0.971 for the senior high school. Further, the coefficients for the 1991 data were: 0.948 for the middle school regular and 0.950 for the senior high school.

5. Grades 11 and 12 student mean score rating of teachers was nearly five points higher than grades 6 and 7 student mean score rating of teachers.

6. The student mean score ratings of teachers were significantly different according to the subject area being taught. No significant differences existed between the student mean score ratings in the core subjects of math and science, and English and social studies. However, significant differences did exist between the student mean score ratings in fine arts classes and the other remaining subjects.

7. No significant differences were found in the student mean score ratings of teacher performance for the three groups or levels of questionnaires.

8. No significant differences were found in the student mean score ratings of teacher performance for the various periods of the day.

9. The student mean score rating for core subjects was significantly different than the student mean score rating for elective subjects. Specifically, the mean score rating for the core subjects was nearly four points higher than the mean score rating for elective subjects.

10. A significant difference was found between student mean score ratings for career ladder participants and non-career ladder teachers.
Specifically, the mean score rating for the career ladder participants was over four points higher than for non-career ladder teachers.

11. The student mean score rating for female teachers was nearly seven points higher than for male teachers. Consequently, a significant difference was found to exist in student mean ratings when grouped by gender.

12. The factor analysis identified two clusters for the senior high data and three clusters for the middle school data. Therefore, it was possible to identify factors which could be used to represent relationships among sets of the discriminating items. Specifically, Factor 1 was Productive Teaching Techniques, Factor 2 was Organized, Structured Classroom Management, and Factor 3 was Positive Interpersonal Relations.

Conclusions

The following conclusions are offered based on the analysis of data collected in this investigation.

1. The Menne and Tolsma methodology (1971) for determining the discrimination power of items on instruments using group rater responses can be used to identify discriminating items for the purpose of refining student feedback items and student rating of teacher performance instruments.

2. All items, except four of the 1990 middle school floater special-area items (item 4, "Gives related assignments"; item 14, "Do the same thing everyday"; item 17, "Gives tests and quizzes"; and item 18, "Returns
tests and assignments quickly"), were identified as being able to discriminate or measure differences among teacher performance based on groups of 15 or more raters. Further, there is a large range in the discriminating power of the items on the student rating of teacher performance instruments used in this study.

3. A Cronbach alpha reliability coefficient was calculated for each set of discriminating items. Further, Cronbach alpha coefficients for pools of items on similar studies utilizing the same methodologies were available for comparison. Specifically, Green (1990) reported a coefficient of .974; Look (1983) reported a coefficient for the combined rater group of .982; and Judkins (1987) reported coefficients of .859 for high school and .839 for junior high school. These high reliability coefficients indicated the items contained in the sets of items were consistently measuring what they intended to measure.

The Cronbach alpha reliability coefficients computed for the 1990 data in this study were: 0.935 for the middle school regular questionnaire, 0.950 for the middle school floater special-area questionnaire, and 0.971 for the senior high school questionnaire. Further, the coefficients for the 1991 data were: 0.948 for the middle school regular and 0.950 for the senior high school. By comparison, the items used in the student ratings of teacher performance in this current investigation also consistently measured what they intended to measure.

4. The majority of items previously identified by Judkins (1987) continued to demonstrate discriminating power when used by students to rate teacher performance. Further, local teachers and administrators were
able to develop items that also possessed discriminating power. Consequently, volunteer and ad hoc committees can, with appropriate assistance, be useful in developing discriminating items which can be used by students to rate teacher performance.

5. Students can reliably serve as one important source of information for the rating of teacher performance.

6. Grade level, core and elective nature of the course, subject area taught, career ladder participation status, and teacher gender were extraneous variables found to have an effect upon the students' mean score ratings of teachers. The findings suggest that the highest student ratings were assigned to career ladder, core subject, and female teachers in the senior grades. It is not possible to determine whether this result was due to optimum teacher performance or the effect of other factors.

7. The level of questionnaire being used, and the period or time of day, did not have an extraneous effect on the student mean score ratings of teacher performance. Therefore, similar teaching performance can be considered to occur at all times of the day for each grade level.

8. Factor analysis can be used to identify clusters of items or factors (Productive Teaching Techniques; Organized, Structured Classroom Management; and Positive Interpersonal Relations) which can then be used to represent relationships among sets of the discriminating items.

9. Items originally developed by Judkins (1987) did not necessarily load upon the expected set of three performance criterion areas or categories. Such a finding is not unexpected due to the processes used to select the items and to develop the questionnaires.
Limitations

Certain limitations were imposed due to the design of this study. They were:

1. All subjects were members of a single Arizona school system where participation in the career ladder was voluntary. There may be unique attributes, not controlled for in this study, of career ladder teachers in a school system that implement a voluntary career ladder process.

2. In order to obtain a total sample of non-volunteers, a school system in which all teachers were rated by their students was chosen as the study population. No previous studies focusing upon students' ratings of teacher performance have been conducted using a non-voluntary sample consisting of a total school system population. Consequently, the stakeholders committee had a significant and direct influence upon the questionnaires to be administered and the processes for administering the questionnaires.

3. The performance level of the teachers was not assessed independent of the questionnaire results. That is to say, this investigation focused on the items, not the teacher as the unit of study.

4. No attempt was made to determine whether the ability and performance level of students affected their rating of teacher performance. Student demographic data were not available or collected. Again, this investigation focused on the items and not the student as the unit of study.

5. The discrimination value of an item does not reflect high or low performance, nor does it indicate which teacher behaviors are most
associated with effective teaching. This methodology simply provides a means to determine how well an item measures differences between teachers' performances. At the same time, the combined total of student mean score ratings for each teacher was used as one component of the career ladder algorithm.

6. Teacher subjects were able to choose whether to participate in the career ladder. Different levels of motivation and attitudes towards student ratings of teachers may have existed as a result of the teachers' choice.

7. Teachers were rated by at least 15 students. Due to the large number of raters for whom data were analyzed, it is believed that the items identified as possessing discriminating power are representative of items which measure differences between teachers. However, the same items may not be discriminating among teachers rated by fewer raters or for any two specific teachers within the group.

8. Judkins (1987) developed items for all three SIM identified teacher performance criteria areas; however, in this study no items designated as belonging to performance area three, were chosen from the available pool of items. Therefore, when conducting the factor analysis, there was no expectation that items would cluster or load upon the factor known as Positive Interpersonal Relations.

9. Elimination of the use of the middle school floater special-area questionnaire in 1991 prevented the comparative analysis of two successive years of data. Since the middle school population contained all teachers
in the 1991 data, any comparisons could be simply attributed to the addition of the new teachers.

10. This study was conducted during the initial stages of implementing a revised career ladder plan. External and internal pressures may have influenced the results.

Discussion

The State of Arizona Legislature responded to students' declining academic performance in 1984 by passing a bill intended to support incentive pay for teachers. Specifically, this initiative originally produced 16 career ladder pilot projects and currently 14 projects remain operational. On the basis of data collected during an evaluative investigation of the Arizona career ladder initiative and the effects of incentive pay on student achievement, Braver and Helmstadter (1990) concluded that,

1) An increase in student achievement occurred in the Career Ladder districts following the implementation of the Career Ladder Pilot Program, whereas the non-Career Ladder districts showed no increase.
2) Further, this increase remained consistent from year to year following Career Ladder implementation.

These findings, reported by Braver and Helmstadter (1990), affirm L'Hommedieu's (1990) assertion that researchers who address the question of improving instruction through systematic feedback are exploring an issue of immense practical value. The rationale for including student ratings of teacher performance in a career ladder algorithm is enhanced when it can be demonstrated that valid, reliable and legally
discriminating items and instruments can be developed for use by the students.

The original items developed by Judkins (1987) continued to demonstrate discriminating power. In addition, the high reliability coefficients, consistent with the previous findings of Green (1990), Judkins (1987), and Look (1983), indicated that the items used in this current investigation also continued to measure what they intended to measure. At the same time, such high reliability coefficients may be used by some to argue that the items simply measure student opinion to a single generalized criteria such as, "I like my teacher." The credibility of this argument is lowered by the results of the factor analysis which indicated the tendency of the items to cluster into identifiable groups.

Those few items that did not retain discrimination power were tested in the middle school floater special-area questionnaire. The homogeneous nature of the students, the smaller class sizes, unique instructional structures, and procedures or other unidentified factors may have been responsible for the inability of four items to discriminate. However, it is interesting to note that all five items developed by local teachers and administrators for use in the middle school floater special-area questionnaire were identified as possessing discriminating power.

It was expected that the selection of a non-voluntary sample consisting of all teachers and students in a single school system would enhance the ability of the original Judkins (1987) items to allow student raters to discriminate between teachers' performances. Further, it was expected that with the support and assistance of the SIM Team, local
teachers and administrators could develop items which would be shown to possess valid and reliable discriminating power. This was a desirable action to promote local ownership of the student feedback process and instruments. The findings of this study affirm that both expectations were reasonable.

Ideally, according to Aleamoni (1987b), one would expect that extraneous variables such as grade level, subject area, core versus elective, career ladder status, and teacher gender would not have an undue effect upon the students' mean score ratings of teachers. However, it was possible to speculate that senior students and their teachers might both approach the acts of teaching, learning, and the rating of teacher performance in a more serious manner than would younger students.

In contrast, a common expectation is that students in elective courses would rate their teachers higher than in other classes because of the students' interest in the self-selected course. The findings of this study have shown that such a popular expectation is not supported. Two possible explanations exist which might be used to explain the findings: 1) perhaps students are not able to truly choose their electives, or 2) perhaps the core subject teachers and their students, as a group, approach their teaching and learning in a more serious and effective manner.

It was expected that career ladder participants would receive higher ratings than non-career ladder teachers. An initial speculation was that teachers who chose to voluntarily participate in the career ladder plan would posses higher levels of motivation, increased confidence and
ability, and would approach all aspects of the algorithm, specifically student ratings, more seriously. An additional speculation that occurred during the analysis of these data was made concerning certain non-career ladder teachers who received unusually high mean score ratings from students. Perhaps these teachers possess some unique reasons, such as length of service, proximity to retirement age, or other commitments, that have caused them to refrain from participating in the career ladder plan.

The findings of the 1991 NEA study conducted by Chopade and Lackritz, "Student Evaluations: Equal Opportunity Concerns," indicated that female faculty were favored over male faculty by the student raters. Therefore, an expectation existed for a difference in the students' mean score ratings by gender of teacher. Having identified that a difference did exist, it would be useful to determine through further research whether this difference can be best attributed to increased teacher performance by female teachers or a possible rater bias on the part of students.

And finally, an initial attempt was made to identify specific clusters on the basis of underlying common characteristics of discriminating items. It is likely that the factors that the items load upon represent unique clusters or teaching performance criteria. Such findings are especially useful when revisiting the original classification of items according to the SIM performance areas and criteria.

Overall, the findings of this study complement the position taken by Scriven (1990) that student ratings, if gathered in a suitable secure way, using a suitably designed form, can provide a useful basis for rating teachers. Even with the additional questions raised concerning the effect
of extraneous variables, this study clearly reaffirms the ability of students to rate teacher performance when using feedback questionnaires consisting of valid and reliable discriminating items. Students can provide valuable information regarding teacher performance. Such a valuable source of information should be used to develop a more complete picture of teacher performance in order that gains in student performance can be achieved.

Recommendations for Practice

The results of this study point to several suggestions for teachers, administrators, and superintendents.

1. When selecting items for students' use in rating teacher performance, only those items shown to possess discriminating power should be selected. Further, items should be selected because they match the district's philosophy, policies, and beliefs regarding effective teaching practice.

2. Discriminating items should be used within the same grade levels for which they were tested. If these items are to provide information that will discriminate between teacher performances then, due to concerns of readability, the items should be used at the grade levels for which they were developed. (See suggested sets in Appendix E and Appendix G.)

3. Student ratings should be used as one component of a total teacher performance evaluation system. Students are able to provide valuable feedback that can be used to improve and assess teacher performance.
4. Each teacher should be rated by at least 15 students in each class, program, or subject area. Whenever fewer than 15 raters are used, the resulting data should be analyzed to determine whether the items continue to possess discriminating power.

5. It is important that all aspects of effective teaching practice are rated. Even though current concerns exist regarding child abuse, to the extent that some institutions require that office doors remain open during interviews or that a witness be present during conferences, the practice of teachers serving as confidants and a source of caring remains highly desirable. Therefore, consideration should be given to developing, validating, and subsequently adding discriminating items which match the criteria from the Positive Interpersonal Relations performance area. Essentially, such items should rate the teacher’s ability to provide for positive student-teacher interaction, communication, and student confidentiality.

6. The directions and procedures developed for administering the questionnaires for this study should be followed. These directions and procedures are intended to protect the integrity of the student ratings of teacher performance.

7. District and national norms should be recorded and established. Students’ mean score ratings of teacher performance can be appropriately compared once a sufficient number of mean scores have been collected for each grade level.

8. Teachers should be encouraged to use the students' rating information as one source of valuable information when developing teacher
improvement plans. Detailed information can be made available to each individual teacher by question and class period.

Recommendations for Further Research

The findings of this study suggest further research. In each suggested study, the sample size should be as large as possible. In addition, whenever possible non-volunteers should be used as the study population.

1. This study was intended to be a refinement of the work undertaken by Hildebaugh (1973) and Judkins (1987). Further research is needed to ascertain if the items identified as being discriminating in this study would also possess discriminating power when used in other school districts. Such investigations could be conducted in other career ladder districts, and if possible, with a non-voluntary sample of teachers and students.

2. The relationship between teacher effectiveness and ratings of the items should be explored. Specifically, the other components of the Cave Creek (Arizona) career ladder algorithm should be compared with student ratings. It would be highly desirable to explore the correlations among principal and peer evaluation scores, student achievement scores, parental ratings, and student ratings of teacher performance. By examining such correlations, it would be possible to describe whether the significant differences are best attributed to the effect of specific extraneous variables or specific levels of individual teacher performance. Such an investigation would be especially useful in examining the question: Are
senior high school female teachers more effective than male senior high school teachers?

3. Additional student demographic data should be collected and used to determine whether the mean score ratings of teachers are affected by the students' age, sex, academic achievement, or gender. Limited research has been conducted on student leniency bias at the secondary level. An associated study (Weber, 1992) could be used as a model research design for such an investigation.

4. Further exploration of the use of factor analysis should be applied to the data which have been collected to more closely examine the similarities and differences with Judkins' (1987) teacher performance criteria areas. The information developed from such further investigation could then be used to continue the refinement of the existing questionnaires.

5. A longitudinal study should be conducted to determine whether teachers are able to improve their performance in areas identified as requiring improvement. The use of students as a valuable source of such information would be further enhanced if such a study were undertaken.

6. A case study approach should be used to profile the teaching practices of those teachers receiving the highest mean score ratings from students. Subjects could be observed, interviewed, and anecdotal descriptions could be obtained from supervisors, peers, and students. These profiles could be further subdivided by specific items and the information could then be related to specific performance areas and
criteria. Subsequently, such information could be used to enhance staff development programs and teacher training programs.

7. The questions regarding the high reliability coefficients should be addressed in a future investigation. A test and re-test study design could be used to ensure that the items are indeed measuring what they are intended to measure.

8. An investigation should be conducted in a district which uses pre- and post-test measures to assess student gains in learning and achievement. Such a study would explore, validate, and describe the correlation between student ratings of teacher performance and student performance and achievement.
BIBLIOGRAPHY


Many people will rightfully share in the celebration of success which results from the completion of my doctoral program. I wish to openly acknowledge the influence of my father and mother who instilled in me an insatiable desire to learn, teach, and serve. Their beliefs and positive actions have instilled in me an optimistic expectation for each day of life.

Several colleagues and mentors were responsible for planting the seed of belief that I was capable of meeting the challenges of this endeavor. I wish to acknowledge my good fortune and privilege to have been prepared, encouraged, and supported by such notable educators as Dr. Kathy Skau, Dr. Lawrence Tymko, Dr. Reno Bosetti, Dr. Bill Duke, Dr. Ken Sauer, and Dr. Harold Storlien. I sincerely hope that my efforts and work meet their high expectations and standards of quality.

I especially want to extend my sincerest appreciation to the members of my doctoral committee. It was because of their efforts, commitment, and support that I consider my experiences at Iowa State University to be truly exceptional. Dr. Richard P. Manatt personified everything I had ever envisioned in a major professor. Professor Manatt's commitment to being available, regardless of the time, day, or personal schedule, was without precedent. Further, by socializing me into the culture of "The Long Grey Line," my life has been positively changed forever. Dr. Shirley Stow represented all that was good and right in a professional educator. She was a source of inspiration, sound counsel, and constant
encouragement. Dr. Anton Netusil caused me to draw upon every ability I possessed in order to successfully comprehend and apply research statistics. At the same time, his individual and personal support of my research efforts was sincere, invaluable, and very much appreciated. His goal of having me use what was learned, in the real world of schools, has been attained. Dr. Dan Reschly caused me to think about the lives of students. He provided me with the opportunity to reflect upon what was important in our daily work. Dr. Russ Mullen reminded me that excellent teachers continue to demonstrate a desire to learn. Regardless of his own needs and time constraints, Professor Mullen was always willing to listen and engage in an ongoing dialogue regarding strategies for effective teaching. I arrived at Iowa State University being a believer in teamwork and collaborative effort. I leave the doctoral program grateful for the opportunity of having so many high quality cooperative and collaborative experiences.

I also wish to extend my appreciation to the members of the education administration team. The support and friendship of fellow R.A.'s will be a part of me forever. In particular I wish to formally acknowledge my study partners, Pete Price and Barb Weber, who exemplified the theme of the Three Musketeers--"All for one and one for all!" To Katy Rice, Judy Welland, Marjorie Smith, Kris Benyshek, and Jamie Blomgren countless words of thanks and appreciation for all of the ongoing support, extra effort, and the personal touch. To all of my professors my sincerest appreciation for sharing of your knowledge, skills, and experiences. I trust I have learned your lessons well.
A special note of recognition and appreciation to Dr. Jim Sweeney, who encouraged me to think differently and taught me the true meaning of reflective leadership. Professor Sweeney also provided me with the opportunity to experience the Danforth Leadership T.E.A.M. To Dave Black and his colleagues—"Expect the Best!"

To Steve Williams and the staff of Marshalltown Community Schools, I wish to extend my appreciation for their openness, support, and friendship. The opportunity to observe, from the inside, a successful Iowa school system and NEA Shared Decision Making Learning Lab was a privilege and an invaluable learning experience. I hope that I was able to contribute as much as I received.

I would also like to thank Dr. David Alexander and the staff of Cave Creek Unified School District No. 93 for their cooperation and willingness to serve as a Beta site for the student ratings of teacher performance items. And finally, I wish to acknowledge the time, effort, and work of Bonnie Trede, who provided exceptional service in the final preparation of this dissertation.

My time and experiences at Iowa State University have been exceptional. I have grown personally and professionally. Further, I have renewed my commitment to the values of caring, trust, loyalty, teaching, and service. I now value friendship and each day of life more than ever before.
APPENDIX A.

K-2 CAVE CREEK REGULAR STUDENT RATING
OF TEACHERS QUESTIONNAIRE
**Student Feedback to Teachers (Lower Elementary, K-2)**

0. I like the color red. {No O} {? O} {Yes O}

1. My school day is interesting. {No O} {? O} {Yes O}

2. We do the same thing every day in class. {No O} {? O} {Yes O}

3. I pay attention in class. {No O} {? O} {Yes O}

4. Our discussions are about the lesson being studied. {No O} {? O} {Yes O}

5. Our work is too hard for us. {No O} {? O} {Yes O}

6. My teacher is usually prepared for class. {No O} {? O} {Yes O}

7. My teacher comes to class on time. {No O} {? O} {Yes O}

8. My teacher makes me follow the rules. {No O} {? O} {Yes O}

9. My teacher is fair with everybody. {No O} {? O} {Yes O}

10. My teacher cares if I waste time in class. {No O} {? O} {Yes O}

11. I work in this class even if the teacher is not watching. {No O} {? O} {Yes O}

12. I can get help from my teacher when I need it. {No O} {? O} {Yes O}

13. My teacher tells me that I do good work. {No O} {? O} {Yes O}

14. My teacher tells me where I can find more information to help me learn about the lesson. {No O} {? O} {Yes O}

15. My teacher is ready for class when it is time to begin. {No O} {? O} {Yes O}

16. I know what the teacher wants us to do. {No O} {? O} {Yes O}

17. My teacher is easy to understand. {No O} {? O} {Yes O}

18. My teacher has us learn hard lessons in small steps. {No O} {? O} {Yes O}

19. My teacher will explain new things in a way that is easy to learn. {No O} {? O} {Yes O}

20. My teacher tells us what new things we can learn in each lesson. {No O} {? O} {Yes O}
STUDENT FEEDBACK TO SPECIAL AREA TEACHERS
LOWER ELEMENTARY SCHOOL QUESTIONNAIRE (K-2)

Because lower elementary students may experience difficulty in reading their own directions, the adult proctor will read:

Note to students: Please remember that completing this form is voluntary. You may keep this form if you decide not to participate.

Directions: The statements on your sheet are designed to find out more about your class and teacher. For each question or statement, fill in the circle after each statement that best describes this class or teacher. This is not a test. Do not put your name on this paper or answer sheet. Please answer all the statements. Carefully listen to directions for marking answers. Students are not to ask any questions during the survey.

CAREFULLY FILL IN THE "NO" CIRCLE  
if the statement does not describe your class or teacher at all.

CAREFULLY FILL IN THE "?" CIRCLE  
if the statement describes your class or teacher the way it is sometimes.

CAREFULLY FILL IN THE "YES" CIRCLE  
if the statement describes your class or teacher the way it is almost all of the time.

NOW LET'S PRACTICE on the first item marked 0 (zero) at the top of your sheet.

0. I like the color red. {No O} {? O} {Yes O}

Notice that some of you may mark "yes" and some of you may mark "no", while others may mark "?" because each of you may have a different opinion about red. All of the questions you will answer today are your opinions and you may each answer differently for each question.
Student Feedback to Teachers (Lower Elementary, K-2)

0. I like the color red. {No O} {? O} {Yes O}
1. My school day is interesting. {No O} {? O} {Yes O}
2. My teacher gives us enough time to do our work. {No O} {? O} {Yes O}
3. I pay attention in class. {No O} {? O} {Yes O}
4. Our discussions are about the lesson being studied. {No O} {? O} {Yes O}
5. Our work is too hard for us. {No O} {? O} {Yes O}
6. My teacher gives us homework. {No O} {? O} {Yes O}
7. My teacher comes to class on time. {No O} {? O} {Yes O}
8. My teacher makes me follow the rules. {No O} {? O} {Yes O}
9. We often have to take a test in class. {No O} {? O} {Yes O}
10. My teacher cares if I waste time in class. {No O} {? O} {Yes O}
11. I work in this class even if the teacher is not watching. {No O} {? O} {Yes O}
12. I can get help from my teacher when I need it. {No O} {? O} {Yes O}
13. My teacher gives me new work to do when I am ready for it. {No O} {? O} {Yes O}
14. My teacher tells me where I can find more information to help me learn about the lesson. {No O} {? O} {Yes O}
15. My teacher is ready for class when it is time to begin. {No O} {? O} {Yes O}
16. I know what the teacher wants us to do. {No O} {? O} {Yes O}
17. My teacher gives me interesting work if I finish my work before class is over. {No O} {? O} {Yes O}
18. My teacher has us learn hard lessons in small steps. {No O} {? O} {Yes O}
19. My teacher gives our work back to us quickly. {No O} {? O} {Yes O}
20. My teacher tells us what new things we can learn in each lesson. {No O} {? O} {Yes O}
APPENDIX C.

3-5 CAVE CREEK REGULAR STUDENT RATING OF TEACHERS QUESTIONNAIRE
### STUDENT FEEDBACK TO TEACHERS
#### UPPER ELEMENTARY SCHOOL QUESTIONNAIRE (3-5)

**NOTE TO STUDENTS:** Please remember that completing this form is voluntary. You may keep this form if you decide not to participate.

**Directions:** The statements below are designed to find out more about your class and teacher. This is not a test. Do not put your name on this paper. Please answer all the statements. Students are not to ask any questions during the survey.

1 = Never  
2 = Not often  
3 = Sometimes  
4 = Usually  
5 = Almost always

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0. I like to eat ice cream.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>1. My teacher makes our work interesting.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>2. My school day is interesting.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>3. We go back over each lesson when we finish it.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>4. My teacher gives us work to do at home.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>5. Our discussions are about the subject being studied.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>6. My teacher gives our work back to us quickly.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>7. My teacher makes me feel good when I do good work.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>8. I can get help from my teacher.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>9. I finish my work before class is over.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>10. My teacher makes me follow the rules.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>11. My teacher gives me new work to do without having to wait a long time for it.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>12. My teacher explains the lesson clearly.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>13. My teacher knows me well.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>14. My teacher has work for me to do if I finish my assignment before class is over.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>15. My teacher has us work at the right pace.</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>16. My teacher tells us what new things we can learn in each lesson.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>17. My teacher will explain new things in a way that is easy to understand.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>18. My teacher is available to help me during class time and other times during the school day.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>19. My teacher uses a variety of classroom activities and resources.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>20. My teacher is well-prepared.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
</tbody>
</table>
APPENDIX D.

3-5 CAVE CREEK FLOATER STUDENT RATING OF TEACHERS QUESTIONNAIRE
**STUDENT FEEDBACK TO SPECIAL AREA TEACHERS**
**UPPER ELEMENTARY SCHOOL QUESTIONNAIRE (3-5)**

NOTE TO STUDENTS: Please remember that completing this form is voluntary. You may keep this form if you decide not to participate.

**Directions:** The statements below are designed to find out more about your class and teacher. This is not a test. Do not put your name on this paper. Please answer all the statements. Students are not to ask any questions during the survey.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Scale (1 = Never, 2 = Not often, 3 = Sometimes, 4 = Usually, 5 = Almost always)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0. I like to eat ice cream.</td>
<td>0 0 0 0 0</td>
</tr>
<tr>
<td>1. My teacher makes our work interesting.</td>
<td>0 0 0 0 0</td>
</tr>
<tr>
<td>2. My school day is interesting.</td>
<td>0 0 0 0 0</td>
</tr>
<tr>
<td>3. We go back over each lesson when we finish it.</td>
<td>0 0 0 0 0</td>
</tr>
<tr>
<td>4. We do the same thing everyday in class.</td>
<td>0 0 0 0 0</td>
</tr>
<tr>
<td>5. Our discussions are about the subject being studied.</td>
<td>0 0 0 0 0</td>
</tr>
<tr>
<td>6. My teacher is usually prepared for class.</td>
<td>0 0 0 0 0</td>
</tr>
<tr>
<td>7. My teacher makes me feel good when I do good work.</td>
<td>0 0 0 0 0</td>
</tr>
<tr>
<td>8. I can get help from my teacher.</td>
<td>0 0 0 0 0</td>
</tr>
<tr>
<td>9. My teacher is fair with everybody.</td>
<td>0 0 0 0 0</td>
</tr>
<tr>
<td>10. My teacher makes me follow the rules.</td>
<td>0 0 0 0 0</td>
</tr>
<tr>
<td>11. My teacher tells me that I do good work.</td>
<td>0 0 0 0 0</td>
</tr>
<tr>
<td>12. My teacher explains the lesson clearly.</td>
<td>0 0 0 0 0</td>
</tr>
<tr>
<td>13. My teacher is easy to understand.</td>
<td>0 0 0 0 0</td>
</tr>
<tr>
<td>14. My teacher stays in our classroom.</td>
<td>0 0 0 0 0</td>
</tr>
<tr>
<td>15. My teacher has us work at the right pace.</td>
<td>0 0 0 0 0</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>16.</td>
<td>My teacher tells us what new things we can learn in each lesson.</td>
</tr>
<tr>
<td>17.</td>
<td>My teacher will explain new things in a way that is easy to understand.</td>
</tr>
<tr>
<td>18.</td>
<td>My teacher is available to help me during class time and other times during the school day.</td>
</tr>
<tr>
<td>19.</td>
<td>My teacher knows a lot about the lesson being taught.</td>
</tr>
<tr>
<td>20.</td>
<td>My teacher is well-prepared.</td>
</tr>
</tbody>
</table>
APPENDIX E.

6-8 CAVE CREEK REGULAR STUDENT RATING OF TEACHERS QUESTIONNAIRE
NOTE TO STUDENTS: Please remember that completing this form is voluntary. You may keep this form if you decide not to participate.

Directions: The statements below are designed to find out more about your class and teacher. This is not a test. Do not put your name on this paper. Please answer all the statements. Students are not to ask any questions during the survey.

1 = Never
2 = Not often
3 = Sometimes
4 = Usually
5 = Almost always

1. My teacher makes class work interesting.  
2. My teacher is fair with all.  
3. My teacher maintains discipline in our classroom.  
4. My teacher is well-prepared for our class.  
5. My teacher gives assignments related to the subject we are studying.  
6. We discuss and summarize each lesson just studied.  
7. Our discussions focus on the topic of the lesson.  
8. My teacher likes it when we ask questions.  
9. I have more time to do my work than I need.  
10. My teacher starts lessons explaining what we are going to do and why we are going to do it.  
11. My teacher asks us questions in class to see if we understand what has been taught.  
12. My teacher explains new ideas in a way that is easy to understand.  
13. My teacher looks at our work, as we are doing it, to see if we understand the lesson.
14. My teacher knows more about the subject than other teachers I have had.

15. My teacher has work for me to do if I finish an assignment before the class is over.

16. My teacher often makes materials and worksheets for us to use.

17. My teacher gives tests and quizzes.

18. My teacher returns tests and assignments quickly.

19. My teacher uses a variety of classroom activities and resources.

20. My teacher gives enough time to do our work.
APPENDIX F.

6-8 CAVE CREEK FLOATER STUDENT RATING OF TEACHERS QUESTIONNAIRE
NOTE TO STUDENTS: Please remember that completing this form is voluntary. You may keep this form if you decide not to participate.

Directions: The statements below are designed to find out more about your class and teacher. This is not a test. Do not put your name on this paper. Please answer all the statements. Students are not to ask any questions during the survey.

1 = Never
2 = Not often
3 = Sometimes
4 = Usually
5 = Almost always

1. My teacher makes class work interesting.  O O O O O
2. My teacher is fair with all.  O O O O O
3. My teacher maintains discipline in our classroom.  O O O O O
4. My teacher is well-prepared for our class.  O O O O O
5. My teacher gives assignments related to the subject we are studying.  O O O O O
6. We discuss and summarize each lesson just studied.  O O O O O
7. Our discussions focus on the topic of the lesson.  O O O O O
8. My teacher likes it when we ask questions.  O O O O O
9. My teacher explains the rules for classroom behavior very clearly.  O O O O O
10. My teacher starts lessons explaining what we are going to do and why we are going to do it.  O O O O O
11. My teacher asks us questions in class to see if we understand what has been taught.  O O O O O
12. My teacher explains new ideas in a way that is easy to understand.  O O O O O
13. My teacher looks at our work, as we are doing it, to see if we understand the lesson.  O O O O O
14. We do the same thing everyday in class.  
15. My teacher is easy to understand when talking.  
16. My teacher's tests are fair.  
17. My teacher gives tests and quizzes.  
18. My teacher returns tests and assignments quickly.  
19. My teacher uses a variety of classroom activities and resources.  
20. My teacher expects me to do the best work I can.
APPENDIX G.

9-12 CAVE CREEK STUDENT RATING OF TEACHERS QUESTIONNAIRE
**STUDENT FEEDBACK TO TEACHERS**  
**SENIOR HIGH SCHOOL QUESTIONNAIRE (9-12)**

**NOTE TO STUDENTS:** Please remember that completing this form is voluntary. You may keep this form if you decide not to participate.

**Directions:** The statements below are designed to find out more about your class and teacher. This is not a test. Do not put your name on this paper. Please answer all the statements. Students are not to ask any questions during the survey.

<table>
<thead>
<tr>
<th></th>
<th>1 = Never</th>
<th>2 = Not often</th>
<th>3 = Sometimes</th>
<th>4 = Usually</th>
<th>5 = Almost always</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>My teacher makes class work interesting.</td>
<td>O O O O O</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>My teacher asks questions to see if we understand what has been taught.</td>
<td>O O O O O</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>My teacher gives assignments related to the subject we are studying.</td>
<td>O O O O O</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>We discuss and summarize each lesson we have just studied.</td>
<td>O O O O O</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>My teacher tells us how we can use what we have already learned to learn new things.</td>
<td>O O O O O</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td>My teacher maintains discipline in our classroom.</td>
<td>O O O O O</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td>My teacher returns tests and assignments quickly.</td>
<td>O O O O O</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td>My teacher gives me feedback about my performance.</td>
<td>O O O O O</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td>My teacher knows more about the subject than other teachers I have had.</td>
<td>O O O O O</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td>My homework helps me to learn the subject being taught.</td>
<td>O O O O O</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11.</td>
<td>My teacher makes materials and worksheets for us to use.</td>
<td>O O O O O</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12.</td>
<td>My teacher uses a variety of classroom activities and resources.</td>
<td>O O O O O</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
13. The films or videotapes we watch help us learn about the subject we are studying.

14. My teacher tells the class about library/media materials that will help us learn about the subject we are studying, when appropriate.

15. My teacher is well-organized.

16. My teacher likes it when we ask questions.

17. We work in different groups depending upon the activity in which we are involved.

18. My teacher encourages us to look at problems in new ways and find new ways to solve problems.

19. My teacher is available to help me during class time and other times during the school day.

20. My teacher looks at our work, as we are doing it, to see if we understand the lesson.
APPENDIX H.

ORIGINAL PROCEDURES FOR ADMINISTERING THE QUESTIONNAIRES
INSTRUCTIONS FOR STUDENT FEEDBACK SURVEYS

Student feedback surveys are to be conducted late in each course offered at times set by each building faculty. Following are the instructions for those teachers who will administer the survey.

1. All of your students will be surveyed.

2. For primary grades (K-2), exchanging classes is necessary because all questions must be read to students. Exchanging classes is a recommended procedure for all grade levels. A schedule will be determined in each building.

3. General purpose, machine-scored answer sheets (bubble sheets) will be used for recording answers. These will be supplied by each principal's office prior to the survey. A return envelope will also be provided. This should be sealed after inserting the answer sheets and forwarded through the principal's office for tabulation.

4. Read each direction to every class regardless of age (see special instructions for K-2 students on instrument). Students are not to ask any questions during the survey.

5. Refrain from making any comments other than the specified directions.

6. Insist that no names be written on the form and that "personalized" pen or pencil colors be avoided. Only No. 2 lead pencils should be used. Say that you want "confidential" answers which you will add together to "get the big picture." Make it clear this is voluntary. If students prefer to not participate, they simply do not return the questionnaire.

7. Ask a student to pick up the completed forms (again the reason is to assure anonymity) and place them in the envelope. The student is to seal the envelope.

ADDITIONAL INSTRUCTIONS FOR SPECIAL AREA TEACHERS.

1. Special area teachers include Art, P.E., Band, Music, Special Education, Spanish (K-5), and ESL.

2. Special area teachers are to randomly sample 100 of their students. Use the roll book and pick every third student until a total of 100 is attained.

3. Special education/resource students will have the survey questions read to them as a class.
APPENDIX I.

50-ITEM BUBBLE SHEET
STUDENT RESPONSE RATING FORM
MARKING INSTRUCTIONS

- Do not use ink or ballpoint pens.
- Make heavy black marks that fill the circle completely.
- Erase cleanly any answer you wish to change.
- Make no stray marks on the answer sheet.

EXAMPLES

1. (1) (2) (3) (4) (5)
   2. (1) (2) (3) (4) (5)
   3. (1) (2) (3) (4) (5)
   4. (1) (2) (3) (4) (5)

A B C D E

Name

A B C D E

A B C D E

A B C D E

A B C D E

A B C D E

A B C D E

A B C D E

A B C D E

A B C D E

A B C D E

A B C D E

A B C D E

A B C D E

A B C D E

A B C D E

A B C D E

A B C D E

A B C D E

A B C D E

A B C D E

A B C D E

A B C D E

A B C D E

A B C D E
APPENDIX J.

REVISED INSTRUCTION SHEETS
FOR ADMINISTERING QUESTIONNAIRES
In order that in the future student feedback data can be processed most effectively, please include the following information regarding teachers and/or students participating in Cave Creek's career ladder student feedback process.

A. TO BE INCLUDED ON BUBBLE SHEET

1. Teacher Information:
   Code (6 digits)--The first five digits of the teacher code should be listed under "special codes" and may be the teacher's social security number or any other self-selected number. (For teachers participating for the second year, use the same code as was used previously.) The sixth digit of the code must be an identification number for the class (period of the day at the middle/high school and an appropriate identification number at the elementary level.) If there is no period number or other appropriate identification number, use 9 for the sixth digit.

   When sending the evaluation data, please include a list of class code numbers. (See Category B, below)

   2. Student Information
      Sex of Student--To be listed in the first column under the area designated for identification number. (Code: 0=Male, 1=Female)
      The remaining 9 columns should be left blank.

   3. Date
      Do not fill in any information pertaining to month, day, or year.

B. ON SEPARATE SHEET, LISTED BY TEACHER CODE NUMBERS

1. Sex of Teacher

2. Years of Teaching Experience

3. Grade Level or Subject Taught--for each grade or subject for which the teacher is being evaluated

4. Designation: Required or Elective Class (middle school and high school)

*If possible, we would like to have the above information on each participant in the 1989-90 program.
To facilitate processing of the data, bubble sheets should be packaged in the following manner:

1. Group all of each teacher's classes together—for example, put Teacher A's period 1 together, then period 2 etc., making certain that the sixth digit class identifier code is correct. Put all of Teacher A's bubble sheets inside of a large manila envelope. (Avoid wrapping bubble sheets with rubber bands.) On the outside of the envelope, specify:
   A) Name of Teacher
   B) Name of School
   C) Teacher Code
   D) Grade Level Designation (See Below)
   E) Designation: Regular or Floater

2. Group manila envelopes into 7 bundles according to the following designations. To avoid having to mail 7 boxes, we suggest that you might wrap each bundle (for example, K-2 or 3-5) in butcher paper or a similar packaging material. Pack bundles in boxes and mail.
   A) K-2 Regular Teachers
   B) 3-5 Regular Teachers
   C) 6-8 Regular Teachers
   D) 9-12 Teachers (All)
   E) K-2 Floater Teachers
   F) 3-5 Floater Teachers
   G) 6-8 Floater Teachers
APPENDIX K.

SPSS-X COMPUTER PROGRAMS
FOR PROCESSING AND ANALYZING DATA
61. NNTQ1  '9 12 DIFFERENT GROUPS DEPEND ACTIVITY'
62. NNTQ1 8  '9 12 LOOK PROBLEMS/NEW WAYS TO SOLVE'
63. NNTQ19 '9 12 AVAILABLE CLASS/OTHER TIMES'
64. NNTQ20 '9 12 LOOKS AT WORK/SEE IF UNDERSTAND'

66. VALUE LABELS
67. NNTQ1 TO NNTQ20
68. 0 'NEVER'
69. 1 'NOT OFTEN'
70. 2 'SOMETIMES'
71. 3 'USUALLY'
72. 4 'ALMOST ALWAYS'
73. CLPERN12
74. 1 'PERIOD 1'
75. 2 'PERIOD 2'
76. 3 'PERIOD 3'
77. 4 'PERIOD 4'
78. 5 'PERIOD 5'
79. 6 'PERIOD 6'
80. 7 'PERIOD 7'
81. 8 'PERIOD 8'
82. 9 'NO PERIOD LISTED'

83. COMPUTE TOTSCORE = SUM (NNTQ1 TO NNTQ20)
84. DESCRIPTIVES VARIABLES = TOTSCORE /
85. STATISTICS = DEFAULT
86. SORT CASES BY TCHN12
87. SPLIT FILE BY TCHN12
88. DESCRIPTIVES VARIABLES = TOTSCORE /
89. STATISTICS = DEFAULT
90. FINISH

92. SELECT IF (TCHN12 EQ 40923) /
93. FREQUENCIES VARIABLES = NNTQ1 TO NNTQ20 /
94. STATISTICS = DEFAULT
95. TEMPORARY
96. SELECT IF (TCHN12 EQ 66587) /
97. FREQUENCIES VARIABLES = NNTQ1 TO NNTQ20 /
98. STATISTICS = DEFAULT
99. TEMPORARY
100. SELECT IF (TCHN12 EQ 25281) /
101. FREQUENCIES VARIABLES = NNTQ1 TO NNTQ20 /
102. STATISTICS = DEFAULT
103. TEMPORARY
104. SELECT IF (TCHN12 EQ 88069) /
105. FREQUENCIES VARIABLES = NNTQ1 TO NNTQ20 /
106. STATISTICS = DEFAULT
107. TEMPORARY
108. SELECT IF (TCHN12 EQ 81349) /
109. FREQUENCIES VARIABLES = NNTQ1 TO NNTQ20 /
110. STATISTICS = DEFAULT
111. TEMPORARY
112. SELECT IF (TCHN12 EQ 60122) /
113. FREQUENCIES VARIABLES = NNTQ1 TO NNTQ20 /
114. STATISTICS = DEFAULT
115. TEMPORARY
116. SELECT IF (TCHN12 EQ 06805) /
117. FREQUENCIES VARIABLES = NNTQ1 TO NNTQ20 /
118. STATISTICS = DEFAULT
119. TEMPORARY
120. SELECT IF (TCHN12 EQ 26052)
I //Cave job
2 /*JOBPARM LINES=80
3 //SI EXEC SPSSX
4 //DATA RAW DD DSN=E152ABM.DAT44RAW.UNIT=DISK.DISP=SHR
5.
6. TITLE CAVE CREEK 9-12 PROGRAM
7. DATA LIST FILE=DATA RAW RECORDS=2
8. /1 CARDIN12 1
9. SEQU12 2-5
10. TCHN12 3-40
11. CCLPERN12 41
12. /2 CARDIN12 1
13. SEQU12 2-8
14. NT01 6
15. NT02 7
16. NT03 8
17. NT04 9
18. NT05 10
19. NT06 11
20. NT07 12
21. NT08 13
22. NT09 14
23. NT10 15
24. NT11 16
25. NT12 17
26. NT13 18
27. NT14 19
28. NT15 20
29. NT16 21
30. NT17 22
31. NT18 23
32. NT19 24
33. NT20 25
34. NT21 26
35. RECODE NT01 TO NT20 (1 = 0) (2 = 1) (3 = 2) (4 = 3)
36. (5 = 4) INTO NNT01 TO NNT20
37.
38. VARIABLE LABELS
39. CARDIN12 'REGULAR 9-12 CARD 1'
40. SEQU12 'SEQUENCE OF ENTRY CD1'
41. TCHN12 'TEACHER REGULAR 9-12'
42. CCLPERN12 'CLASS PERIOD R 9-12'
43. SEQU212 'SEQUENCE OF ENTRY CD2'
44. NNT01 '9-12 MAKES CLASS WORK INTERESTING'
45. NNT02 '9-12 ASK QUESTIONS UNDERSTAND TAUGHT'
46. NNT03 '9-12 ASSIGNMENTS RELATED TO SUBJECT'
47. NNT04 '9-12 DISCUSS/SUMMARIZE EACH LESSON'
48. NNT05 '9-12 TELLS WHAT LEARNED TO LEARN NEW'
49. NNT06 '9-12 TEACHER MAINTAINS DISCIPLINE'
50. NNT07 '9-12 RETURNS TEST/ASSIGNMENTS QUICKLY'
51. NNT08 '9-12 GIVES FEEDBACK ABOUT PERFORMANCE'
52. NNT09 '9-12 KNOWS MORE THAN OTHER TEACHERS'
53. NNT10 '9-12 HOMEWORK HELPS ME LEARN'
54. NNT11 '9-12 MAKES MATERIALS/WORKSHEETS'
55. NNT12 '9-12 USES VARIETY ACTIVITIES/RESOURCES'
56. NNT13 '9-12 FILMS/VIDEOTAPE HELP ME LEARN'
57. NNT14 '9-12 TELLS LIBRARY/MEDIA RESOURCES'
58. NNT15 '9-12 TEACHER WELL ORGANIZED'
59. NNT16 '9-12 TEACHER LIKES WHEN ASK QUESTIONS'
VALUE LABELS
NNT01 TO NNT020
0 'NEVER'
1 'NOT OFTEN'
2 'SOMETIMES'
3 'USUALLY'
4 'ALMOST ALWAYS'/
5 'PERIOD 1'
6 'PERIOD 2'
7 'PERIOD 3'
8 'PERIOD 4'
9 'PERIOD 5'
10 'PERIOD 6'
11 'PERIOD 7'
12 'PERIOD 8'
13 'PERIOD 9'
14 'NO PERIOD LISTED'/

COMPUTE TOTSCORE = SUM (NNTQ1 TO NNTQ20)
/STATISTICS = DEFAULT
SORT CASES BY TCHN12
SPLIT FILE BY TCHN12
DESCRIPTIVES VARIABLES = TOTSCORE
/STATISTICS = DEFAUlT
FINISH

//
TEMPORARY
SELECT IF (TCHN12 EQ 40923)
FREQUENCIES VARIABLES = NNTQ1 TO NNTQ20/
STATISTICS = DEFAULT
TEMPORARY
SELECT IF (TCHN12 EQ 66587)
FREQUENCIES VARIABLES = NNTQ1 TO NNTQ20 /
STATISTICS = DEFAULT
TEMPORARY
SELECT IF (TCHN12 EQ 25281)
FREQUENCIES VARIABLES = NNTQ1 TO NNTQ20 /
STATISTICS = DEFAULT
TEMPORARY
SELECT IF (TCHN12 EQ 88069)
FREQUENCIES VARIABLES = NNTQ1 TO NNTQ20 /
STATISTICS = DEFAULT
TEMPORARY
SELECT IF (TCHN12 EQ 81348)
FREQUENCIES VARIABLES = NNTQ1 TO NNTQ20 /
STATISTICS = DEFAULT
TEMPORARY
SELECT IF (TCHN12 EQ 60122)
FREQUENCIES VARIABLES = NNTQ1 TO NNTQ20 /
STATISTICS = DEFAULT
TEMPORARY
SELECT IF (TCHN12 EQ 60122)
FREQUENCIES VARIABLES = NNTQ1 TO NNTQ20 /
STATISTICS = DEFAULT
TEMPORARY
SELECT IF (TCHN12 EQ 60122)
FREQUENCIES VARIABLES = NNTQ1 TO NNTQ20 /
STATISTICS = DEFAULT
TEMPORARY
SELECT IF (TCHN12 EQ 26052)
//CAVE JOB
//JOBPARM LINES=80
//SI EXEC SPSSX
//DATA_READ DD DSN=E15WHAD4.DATANRAW.UNIT=DISK,DISP=SHR

TITLE CAVE CREEK 9-12 PROGRAM

DATA LIST FILE=DATA4RAW RECORDS=2

/1 CARDIN12 1
/2 SEQIN12 2-5
/3 TCHN12 36-40
/4 CLPERN12 41

/1 CARD2N12 1
/2 SEQ2N12 2-8
/3 NTQ1 9
/4 NTQ2 6
/5 NTQ3 7
/6 NTQ4 8
/7 NTQ5 9
/8 NTQ6 10
/9 NTQ7 11
/10 NTQ8 12
/11 NTQ9 13
/12 NTQ10 14
/13 NTQ11 16
/14 NTQ12 17
/15 NTQ13 18
/16 NTQ14 19
/17 NTQ15 20
/18 NTQ16 21
/19 NTQ17 22
/20 NTQ18 23
/21 NTQ19 24
/22 NTQ20 25

RECODE NTQ1 TO NTQ20 (1 = 0) (2 = 1) (3 = 2) (4 = 3)

VARIABLE LABELS
CARDIN12 'REGULAR 9-12 CARD 1'
SEQIN12 'SEQUENCE OF ENTRY CD1'
TCHN12 'TEACHER REGULAR 9-12'
CLPERN12 'CLASS PERIOD R 9-12'
SEQ2N12 'SEQUENCE OF ENTRY CD2'
NTQ1 '9-12 MAKES CLASS WORK INTERESTING'
NTQ2 '9-12 ASK QUESTIONS UNDERSTAND TAUGHT'
NTQ3 '9-12 ASSIGNMENTS RELATED TO SUBJECT'
NTQ4 '9-12 DISCUSS/SUMMARIZE EACH LESSON'
NTQ5 '9-12 TELLS WHAT LEARNED TO LEARN NEW'
NTQ6 '9-12 TEACHER MAINTAINS DISCIPLINE'
NTQ7 '9-12 RETURNS TEST/ASSIGNMENTS QUICKLY'
NTQ8 '9-12 GIVES FEEDBACK ABOUT PERFORMANCE'
NTQ9 '9-12 KNOWS MORE THAN OTHER TEACHERS'
NTQ10 '9-12 HOMEWORK HELPS ME LEARN'
NTQ11 '9-12 MAKES MATERIALS/WORKSHEETS'
NTQ12 '9-12 USE VARIETY ACTIVITIES/RESOURCES'
NTQ13 '9-12 FILMS/VIDEO/PE'S HELP US LEARN'
NTQ14 '9-12 TELLS LIBRARY/MEDIA MATERIALS'
NTQ15 '9-12 TEACHER WOULD ORGANIZE'
NTQ16 '9-12 TEACHER LIKES WHEN ASK QUESTIONS'
VALUE LABELS
NFEQ1 TO NFEQ20
0 'NEVER'
1 'NOT OFTEN'
2 'SOMETIMES'
3 'USUALLY'
4 'ALMOST ALWAYS'/

CLPERF8
1 'PERIOD 1'
2 'PERIOD 2'
3 'PERIOD 3'
4 'PERIOD 4'
5 'PERIOD 5'
6 'PERIOD 6'
7 'PERIOD 7'
8 'PERIOD 8'
9 'NO PERIOD LISTED'/

COMPUTE TOTSCORE = SUM(NFEQ1 TO NFEQ20)

SORT CASES BY TCHF8
SELECT IF (TCHF8 EQ 63380)
FREQUENCIES VARIABLES = NFEQ1 TO NFEQ20/
STATISTICS = DEFAULT

SELECT IF (TCHF8 EQ 05671)
FREQUENCIES VARIABLES = NFEQ1 TO NFEQ20/
STATISTICS = DEFAULT

SELECT IF (TCHF8 EQ 02544)
FREQUENCIES VARIABLES = NFEQ1 TO NFEQ20/
STATISTICS = DEFAULT

SELECT IF (TCHF8 EQ 05641)
FREQUENCIES VARIABLES = NFEQ1 TO NFEQ20/
STATISTICS = DEFAULT

SELECT IF (TCHF8 EQ 05641)
FREQUENCIES VARIABLES = NFEQ1 TO NFEQ20/
STATISTICS = DEFAULT

SELECT IF (TCHF8 EQ 05671)
FREQUENCIES VARIABLES = NFEQ1 TO NFEQ20/
STATISTICS = DEFAULT

SELECT IF (TCHF8 EQ 05671)
FREQUENCIES VARIABLES = NFEQ1 TO NFEQ20/
STATISTICS = DEFAULT

SELECT IF (TCHF8 EQ 05641)
FREQUENCIES VARIABLES = NFEQ1 TO NFEQ20/
STATISTICS = DEFAULT

SELECT IF (TCHF8 EQ 05671)
FREQUENCIES VARIABLES = NFEQ1 TO NFEQ20/
STATISTICS = DEFAULT

SELECT IF (TCHF8 EQ 05671)
FREQUENCIES VARIABLES = NFEQ1 TO NFEQ20/
STATISTICS = DEFAULT

SELECT IF (TCHF8 EQ 05641)
FREQUENCIES VARIABLES = NFEQ1 TO NFEQ20/
STATISTICS = DEFAULT

SELECT IF (TCHF8 EQ 05671)
FREQUENCIES VARIABLES = NFEQ1 TO NFEQ20/
STATISTICS = DEFAULT
//CAVE JOB
//JOBPARM LINES=80
//SI EXEC SPSSX
//DATATRAV DD DSN=E15SWhA.DATATRAV,UNIT=DISK,DISP=SHR

TITLE CAVE CREEK FLOATER 6-8 PROGRAM

DATA LIST FILE=DATATRAV RECORDS=2
/1 CARD1F8 1
/2 CARD2F8 1
SEQ1F8 2-5
TCHF8 36-40
CLPERF8 41
SEQ2F8 2-5
FEQ1 6
FEQ2 7
FEQ3 8
FEQ4 9
FEQ5 10
FEQ6 11
FEQ7 12
FEQ8 13
FEQ9 14
FEQ10 15
FEQ11 16
FEQ12 17
FEQ13 18
FEQ14 19
FEQ15 20
FEQ16 21
FEQ17 22
FEQ18 23
FEQ19 24
FEQ20 25

RECODE FEQ1 TO FEQ20
(1 = 0) (2 = 1) (3 = 2) (4 = 3)
(5 = 4) INTO NFEQ1 TO NFEQ20

VARIABLE LABELS
CAROIF8   'FLOATER 6-8 CARD 1'
SEQ1F8   'SEQUENCE OF ENTRY CD1'
TCHF8   'TEACHER FLOATER 6-8'
CLPERF8   'CLASS PERIOD 6-8'
CARD2F8   'FLOATER 6-8 CARD 2'
SEQ2F8   'SEQUENCE OF ENTRY CD2'
NFEQ1   '6-8 TEACHER MAKES INTERESTING'
NFEQ2   '6-8 TEACHER FAIR WITH ALL'
NFEQ3   '6-8 TEACHER MAINTAINS DISCIPLINE'
NFEQ4   '6-8 TEACHER WELL-PREPARED'
NFEQ5   '6-8 GIVES ASSIGNMENTS RELATED SUBJ'
NFEQ6   '6-8 DISCUSS/SUMMARIZE LESSON'
NFEQ7   '6-8 TALKS ABOUT TOPIC OF LESSON'
NFEQ8   '6-8 LIKES WHEN WE ASK QUESTIONS'
NFEQ9   '6-8 EXPLAINS RULES FOR BEHAVIOR'
NFEQ10   '6-8 EXPLAINS LESSON AND WHY DO IT'
NFEQ11   '6-8 ASK QUESTIONS/UNDERSTAND TAUGHT'
NFEQ12   '6-8 TALKS ABOUT IDEAS EASY UNDERSTAND'
NFEQ13   '6-8 LOOKS AT WORK SEE WE UNDERSTAND'
NFEQ14   '6-8 DO SAME THING IN CLASS EVERY DAY'

DATA LIST FILE=DATATRAV RECORDS=2
/1 CARD1F8 1
/2 CARD2F8 1
SEQ1F8 2-5
TCHF8 36-40
CLPERF8 41
SEQ2F8 2-5
FEQ1 6
FEQ2 7
FEQ3 8
FEQ4 9
FEQ5 10
FEQ6 11
FEQ7 12
FEQ8 13
FEQ9 14
FEQ10 15
FEQ11 16
FEQ12 17
FEQ13 18
FEQ14 19
FEQ15 20
FEQ16 21
FEQ17 22
FEQ18 23
FEQ19 24
FEQ20 25

RECODE FEQ1 TO FEQ20
(1 = 0) (2 = 1) (3 = 2) (4 = 3)
(5 = 4) INTO NFEQ1 TO NFEQ20

VARIABLE LABELS
CAROIF8   'FLOATER 6-8 CARD 1'
SEQ1F8   'SEQUENCE OF ENTRY CD1'
TCHF8   'TEACHER FLOATER 6-8'
CLPERF8   'CLASS PERIOD 6-8'
CARD2F8   'FLOATER 6-8 CARD 2'
SEQ2F8   'SEQUENCE OF ENTRY CD2'
NFEQ1   '6-8 TEACHER MAKES INTERESTING'
NFEQ2   '6-8 TEACHER FAIR WITH ALL'
NFEQ3   '6-8 TEACHER MAINTAINS DISCIPLINE'
NFEQ4   '6-8 TEACHER WELL-PREPARED'
NFEQ5   '6-8 GIVES ASSIGNMENTS RELATED SUBJ'
NFEQ6   '6-8 DISCUSS/SUMMARIZE LESSON'
NFEQ7   '6-8 TALKS ABOUT TOPIC OF LESSON'
NFEQ8   '6-8 LIKES WHEN WE ASK QUESTIONS'
NFEQ9   '6-8 EXPLAINS RULES FOR BEHAVIOR'
NFEQ10   '6-8 EXPLAINS LESSON AND WHY DO IT'
NFEQ11   '6-8 ASK QUESTIONS/UNDERSTAND TAUGHT'
NFEQ12   '6-8 TALKS ABOUT IDEAS EASY UNDERSTAND'
NFEQ13   '6-8 LOOKS AT WORK SEE WE UNDERSTAND'
NFEQ14   '6-8 DO SAME THING IN CLASS EVERY DAY'
//CAVE J00
//JOBPARM LINES=80
//EXEC SPSS
//DATA TRAM DD DSN=E1SLMD.DATATRAM.UNIT=DISK,DISP=SHR

TITLE CAVE CREEK FLOATER 6-8 PROGRAM

DATA LIST FILE=DATATRAM RECORDS=2
/CARD1F8 1
SEQ1F8 2-5
TCHF8 36-40
CLPERF8 41
/CARD2F8 1
SEQ2F8 2-9
FEQ1 6
FEQ2 7
FEQ3 8
FEQ4 9
FEQ5 10
FEQ6 11
FEQ7 12
FEQ8 13
FEQ9 14
FEQ10 15
FEQ11 16
FEQ12 17
FEQ13 18
FEQ14 19
FEQ15 20
FEQ16 21
FEQ17 22
FEQ18 23
FEQ19 24
FEQ20 25

RECODE FEQ1 TO FEQ20 (1 = 0) (2 = 1) (3 = 2) (4 = 3)
(5 = 4) INTO NFEQ1 TO NFEQ20

VARIABLE LABELS
CARD1F8 'FLOATER 6-8 CARD 1'
SEQ1F8 'SEQUENCE OF ENTRY CD1'
TCHF8 'TEACHER FLOATER 6-8'
CLPERF8 'CLASS PERIOD F 6-8'
CARD2F8 'FLOATER 6-8 CARD 2'
SEQ2F8 'SEQUENCE OF ENTRY CD2'
NFEQ1 'F-8 TEACHER MAKES INTERESTING'
NFEQ2 'F-8 TEACHER FAIR WITH ALL'
NFEQ3 'F-8 TEACHER MAINTAINS DISCIPLINE'
NFEQ4 'F-8 TEACHER WELL-PREPARED'
NFEQ5 'F-8 GIVES ASSIGNMENTS RELATED SUBJ'
NFEQ6 'F-8 DISCUSS/SUMMARIZE LESSON'
NFEQ7 'F-8 DISCUSSIONS ON TOPIC OF LESSON'
NFEQ8 'F-8 LIKES WHEN WE ASK QUESTIONS'
NFEQ9 'F-8 EXPLAINS RULES FOR BEHAVIOR'
NFEQ10 'F-8 EXPLAINS LESSON AND WHY DO IT'
NFEQ11 'F-8 ASK QUESTIONS/UNDERSTAND TAUGHT'
NFEQ12 'F-8 EXPLAINS IDEAS EASY UNDERSTAND'
NFEQ13 'F-8 LOOKS AT WORK SEE WE UNDERSTAND'
NFEQ14 'F-8 DO SAME THING IN CLASS EVERY DAY'
NFEQ15 'F-8 TEACHER EASY TO UNDERSTAND'

6.
6.8 KNOWS MORE THAN OTHER TEACHERS'
6.8 WORKS TO DO IF FINISH CLASS OVER
6.8 MAKES MATERIALS WORKSHEETS USE
6.8 GIVES TESTS AND QUIZES
6.8 RETURNS TESTS/ASSIGNMENTS QUICKLY
6.8 VARIETY OF ACTIVITIES/RESOURCES

VALUE LABELS
NSEQ10 TO NSEQ20

NEVER
"NOT OFTEN"
"SOMETIMES"
"USUALLY"
"ALMOST ALWAYS"
"PERIOD 1"
"PERIOD 2"
"PERIOD 3"
"PERIOD 4"
"PERIOD 5"
"PERIOD 6"
"PERIOD 7"
"PERIOD 8"
"NO PERIOD LISTED"

COMPUTE TDTSCORE = SUM (NSEQ10 TO NSEQ20)

DESCRIPTIVES VARIABLES = TDTSCORE
/STATISTICS = DEFAULT
SORT CASES BY TCHSS
SPLIT FILE BY TCHSS
DESCRIPTIVES VARIABLES = TOTSCORE
/STATISTICS = DEFAULT
FINISH

SELECT IF (TCHSS EQ 411821) .
FREQUENCIES VARIABLES = NSEQ10 TO NSEQ20 /
STATISTICS = DEFAULT
TEMPORARY
SELECT IF (TCHSS EQ 476861) .
FREQUENCIES VARIABLES = NSEQ10 TO NSEQ20 /
STATISTICS = DEFAULT
TEMPORARY
SELECT IF (TCHSS EQ 231361) .
FREQUENCIES VARIABLES = NSEQ10 TO NSEQ20 /
STATISTICS = DEFAULT
TEMPORARY
SELECT IF (TCHSS EQ 879982) .
FREQUENCIES VARIABLES = NSEQ10 TO NSEQ20 /
STATISTICS = DEFAULT
TEMPORARY
DATA LIST FILE=DATA\RAW \TITLE CAVE CREEK 6-8 REGULAR PROGRAM
/1 CARD\$B 1
  SEQ\$B 2-5
  TCH\$B 36-40
  CL\$B 41
/2 CARD\$B 1
  SEQ\$B 2-5
  SEQ  6
  SEQ  7
  SEQ  8
  SEQ  9
  SEQ 10
  SEQ 11
  SEQ 12
  SEQ 13
  SEQ 14
  SEQ 15
  SEQ 16
  SEQ 17
  SEQ 18
  SEQ 19
  SEQ 20
  SEQ 21
  SEQ 22
  SEQ 23
  SEQ 24
  SEQ 25
  SEQ 26
  SEQ 27
  SEQ 28
  SEQ 29
  SEQ 30
  SEQ 31
  SEQ 32
  SEQ 33
  SEQ 34
  SEQ 35
  SEQ 36
  RECODE SEQ1 TO SEQ20 (1 = 0) (2 = 1) (3 = 2) (4 = 3)
  (5 = 4) INTO NSEQ1 TO NSEQ20
  /VARIABLE LABELS
  CARD\$ '6-8 REGULAR CARD 1'
  SEQ\$ 'SEQUENCE OF ENTRY CO1'
  TCH\$ 'TEACHER REGULAR 6-8'
  CL\$ 'CLASS PERIOD R 6-8'
  CARD\$B '6-8 REGULAR CARD 2'
  SEQ\$B 'SEQUENCE OF ENTRY CO2'
  NSEQ1 '6-8 MAKES CLASS WORK INTERESTING'
  NSEQ2 '6-8 TEACHER FAIR WITH ALL'
  NSEQ3 '6-8 TEACHER MAINTAINS DISCIPLINE'
  NSEQ4 '6-8 TEACHER WELL-PREPARED'
  NSEQ5 '6-8 GIVES ASSIGNMENTS RELATED SUBJ'
  NSEQ6 '6-8 DISCUSS/SUMMARIZE LESSON'
  NSEQ7 '6-8 DISCUSSIONS ON TOPIC OF LESSON'
  NSEQ8 '6-8 LIKES WHEN WE ASK QUESTIONS'
  NSEQ9 '6-8 MORE TIME TO WORK THAN NEED'
  NSEQ10 '6-8 EXPLAINS LESSON AND WHY DO IT'
  NSEQ11 '6-8 ASK QUESTIONS/UNDERSTAND TAUGHT'
  NSEQ12 '6-8 EXPLAINS IDEAS EASY UNDERSTAND'
  NSEQ13 '6-8 LOOKS AT WORK SEE WE UNDERSTAND'
1. NSEQ14: '6-8 KNOWS MORE THAN OTHER TEACHERS'
2. NSEQ15: '6-8 WORK TO DO IF FINISH CLASS OVER'
3. NSEQ16: '6-8 MAKES MATERIALS/WORKSHEETS USE'
4. NSEQ17: '6-8 GIVES TESTS AND QUIZZES'
5. NSEQ18: '6-8 RETURNS TESTS/ASSIGNMENTS QUICK'
6. NSEQ19: '6-8 VARIETY OF ACTIVITIES/RESOURCES'
7. NSEQ20: '6-8 ENOUGH TIME TO DO OUR WORK'

VALUE LABELS
NSEQ1 TO NSEQ20
  0  'NEVER'
  1  'NOT OFTEN'
  2  'SOMETIMES'
  3  'USUALLY'
  4  'ALMOST ALWAYS'

CLPERSON
  1  'PERIOD 1'
  2  'PERIOD 2'
  3  'PERIOD 3'
  4  'PERIOD 4'
  5  'PERIOD 5'
  6  'PERIOD 6'
  7  'PERIOD 7'
  8  'PERIOD 8'
  9  'NO PERIOD LISTED'

TEMPORARY
SELECT IF (TCHSB EQ 41182)
FREQUENCIES VARIABLES = NSEQ1 TO NSEQ20/
STATISTICS = DEFAULT

TEMPORARY
SELECT IF (TCHSB EQ 47686)
FREQUENCIES VARIABLES = NSEQ1 TO NSEQ20/
STATISTICS = DEFAULT

TEMPORARY
SELECT IF (TCHSB EQ 48658)
FREQUENCIES VARIABLES = NSEQ1 TO NSEQ20/
STATISTICS = DEFAULT

TEMPORARY
SELECT IF (TCHSB EQ 56012)
FREQUENCIES VARIABLES = NSEQ1 TO NSEQ20/
STATISTICS = DEFAULT

TEMPORARY
SELECT IF (TCHSB EQ 23136)
FREQUENCIES VARIABLES = NSEQ1 TO NSEQ20/
STATISTICS = DEFAULT

TEMPORARY
SELECT IF (TCHSB EQ 87998)
FREQUENCIES VARIABLES = NSEQ1 TO NSEQ20/
STATISTICS = DEFAULT

TEMPORARY
SELECT IF (TCHSB EQ 86917)
FREQUENCIES VARIABLES = NSEQ1 TO NSEQ20/
STATISTICS = DEFAULT

TEMPORARY
SELECT IF (TCHSB EQ 86917)
FREQUENCIES VARIABLES = NSEQ1 TO NSEQ20/
STATISTICS = DEFAULT

TEMPORARY
SELECT IF (TCHSB EQ 40119)
FREQUENCIES VARIABLES = NSEQ1 TO NSEQ20/
STATISTICS = DEFAULT

TEMPORARY
SELECT IF (TCHSB EQ 06858)
FREQUENCIES VARIABLES = NSEQ1 TO NSEQ20/
STATISTICS = DEFAULT
DATA LIST: FILE='DATABRAW' RECORD=2.
/1 CARD158
SEQ158 2-5
TCHS8 36-40
CLPERS8 41
/
/2 CARD258
SEQ258 2-5
SEQ1 6
SEQ2 7
SEQ3 8
SEQ4 9
SEQ5 10
SEQ6 11
SEQ7 12
SEQ8 13
SEQ9 14
SEQ10 15
SEQ11 16
SEQ12 17
SEQ13 18
SEQ14 19
SEQ15 20
SEQ16 21
SEQ17 22
SEQ18 23
SEQ19 24
SEQ20 25
RECODE SEQ1 TO SEQ20 (1 = 0) (2 = 1) (3 = 2) (4 = 3) (5 = 4)
INTO NSEQ1 TO NSEQ20.

VARIABLE LABELS:
CARD158 '6-8 REGULAR CARD 1'
SEQ158 'SEQUENCE OF ENTRY CO1'
TCHS8 'TEACHER REGULAR 6-8'
CLPERS8 'CLASS PERIOD R 6-8'
CARD258 '6-8 REGULAR CARD 2'
SEQ258 'SEQUENCE OF ENTRY CO2'
NSEQ1 '6-8 MAKES CLASS WORK INTERESTING'
NSEQ2 '6-8 TEACHER FAIR WITH ALL'
NSEQ3 '6-8 TEACHER MAINTAINS DISCIPLINE'
NSEQ4 '6-8 TEACHER WELL-PREPARED'
NSEQ5 '6-8 GIVES ASSIGNMENTS RELATED SUBJ'
NSEQ6 '6-8 DISCUSS/SUMMARIZE LESSON'
NSEQ7 '6-8 DISCUSSIONS ON TOPIC OF LESSON'
NSEQ8 '6-8 LIKES WHEN WE ASK QUESTIONS'
NSEQ9 '6-8 MORE TIME TO WORK THAN NEED'
NSEQ10 '6-8 EXPLAINS LESSON AND WHY DO IT'
NSEQ11 '6-8 ASK QUESTIONS/UNDERSTAND TAUGHT'
NSEQ12 '6-8 EXPLAINS IDEAS EASY UNDERSTAND'
NSEQ13 '6-8 TEACHER BAK WORKS WE UNDERSTAND'
NSEQ14 '6-8 TEACHER WORKS LESSON TO UNDERSTAND'
NSEQ15 '6-8 TEACHER WORKS LESSON TO UNDERSTAND'
NSEQ16 '6-8 TEACHER WORKS LESSON TO UNDERSTAND'
NSEQ17 '6-8 TEACHER WORKS LESSON TO UNDERSTAND'
NSEQ18 '6-8 TEACHER WORKS LESSON TO UNDERSTAND'
NSEQ19 '6-8 TEACHER WORKS LESSON TO UNDERSTAND'
NSEQ20 '6-8 TEACHER WORKS LESSON TO UNDERSTAND'
```
VALUE LABELS
NFFQ1 TO NFFQ20
0 'NEVER'
1 'NOT OFTEN'
2 'SOMETIMES'
3 'USUALLY'
4 'ALMOST ALWAYS'

CLPERFS
1 'PERIOD 1'
2 'PERIOD 2'
3 'PERIOD 3'
4 'PERIOD 4'
5 'PERIOD 5'
6 'PERIOD 6'
7 'PERIOD 7'
8 'PERIOD 8'
9 'NO PERIOD LISTED'

COMPUTE TOTSCORE = SUM (NFFQ1 TO NFFQ20)
/STATISTICS = DEFAULT
SORT CASES BY TCHFS
SPLIT FILE BY TCHFS
DESCRIPTIVES VARIABLES = TOTSCORE
/STATISTICS = DEFAULT
FINISH
TEMPORARY
SELECT IF (TCHFS EQ 07481)
FREQUENCIES VARIABLES = NFFQ1 TO NFFQ20/
STATISTICS = DEFAULT
TEMPORARY
SELECT IF (TCHFS EQ 22106)
FREQUENCIES VARIABLES = NFFQ1 TO NFFQ20/
STATISTICS = DEFAULT
TEMPORARY
SELECT IF (TCHFS EQ 60668)
FREQUENCIES VARIABLES = NFFQ1 TO NFFQ20/
STATISTICS = DEFAULT
TEMPORARY
SELECT IF (TCHFS EQ 76111)
FREQUENCIES VARIABLES = NFFQ1 TO NFFQ20/
STATISTICS = DEFAULT
TEMPORARY
SELECT IF (TCHFS EQ 84801)
FREQUENCIES VARIABLES = NFFQ1 TO NFFQ20/
STATISTICS = DEFAULT
TEMPORARY
SELECT IF (TCHFS EQ 22106)
FREQUENCIES VARIABLES = NFFQ1 TO NFFQ20/
STATISTICS = DEFAULT
TEMPORARY
SELECT IF (TCHFS EQ 07481)
FREQUENCIES VARIABLES = NFFQ1 TO NFFQ20/
STATISTICS = DEFAULT
TEMPORARY
SELECT IF (TCHFS EQ 22106)
FREQUENCIES VARIABLES = NFFQ1 TO NFFQ20/
STATISTICS = DEFAULT
```

DATA LIST FILE=DAFRAW RECORDS=2
/1 CARDIFS 1
SEQIFS 2-5
TCHIFS 36-40
CLPERFS 41
/2 CARD2IFS 1
SEQ2IFS 2-5
FFQ1 6
FFQ2 7
FFQ3 8
FFQ4 9
FFQ5 10
FFQ6 11
FFQ7 12
FFQ8 13
FFQ9 14
FFQ10 15
FFQ11 16
FFQ12 17
FFQ13 18
FFQ14 19
FFQ15 20
FFQ16 21
FFQ17 22
FFQ18 23
FFQ19 24
FFQ20 25

RECODE FFQ1 TO FFQ20 (1 = 0) (2 = 1) (3 = 2) (4 = 3) (5 = 4) INTO NFFQ1 TO NFFQ20

VARIABLE LABELS
CARDIFS 'FLOATER 3-5 CARD 1'
SEQIFS 'SEQUENCE OF ENTRY CARD1'
TCHIFS 'TEACHER FLOATER 3-5'
CLPERFS 'CLASS PERIOD F 3-5'
CARD2IFS 'FLOATER 3-5 CARD 2'
SEQ2IFS 'SEQUENCE OF ENTRY CARD2'
NFFQ1 'F-5 MAKES WORK INTERESTING'
NFFQ2 'F-5 SCHOOLDAY INTERESTING'
NFFQ3 'F-5 TEACHER USUALLY PREPARED'
NFFQ4 'F-5 SCHOOL DAY INTERESTING'
NFFQ5 'F-5 MAKES FEEL GOOD / GOOD WORK'
NFFQ6 'F-5 MAKES FEEL GOOD / GOOD WORK'
NFFQ8 'F-5 CAN GET HELP FROM TEACHER'
NFFQ9 'F-5 TEACHER FAIR WITH EVERYBODY'
NFTO10 'F-5 MAKES FOLLOW RULES'
NFTO11 'F-5 TELLS ME I DO GOOD WORK'
NFTO12 'F-5 EXPLAINS LESSON CLEARLY'
NFTO13 'F-5 TEACHER EASY TO UNDERSTAND'
NFTO14 'F-5 TEACHER STAYS IN CLASSROOM'
//CAVE JOB
//JOBPARM LINES=80
//S1 EXEC SPSSX
//DATA/AGRA DD DSN=EISO1MDATA/AGRA,UNIT=DISK,DISP=SHR

TITLE CAVE CREEK FLOATER 3-5 PROGRAM

DATA LIST FILE=DATA/AGRA RECORDS=2

/1 CARD1FS 1
SEQ1FS 2-5
TCHFS 36-40
CLPERFS 41
/2 CARD2FS 1
SEQ2FS 2-5
FFQ1 6
FFQ2 7
FFQ3 8
FFQ4 9
FFQ5 10
FFQ6 11
FFQ7 12
FFQ8 13
FFQ9 14
FFQ10 15
FFQ11 16
FFQ12 17
FFQ13 18
FFQ14 19
FFQ15 20
FFQ16 21
FFQ17 22
FFQ18 23
FFQ19 24
FFQ20 25

RECODE FFQ1 TO FFQ20
(1 = 0) (2 = 1) (3 = 2) (4 = 3)
(8 = 4) INTO NFFQ1 TO NFFQ20

VARIABLE LABELS
CARD1FS 'FLOATER 3-5 CARD 1'
SEQ1FS 'SEQUENCE OF ENTRY CD1'
TCHFS 'TEACHER FLOATER 3-5'
CLPERFS 'CLASS PERIOD F 3-5'
CARD2FS 'FLOATER 3-5 CARD 2'
SEQ2FS 'SEQUENCE OF ENTRY CD2'
NFFQ1 'F-5 MAKES WORK INTERESTING'
NFFQ2 'F-5 SCHOOL DAY INTERESTING'
NFFQ3 'F-5 GO OVER EACH LESSON FINISHED'
NFFQ4 'F-5 DO SAME THING EVERY DAY'
NFFQ5 'F-5 DISCUSS ABOUT SUBJECT STUDIED'
NFFQ6 'F-5 TEACHER USUALLY PREPARED'
NFFQ7 'F-5 MAKES FEEL GOOD / GOOD WORK'
NFFQ8 'F-5 CAN GET HELP FROM TEACHER'
NFFQ9 'F-5 TEACHER FAIR WITH EVERYBODY'
NFFQ10 'F-5 MAKES FOLLOW RULES'
NFFQ11 'F-5 TILTS ME I DO GOOD WORK'
NFFQ12 'F-5 EXPLAINS LESSON CLEARLY'
NFFQ13 'F-5 TEACHER EASY TO UNDERSTAND'
NFFQ14 'F-5 TEACHER STAYS IN CLASSROOM'
VALUE LABELS
NT501 TO NT5020
  0 'NEVER'
  1 'NOT OFTEN'
  2 'SOMETIMES'
  3 'USUALLY'
  4 'ALMOST ALWAYS'/

CLPERTE
  1 'PERIOD 1'
  2 'PERIOD 2'
  3 'PERIOD 3'
  4 'PERIOD 4'
  5 'PERIOD 5'
  6 'PERIOD 6'
  7 'PERIOD 7'
  8 'PERIOD 8'
  9 'NO PERIOD LISTED'/

COMPUTE TOTSCORE = SUM (NT501 TO NT5020)
DESCRIPTIVES VARIABLES = TOTSCORE
/STATISTICS = DEFAULT

SORT CASES BY TCHTS
SPLIT FILE BY TCHTS
DESCRIPTIVES VARIABLES = TOTSCORE
/STATISTICS = DEFAULT
FINISH

TEMPORARY:
SELECT IF (TCHTS EQ 26671)
FREQUENCIES VARIABLES = NT501 TO NT5020/
STATISTICS = DEFAULT

TEMPORARY:
SELECT IF (TCHTS EQ 01552)
FREQUENCIES VARIABLES = NT501 TO NT5020/
STATISTICS = DEFAULT

TEMPORARY:
SELECT IF (TCHTS EQ 04699)
FREQUENCIES VARIABLES = NT501 TO NT5020/
STATISTICS = DEFAULT

TEMPORARY:
SELECT IF (TCHTS EQ 80856)
FREQUENCIES VARIABLES = NT501 TO NT5020/
STATISTICS = DEFAULT

TEMPORARY:
SELECT IF (TCHTS EQ 42558)
FREQUENCIES VARIABLES = NT501 TO NT5020/
STATISTICS = DEFAULT
DATA LIST FILE=DATA2RAW RECORDS=2
/1 CAROITS 1
SEQITS 2-5
TCHITS 36-40
CLPETITS 41
/2 CAR2ITS 1
SEQ2ITS 2-5
TSQ1 6
TSQ2 7
TSQ3 8
TSQ4 9
TSQ5 10
TSQ6 11
TSQ7 12
TSQ8 13
TSQ9 14
TSQ10 15
TSQ11 16
TSQ12 17
TSQ13 18
TSQ14 19
TSQ15 20
TSQ16 21
TSQ17 22
TSQ18 23
TSQ19 24
TSQ20 25
RECODE TSQ1 TO TSQ20 (1 = 0) (2 = 1) (3 = 2) (4 = 3)
(5 = 4) INTO NTSQ1 TO NTSQ20

VARIABLE LABELS
CAROITS 'REGULAR 3-5 CARD 1'
SEQITS 'SEQUENCE OF ENTRY CD1'
TCHITS 'TEACHER REGULAR 3-5'
CLPETITS 'CLASS PERIOD R 3-5'
CAR2ITS 'REGULAR 3-5 CARD 2'
SEQ2ITS 'SEQUENCE OF ENTRY CD2'
NTSQ1 '3-5 MAKES WORK INTERESTING'
N1502 '3-5 SCHOOL DAY INTERESTING'
N1503 '3-5 GO OVER EACH LESSON FINISHED'
N1504 '3-5 GIVES WORK TO DO AT HOME'
N1505 '3-5 DISCUS ABOUT SUBJECT STUDIED'
N1506 '3-5 GIVES WORK BACK QUICKLY'
N1507 '3-5 MAKES FEEL GOOD / GOOD WORK'
N1508 '3-5 CAN GET HELP FROM TEACHER'
N1509 '3-5 FINISH WORK BEFORE CLASS OVER'
N1510 '3-5 MAKES FOLLOW RULES'
N1511 '3-5 NEW WORK WITHOUT WAITING'
N1512 '3-5 EXPLAINS LESSON CLEARLY'
N1513 '3-5 TEACHER KNOWS ME WELL'
DATA LIST FREE / TCHTS, NTSO1 TO NT5Q20.

BEGIN DATA

TCHTS NTSO1 NTSO2 NTSO3 NTSO4 NTSO5 NTSO6 NTSO7 NTSO8 NTSO9 NTSO10 NTSO11 NTSO12 NTSO13 NTSO14 NTSO15 NTSO16 NTSO17 NTSO18 NTSO19 NTSO20

END DATA.

FORMATS NTSO1 TO NT5Q20 (A2) / TCHTS (F6).

VALUE LABELS NTSO1 TO NT5Q20 0 'NEVER' 1 'NOT OBTEN' 2 'SOMETIMES' 3 'USUALLY' 4 'ALMOST ALWAYS'.

COMPUTE TOTSCORE = SUM (NTSO1 TO NT5Q20).

DESCRIPTIVES VARIABLES = TOTSCORE /STATISTICS = DEFAULT.

SORT CASES BY TCHTS.

FREQUENCIES VARIABLES = NTSO1 TO NT5Q20 /STATISTICS = DEFAULT.

TEMPORARY.

SELECT IF (TCHTS EQ 26670).

FREQUENCIES VARIABLES = NTSO1 TO NT5Q20 /STATISTICS = DEFAULT.

SELECT IF (TCHTS EQ 01552).

FREQUENCIES VARIABLES = NTSO1 TO NT5Q20 /STATISTICS = DEFAULT.

SELECT IF (TCHTS EQ 20740).

FREQUENCIES VARIABLES = NTSO1 TO NT5Q20 /STATISTICS = DEFAULT.

SELECT IF (TCHTS EQ 44699).

FREQUENCIES VARIABLES = NTSO1 TO NT5Q20 /STATISTICS = DEFAULT.

SELECT IF (TCHTS EQ 08055).

FREQUENCIES VARIABLES = NTSO1 TO NT5Q20 /STATISTICS = DEFAULT.

SELECT IF (TCHTS EQ 42558).

FREQUENCIES VARIABLES = NTSO1 TO NT5Q20 /STATISTICS = DEFAULT.

FINISH.
TITLE CAVE CREEK Regular 3-5 Program

DATA LIST FILE=DATA2RAW RECORDS=2
/1 CARD1TS 1
/2 CARD2TS 1
SEQ1TS 2-5
TCH1TS 36-40
CLPER1TS 41
/2 CARD2TS 1
SEQ2TS 2-5
T501 6
T502 7
T503 8
T504 9
T505 10
T506 11
T507 12
T508 13
T509 14
T5010 15
T5011 16
T5012 17
T5013 18
T5014 19
T5015 20
T5016 21
T5017 22
T5018 23
T5019 24
T5020 25

RECODE T501 TO T5020 (1 = 0) (2 = 1) (3 = 2) (4 = 3) (5 = 4) INTO NT501 TO NT5020

VARIABLE LABELS
CARD1TS 'Regular 3-5 Card 1'
SEQ1TS 'Sequence of Entry C01'
TCH1TS 'Teacher Regular 3-5'
CLPER1TS 'Class Period R 3-5'
CARD2TS 'Regular 3-5 Card 2'
SEQ2TS 'Sequence of Entry C02'
NT501 '3-5 Makes Work Interesting'
NT502 '3-5 School Day Interesting'
NT503 '3-5 GD Over Each Lesson Finished'
NT504 '3-5 Gives Work To Do At Home'
NT505 '3-5 Discuss About Subject Studied'
NT506 '3-5 Gives Work Back Quickly'
NT507 '3-5 Makes Feel Good / Good Work'
NT508 '3-5 Can Get Help From Teacher'
NT509 '3-5 Finish Work Before Class Over'
NT5010 '3-5 Makes Follow Rules'
NT5011 '3-5 New Work Without Waiting'
NT5012 '3-5 Explains Lesson Clearly'
NT5013 '3-5 Teacher Knows Me Well'
VALUE LABELS

NF2Q1 TO NF2Q20

0 'NEVER'
1 'SOMETIMES'
2 'ALMOST ALWAYS'/

CLPERF2

1 'PERIOD 1'
2 'PERIOD 2'
3 'PERIOD 3'
4 'PERIOD 4'
5 'PERIOD 5'
6 'PERIOD 6'
7 'PERIOD 7'
8 'PERIOD 8'
9 'NO PERIOD LISTED'/

COMPUTE TOTSCLRE = SUM (NF2Q1 TO NF2Q20)

DESCRIPTIVES VARIABLES = TOTSCLRE

STATISTICS = DEFAULT

SORT CASES BY CLPERF2
SPLIT FILE BY CLPERF2

DESCRIPTIVES VARIABLES = TOTSCLRE

STATISTICS = DEFAULT

FINISH

TEMPORARY

SELECT IF (TCHF2 EQ 31212)
FREQUENCIES VARIABLES = NF2Q1 TO NF2Q20/

STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHF2 EQ 34281)
FREQUENCIES VARIABLES = NF2Q1 TO NF2Q20/

STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHF2 EQ 38763)
FREQUENCIES VARIABLES = NF2Q1 TO NF2Q20/

STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHF2 EQ 88767)
FREQUENCIES VARIABLES = NF2Q1 TO NF2Q20/

STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHF2 EQ 88800)
FREQUENCIES VARIABLES = NF2Q1 TO NF2Q20/

STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHF2 EQ 88800)
FREQUENCIES VARIABLES = NF2Q1 TO NF2Q20/

STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHF2 EQ 31212)
FREQUENCIES VARIABLES = NF2Q1 TO NF2Q20/

STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHF2 EQ 34281)
FREQUENCIES VARIABLES = NF2Q1 TO NF2Q20/

STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHF2 EQ 38763)
FREQUENCIES VARIABLES = NF2Q1 TO NF2Q20/

STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHF2 EQ 88767)
FREQUENCIES VARIABLES = NF2Q1 TO NF2Q20/

STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHF2 EQ 88800)
FREQUENCIES VARIABLES = NF2Q1 TO NF2Q20/

STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHF2 EQ 31212)
FREQUENCIES VARIABLES = NF2Q1 TO NF2Q20/

STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHF2 EQ 34281)
FREQUENCIES VARIABLES = NF2Q1 TO NF2Q20/

STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHF2 EQ 38763)
FREQUENCIES VARIABLES = NF2Q1 TO NF2Q20/

STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHF2 EQ 88767)
FREQUENCIES VARIABLES = NF2Q1 TO NF2Q20/

STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHF2 EQ 88800)
FREQUENCIES VARIABLES = NF2Q1 TO NF2Q20/

STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHF2 EQ 31212)
FREQUENCIES VARIABLES = NF2Q1 TO NF2Q20/

STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHF2 EQ 34281)
FREQUENCIES VARIABLES = NF2Q1 TO NF2Q20/

STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHF2 EQ 38763)
FREQUENCIES VARIABLES = NF2Q1 TO NF2Q20/

STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHF2 EQ 88767)
FREQUENCIES VARIABLES = NF2Q1 TO NF2Q20/

STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHF2 EQ 88800)
FREQUENCIES VARIABLES = NF2Q1 TO NF2Q20/

STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHF2 EQ 31212)
FREQUENCIES VARIABLES = NF2Q1 TO NF2Q20/

STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHF2 EQ 34281)
FREQUENCIES VARIABLES = NF2Q1 TO NF2Q20/

STATISTICS = DEFAULT
TITLE CAVE CREEK FLOATER K-2 PROGRAM
DATA LIST FILE=DATASRAW RECORDS=2
/1 CARDIF2 1
/2 CARD2F2 1
SEQIF2 1-5
TCHF2 .36-40
CLPERF2 41
SEQ2F2 2-8
FEQ1 6
FEQ2 7
FEQ3 8
FEQ4 9
FEQ5 10
FEQ6 11
FEQ7 12
FEQ8 13
FEQ9 14
FEQ10 15
FEQ11 16
FEQ12 17
FEQ13 18
FEQ14 19
FEQ15 20
FEQ16 21
FEQ17 22
FEQ18 23
FEQ19 24
FEQ20 25
RE CODE FEQ1 TO FEQ20 (1=0) (2=2) (3=4)
INTO NFQ01 TO NFQ20
VARIABLE LABELS
CARD1F2 'FLOATER K-2 CARD 1'
SEQ1F2 'SEQUENCE OF ENTRY CD1'
TCHF2 'TEACHER FLOATER K-2'
CLPERF2 'CLASS PERIOD F K-2'
CARD2F2 'FLOATER K-2 CARD 2'
SEQ2F2 'SEQUENCE OF ENTRY CD2'
NFQ01 'F-2 MY SCHOOL DAY INTERESTING'
NFQ02 'F-2 DD SAME THING EVERY DAY'
NFQ03 'F-2 PAY ATTENTION IN CLASS'
NFQ04 'F-2 DISCUSSIONS LESSONS STUDIED'
NFQ05 'F-2 WORK TOO HARD'
NFQ06 'F-2 TEACHER USUALLY PREPARED'
NFQ07 'F-2 TEACHER COMES ON TIME'
NFQ08 'F-2 MAKES FOLLOW RULES'
NFQ09 'F-2 FAIR WITH EVERYBODY'
NFQ10 'F-2 CARES IF WASTE TIME'
NFQ11 'F-2 WORK IF TEACHER NOT WATCHING'
NFQ12 'F-2 CAN GET HELP WHEN NEED IT'
NFQ13 'F-2 TELLS ME I DO GOOD WORK'
NFQ14 'F-2 TELLS ME WHERE TO FIND INFORMATION'
NFQ15 'F-2 TEACHER READY FOR CLASS'
NFQ16 'F-2 KNOW WHAT TEACHER WANTS'
NK2015 'K 2 TEACHER READY FOR CLASS'
NK2016 'K-2 KNOW WHAT TEACHER WANTS'
NK2017 'K-2 INTEREST WORK BEFORE CLASS OVER'
NK2018 'K-2 LEARN HARD LESSONS SMALL STEPS'
NK2019 'K-2 GIVES WORK BACK QUICKLY'
NK2020 'K-2 TELLS NEW THINGS TO LEARN'

VALUE LABELS
NK201 TO NK2020
0 'NEVER'
2 'SOMETIMES'
4 'ALMOST ALWAYS'

CLPERK2
1 'PERIOD 1'
2 'PERIOD 2'
3 'PERIOD 3'
4 'PERIOD 4'
5 'PERIOD 5'
6 'PERIOD 6'
7 'PERIOD 7'
8 'PERIOD 8'
0 'NO PERIOD LISTED'

COMPUTE TOTSCORE = SUM (NK201 TO NK2020)

DESCRIPTIVES VARIABLES = TOTSCORE
/STATISTICS = DEFAULT

SORT CASES BY TCHK2

SPLIT FILE BY TCHK2

DESCRIPTIVES VARIABLES = TOTSCORE
/STATISTICS = DEFAULT

FINISH

TEMPORARY

SELECT IF (TCHK2 EQ 24190)
FREQUENCIES VARIABLES = NK201 TO NK2020
/STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHK2 EQ 25238)
FREQUENCIES VARIABLES = NK201 TO NK2020
/STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHK2 EQ 26597)
FREQUENCIES VARIABLES = NK201 TO NK2020
/STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHK2 EQ 36103)
FREQUENCIES VARIABLES = NK201 TO NK2020
/STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHK2 EQ 44634)
FREQUENCIES VARIABLES = NK201 TO NK2020
/STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHK2 EQ 28084)
FREQUENCIES VARIABLES = NK201 TO NK2020
/STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHK2 EQ 60398)
FREQUENCIES VARIABLES = NK201 TO NK2020
/STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHK2 EQ 08834)
FREQUENCIES VARIABLES = NK201 TO NK2020
/STATISTICS = DEFAULT
TITLE CAVE CREEK REGULAR K-2 PROGRAM

DATA LIST FILE=DATAIRAW RECORDS=2
/1 CARD1K2 1
/1 SEQ1K2 2-5
/1 TCHK2 36-40
/2 CARD2K2 1
/2 SEQ2K2 2-6
/2 K2Q1 6
/2 K2Q2 7
/2 K2Q3 8
/2 K2Q4 9
/2 K2Q5 10
/2 K2Q6 11
/2 K2Q7 12
/2 K2Q8 13
/2 K2Q9 14
/2 K2Q10 15
/2 K2Q11 16
/2 K2Q12 17
/2 K2Q13 18
/2 K2Q14 19
/2 K2Q15 20
/2 K2Q16 21
/2 K2Q17 22
/2 K2Q18 23
/2 K2Q19 24
/2 K2Q20 25
/2 K2Q21 26
/2 K2Q22 27
/2 K2Q23 28
/2 K2Q24 29
/2 K2Q25 30
/2 K2Q26 31
/2 K2Q27 32
/2 K2Q28 33
/2 K2Q29 34
/2 K2Q30 35
/2 RECODE K2Q1 TO K2Q25 (1 * 0) (2 * 2) (3 * 4)
INTO NK2Q1 TO NK2Q30

VARIABLE LABELS
CARD1K2 'REGULAR K-2 CARD 1'
SEQ1K2 'SEQUENCE OF ENTRY CD1'
TCHK2 'TEACHER REGULAR K-2'
CLPERK2 'CLASS PERIOD R K-2'
CARD2K2 'REGULAR K-2 CARD 2'
SEQ2K2 'SEQUENCE OF ENTRY CD2'
NK2Q1 'K-2 SCHOOL DAY INTERESTING'
NK2Q2 'K-2 TIME TO DO OUR WORK'
NK2Q3 'K-2 PAY ATTENTION IN CLASS'
NK2Q4 'K-2 DISCUSSIONS LESSONS STUDIED'
NK2Q5 'K-2 WORK TOO HARD'
NK2Q6 'K-2 GIVES HOMEWORK'
NK2Q7 'K-2 TEACHER COMES ON TIME'
NK2Q8 'K-2 MAKES FOLLOW RULES'
NK2Q9 'K-2 OFTEN TAKE TEST IN CLASS'
NK2Q10 'K-2 CARES IF WASTE TIME'
NK2Q11 'K-2 WORK IF TEACHER NOT WATCHING'
NK2Q12 'K-2 CAN GET HELP WHEN NEED IT'
NK2Q13 'K-2 GIVES NEW WORK WHEN I AM READY'
NK2Q14 'K-2 TELLS WHERE TO FIND INFORMATION'
I. TEACHER READY FOR CLASS

II. KNOW WHAT TEACHER WANTS

III. INTEREST WORK BEFORE CLASS OVER

IV. LEARN HARD LESSONS SMALL STEPS

V. GIVES WORK BACK QUICKLY

VI. TELLS NEW THINGS TO LEARN

VII. NEVER

VIII. SOMETIMES

IX. ALMOST ALWAYS

X. PERIOD 1

XI. PERIOD 2

XII. PERIOD 3

XIII. PERIOD 4

XIV. PERIOD 5

XV. PERIOD 6

XVI. PERIOD 7

XVII. PERIOD 8

XVIII. NO PERIOD LISTED

VALUE LABELS

NK2Q1 TO NK2Q20

0 'NEVER'

2 'SOMETIMES'

4 'ALMOST ALWAYS'

CLPERK2

1 'PERIOD 1'

2 'PERIOD 2'

3 'PERIOD 3'

4 'PERIOD 4'

5 'PERIOD 5'

6 'PERIOD 6'

7 'PERIOD 7'

8 'PERIOD 8'

9 'NO PERIOD LISTED'

COMPUTE TDTSORE = SUM (NK2Q1 TO NK2Q20)

DESCRIPTIVES VARIABLES = TDTSORE /STATISTICS = DEFAULT

SORT CASES BY TCHK2

SPLIT FILE BY TCHK2

DESCRIPTIVES VARIABLES = TDTSORE /STATISTICS = DEFAULT

FINISH

TEMPORARY

SELECT IF (TCHK2 EQ 24191)

FREQUENCIES VARIABLES = NK2Q1 TO NK2Q20 /STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHK2 EQ 25238)

FREQUENCIES VARIABLES = NK2Q1 TO NK2Q20 /STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHK2 EQ 20597)

FREQUENCIES VARIABLES = NK2Q1 TO NK2Q20 /STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHK2 EQ 36103)

FREQUENCIES VARIABLES = NK2Q1 TO NK2Q20 /STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHK2 EQ 44634)

FREQUENCIES VARIABLES = NK2Q1 TO NK2Q20 /STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHK2 EQ 28084)

FREQUENCIES VARIABLES = NK2Q1 TO NK2Q20 /STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHK2 EQ 60399)

FREQUENCIES VARIABLES = NK2Q1 TO NK2Q20 /STATISTICS = DEFAULT

TEMPORARY

SELECT IF (TCHK2 EQ 08834)

FREQUENCIES VARIABLES = NK2Q1 TO NK2Q20 /STATISTICS = DEFAULT

FINISH
TITLE CAVE CREEK REGULAR K-2 PROGRAM

DATA LIST FILE=DATARAW
RECORDS=2

/1 CARD K2 1
SEQK2 2-5
TCMK2 36-40
CLPERK2 41
/2 CARD K2 1
SEQK2 2-6
K2Q3 6
K2Q4 7
K2Q5 8
K2Q6 9
K2Q7 10
K2Q8 11
K2Q9 12
K2Q10 13
K2Q11 14
K2Q12 15
K2Q13 16
K2Q14 17
K2Q15 18
K2Q16 19
K2Q17 20
K2Q18 21
K2Q19 22
K2Q20 23
K2Q21 24
K2Q22 25
K2Q23 26
K2Q24 27
K2Q25 28
K2Q26 29
K2Q27 30
K2Q28 31
K2Q29 32
K2Q30 33
K2Q31 34
K2Q32 35
K2Q33 36
K2Q34 37
K2Q35 38
K2Q36 39
K2Q37 40
K2Q38 41
K2Q39 42
K2Q40 43
K2Q41 44
K2Q42 45
K2Q43 46
K2Q44 47
K2Q45 48
K2Q46 49
K2Q47 50
K2Q48 51
K2Q49 52
K2Q50 53
K2Q51 54
K2Q52 55
K2Q53 56
K2Q54 57
K2Q55 58
K2Q56 59
K2Q57 60

RECODE K2Q1 TO K2Q20 (1 = 0) (2 = 0) (3 = 4)

VARIABLE LABELS
CARDK2 'REGULAR K-2 CARD 1'
SEQK2 'SEQUENCE OF ENTRY CD1'
TCMK2 'TEACHER REGULAR K-2'
CLPERK2 'CLASS PERIOD & K-2'
GARK2 'REGULAR K-2 CARD 2'
SEQK2 'SEQUENCE OF ENTRY CD2'
K2Q1 'K-2 SCHOOL DAY INTERESTING'
K2Q2 'K-2 TIME TO DO OUR WORK'
K2Q3 'K-2 PAY ATTENTION IN CLASS'
K2Q4 'K-2 DISCUSSIONS LESSONS STUDIED'
K2Q5 'K-2 WORK TOO HARD'
K2Q6 'K-2 GIVES HOMEWORK'
K2Q7 'K-2 TEACHER COMES ON TIME'
K2Q8 'K-2 MAKES FOLLOW RULES'
K2Q9 'K-2 OFTEN TAKE TEST IN CLASS'
K2Q10 'K-2 CARES IF WASTE TIME'
K2Q11 'K-2 WORK IF TEACHER NOT WATCHING'
K2Q12 'K-2 CAN GET HELP WHEN NEED IT'
K2Q13 'K-2 GIVES NEW WORK WHEN I AM READY'
K2Q14 'K-2 TELLS WHERE TO FIND INFORMATION'
APPENDIX L.

CAVE CREEK CAREER LADDER PLAN ALGORITHM
Scoring of Student Feedback Summative Report

Student feedback surveys are conducted by each teacher late in each course offered (semester or year). Primary grades (K-2) are surveyed with an instrument that uses a three-point scale; upper elementary, middle school and high school surveys use a five-point scale.

The primary grades' instrument has 20 questions which will be tabulated with the following values: [This item describes my teacher]

No = 0
Sometimes = 2
Almost always = 4

Thus a teacher receiving an "almost always" rating on each of the 20 items would receive a total rating of 80. Next, all of the ratings of all of his/her students would be averaged and the average rating would be transformed by the following conversion table:

<table>
<thead>
<tr>
<th>Average Rating Total</th>
<th>Composite Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>70-80</td>
<td>10.0</td>
</tr>
<tr>
<td>60-69</td>
<td>7.5</td>
</tr>
<tr>
<td>50-59</td>
<td>5.0</td>
</tr>
<tr>
<td>&lt;50</td>
<td>2.5</td>
</tr>
</tbody>
</table>

Scoring for teachers in the upper elementary school, middle school and high school would be done in a similar manner. The rating scale for these grades uses the following response mode and values: [Describes my class or teacher]

Never = 0
Not Often = 1
Sometimes = 2
Usually = 3
Almost always = 4
APPENDIX M.

HUMAN SUBJECTS RELEASE LETTER
FROM DAVE ALEXANDER
October 25, 1990

Dr. Richard P. Manatt, Director
School Improvement Model
Iowa State University
2926 Monroe Dr.
Ames, Iowa 50010

Dear Dr. Manatt:

This memorandum grants you permission to use Cave Creek School District student feedback data for conducting analysis of same for District purposes.

You also are authorized to use such data, assign such data, and to have analyzed such data, for use in dissertations by appropriate candidates under your supervision.

Sincerely,

David C. Alexander, Ed.D.
Superintendent

DCA:jjr
APPENDIX N.

1990 CAVE CREEK STUDENT RATINGS OF TEACHERS
ANALYSIS AND SUMMARY
### 1990 Cave Creek Student Ratings of Teachers

#### Analysis Based on 465 Subjects in 19 Groups

<table>
<thead>
<tr>
<th>ITEM CODE</th>
<th>DESCRIPTIVE</th>
<th>N</th>
<th>MEAN</th>
<th>VARIANCE</th>
<th>ITEM DISCRIMINATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 S</td>
<td>'K-2 School Day Interesting'</td>
<td>429</td>
<td>2.27792</td>
<td>0.59309</td>
<td>4%</td>
</tr>
<tr>
<td>2 S</td>
<td>'K-2 Time To Do Our Work'</td>
<td>423</td>
<td>2.52246</td>
<td>0.50481</td>
<td>14%</td>
</tr>
<tr>
<td>3 S</td>
<td>'K-2 Pay Attention in Class'</td>
<td>422</td>
<td>2.82938</td>
<td>0.19364</td>
<td>8%</td>
</tr>
<tr>
<td>4 SR</td>
<td>'K-2 Discussions Lessons Studied'</td>
<td>424</td>
<td>2.40569</td>
<td>0.51984</td>
<td>8%</td>
</tr>
<tr>
<td>5 S</td>
<td>'K-2 Work Too Hard'</td>
<td>426</td>
<td>1.44693</td>
<td>0.47738</td>
<td>5%</td>
</tr>
<tr>
<td>6 SR</td>
<td>'K-2 Gives Homework'</td>
<td>428</td>
<td>2.53972</td>
<td>0.50943</td>
<td>23%</td>
</tr>
<tr>
<td>7 SR</td>
<td>'K-2 Teacher Comes On Time'</td>
<td>427</td>
<td>2.61593</td>
<td>0.41435</td>
<td>6%</td>
</tr>
<tr>
<td>8 CC</td>
<td>'K-2 Makes Follow Rules'</td>
<td>427</td>
<td>2.79931</td>
<td>0.28071</td>
<td>8%</td>
</tr>
<tr>
<td>9 S</td>
<td>'K-2 Often Take Test In Class'</td>
<td>428</td>
<td>2.01636</td>
<td>0.70768</td>
<td>20%</td>
</tr>
<tr>
<td>10 SR</td>
<td>'K-2 Cares If Waste Time'</td>
<td>426</td>
<td>2.46714</td>
<td>0.55268</td>
<td>4%</td>
</tr>
<tr>
<td>11 SR</td>
<td>'K-2 Work If Teacher Not Watching'</td>
<td>427</td>
<td>2.56674</td>
<td>0.56405</td>
<td>6%</td>
</tr>
<tr>
<td>12 SR</td>
<td>'K-2 Can Get Help When Need It'</td>
<td>425</td>
<td>2.62353</td>
<td>0.38533</td>
<td>12%</td>
</tr>
<tr>
<td>13 SR</td>
<td>'K-2 Gives New Work When I am Ready'</td>
<td>427</td>
<td>2.43794</td>
<td>0.55328</td>
<td>10%</td>
</tr>
<tr>
<td>14 S</td>
<td>'K-2 Tells Where to Find Information'</td>
<td>425</td>
<td>2.46010</td>
<td>0.61208</td>
<td>7%</td>
</tr>
<tr>
<td>15 SR</td>
<td>'K-2 Teacher Ready for Class'</td>
<td>423</td>
<td>2.72176</td>
<td>0.33623</td>
<td>4%</td>
</tr>
<tr>
<td>16 SR</td>
<td>'K-2 Know What Teacher Wants'</td>
<td>429</td>
<td>2.21445</td>
<td>0.67652</td>
<td>11%</td>
</tr>
<tr>
<td>17 SR</td>
<td>'K-2 Interest Work Before Class Over'</td>
<td>422</td>
<td>2.16568</td>
<td>0.67391</td>
<td>11%</td>
</tr>
<tr>
<td>18 S</td>
<td>'K-2 Learn Hard Lessons Small Steps'</td>
<td>422</td>
<td>2.33469</td>
<td>0.64507</td>
<td>9%</td>
</tr>
<tr>
<td>19 S</td>
<td>'K-2 Gives Work Back Quickly'</td>
<td>424</td>
<td>1.95755</td>
<td>0.68688</td>
<td>23%</td>
</tr>
<tr>
<td>20 SR</td>
<td>'K-2 Tells New Things to Learn'</td>
<td>424</td>
<td>2.59650</td>
<td>0.45434</td>
<td>11%</td>
</tr>
</tbody>
</table>

**Cronbach Alpha Reliability Based on 6 Items with Discrimination >= 13% is 0.968**

### Notes

**Descriptive** = an abbreviated label used to identify the specific student rating question.

**Code** = S = original question (known to be a valid and reliable discriminating item) from the School Improvement Model (SIM).

**SR** = revised question from the original SIM listing of discriminating items.

**CC** = new questions developed by Cave Creek Schools.

**N** = total number of students who completed a teacher rating questionnaire.

**Mean** = the arithmetic average of the total student responses for each item.

**Variance** = the relationship of scores to a central value, such as the mean. Variance is defined as the sum of squared deviations around the mean. This statistic describes how similar or different the scores are, from the mean, for a given group.

**Item Discrimination** = A question is considered to be most effective when it has a high level of item discrimination. This means that the question is useful in separating high teacher performance from that of average and low performance. A percentage value of at least 13% equals discrimination (at the .05 level of significance) and a percentage value of at least 22% equals discrimination at the .01 level of significance.
## CAVE CREEK SPECIAL MENNE-TOLLSMA ANALYSIS BASED ON 339 SUBJECTS IN 4 GROUPS.

### 1990 CAVE CREEK STUDENT RATINGS OF TEACHERS

<table>
<thead>
<tr>
<th>ITEM</th>
<th>CODE</th>
<th>DESCRIPTIVE</th>
<th>N</th>
<th>MEAN</th>
<th>VARIANCE</th>
<th>ITEM DISCRIMINATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SR</td>
<td>'F-2 MY SCHOOL DAY INTERESTING'</td>
<td>286</td>
<td>2.25874</td>
<td>0.59739</td>
<td>1%</td>
</tr>
<tr>
<td>2</td>
<td>S</td>
<td>'F-2 DO SAME THING EVERY DAY'</td>
<td>285</td>
<td>1.56942</td>
<td>0.56913</td>
<td>2%</td>
</tr>
<tr>
<td>3</td>
<td>S</td>
<td>'F-2 PAY ATTENTION IN CLASS'</td>
<td>281</td>
<td>2.76868</td>
<td>0.27034</td>
<td>3%</td>
</tr>
<tr>
<td>4</td>
<td>SR</td>
<td>'F-2 DISCUSSIONS LESSONS STUDIED'</td>
<td>281</td>
<td>2.32028</td>
<td>0.54510</td>
<td>3%</td>
</tr>
<tr>
<td>5</td>
<td>S</td>
<td>'F-2 WORK TOO HARD'</td>
<td>281</td>
<td>1.53381</td>
<td>0.52644</td>
<td>2%</td>
</tr>
<tr>
<td>6</td>
<td>SR</td>
<td>'F-2 TEACHER USUALLY PREPARED'</td>
<td>283</td>
<td>2.53004</td>
<td>0.52472</td>
<td>1%</td>
</tr>
<tr>
<td>7</td>
<td>SR</td>
<td>'F-2 TEACHER COMES ON TIME'</td>
<td>285</td>
<td>2.50877</td>
<td>0.45343</td>
<td>1%</td>
</tr>
<tr>
<td>8</td>
<td>CC</td>
<td>'F-2 MAKES FOLLOW RULES'</td>
<td>277</td>
<td>2.70759</td>
<td>0.40908</td>
<td>4%</td>
</tr>
<tr>
<td>9</td>
<td>S</td>
<td>'F-2 CARES IF WASTE TIME'</td>
<td>280</td>
<td>2.44643</td>
<td>0.63284</td>
<td>2%</td>
</tr>
<tr>
<td>10</td>
<td>SR</td>
<td>'F-2 WORK IF TEACHER NOT WATCHING'</td>
<td>281</td>
<td>2.45196</td>
<td>0.65339</td>
<td>3%</td>
</tr>
<tr>
<td>11</td>
<td>SR</td>
<td>'F-2 CAN GET HELP WHEN NEED IT'</td>
<td>275</td>
<td>2.52000</td>
<td>0.46091</td>
<td>4%</td>
</tr>
<tr>
<td>13</td>
<td>CC</td>
<td>'F-2 TELLS ME I DO GOOD WORK'</td>
<td>271</td>
<td>2.53137</td>
<td>0.41876</td>
<td>5%</td>
</tr>
<tr>
<td>14</td>
<td>S</td>
<td>'F-2 TELLS WHERE TO FIND INFORMATION'</td>
<td>281</td>
<td>2.20976</td>
<td>0.68545</td>
<td>3%</td>
</tr>
<tr>
<td>15</td>
<td>SR</td>
<td>'F-2 TEACHER READY FOR CLASS'</td>
<td>274</td>
<td>2.58999</td>
<td>0.44062</td>
<td>3%</td>
</tr>
<tr>
<td>16</td>
<td>SR</td>
<td>'F-2 KNOW WHAT TEACHER WANTS'</td>
<td>282</td>
<td>2.14094</td>
<td>0.64913</td>
<td>1%</td>
</tr>
<tr>
<td>17</td>
<td>CC</td>
<td>'F-2 TEACHER EASY TO UNDERSTAND'</td>
<td>283</td>
<td>2.47350</td>
<td>0.51788</td>
<td>4%</td>
</tr>
<tr>
<td>18</td>
<td>S</td>
<td>'F-2 LEARN HARD LESSONS SMALL STEPS'</td>
<td>281</td>
<td>2.16370</td>
<td>0.69918</td>
<td>2%</td>
</tr>
<tr>
<td>19</td>
<td>S</td>
<td>'F-2 EXPLAIN WAYS EASY TO UNDERSTAND'</td>
<td>285</td>
<td>2.55088</td>
<td>0.50705</td>
<td>3%</td>
</tr>
<tr>
<td>20</td>
<td>S</td>
<td>'F-2 TELLS NEW THINGS TO LEARN'</td>
<td>285</td>
<td>2.53947</td>
<td>0.40343</td>
<td>3%</td>
</tr>
</tbody>
</table>

**Cronbach Alpha Reliability Based On 16 Items With Discrimination >= 1% Is 0.000**

**Descriptive** = An abbreviated label used to identify the specific student rating question.

**Code** = S = original question (known to be a valid and reliable discriminating item) from the School Improvement Model (SIM).

SR = revised question from the original SIM listing of discriminating items.

CC = new questions developed by Cave Creek Schools.

**N** = Total number of students who completed a teacher rating questionnaire.

**Mean** = The arithmetic average of the total student responses for each item.

**Variance** = The relationship of scores to a central value, such as the mean. Variance is defined as the sum of squared deviations around the mean. This statistic describes how similar or different the scores are, from the mean, for a given group.

**Item Discrimination** = A question is considered to be most effective when it has a high level of item discrimination. This means that the question is useful in separating high teacher performance from that of average and low performance. A percentage value of at least 13% equals discrimination (at the .05 level of significance) and a percentage value of at least 22% equals discrimination at the .01 level of significance.
## 1990 Cave Creek Student Ratings of Teachers

### Analysis Based on 805 Subjects in 36 Groups

<table>
<thead>
<tr>
<th>Item</th>
<th>Code</th>
<th>Descriptive</th>
<th>N</th>
<th>Mean</th>
<th>Variance</th>
<th>Item Discrimination</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>S</td>
<td>'3-5 Makes Work Interesting'</td>
<td>798</td>
<td>3.68546</td>
<td>1.40397</td>
<td>22%</td>
</tr>
<tr>
<td>2</td>
<td>SR</td>
<td>'3-5 School Day Interesting'</td>
<td>799</td>
<td>3.40766</td>
<td>1.46784</td>
<td>19%</td>
</tr>
<tr>
<td>3</td>
<td>SR</td>
<td>'3-5 Go Over Each Lesson Finished'</td>
<td>796</td>
<td>3.47995</td>
<td>1.54784</td>
<td>10%</td>
</tr>
<tr>
<td>4</td>
<td>S</td>
<td>'3-5 Gives Work to Do at Home'</td>
<td>798</td>
<td>3.61153</td>
<td>1.43556</td>
<td>30%</td>
</tr>
<tr>
<td>5</td>
<td>SR</td>
<td>'3-5 Discuss About Subject Studied'</td>
<td>791</td>
<td>4.26070</td>
<td>1.04602</td>
<td>11%</td>
</tr>
<tr>
<td>6</td>
<td>S</td>
<td>'3-5 Gives Work Back Quickly'</td>
<td>798</td>
<td>2.92629</td>
<td>1.44443</td>
<td>20%</td>
</tr>
<tr>
<td>7</td>
<td>S</td>
<td>'3-5 Makes Feel Good / Good Work'</td>
<td>797</td>
<td>4.05646</td>
<td>1.44098</td>
<td>16%</td>
</tr>
<tr>
<td>8</td>
<td>S</td>
<td>'3-5 Can Get Help From Teacher'</td>
<td>798</td>
<td>4.36301</td>
<td>0.90302</td>
<td>11%</td>
</tr>
<tr>
<td>9</td>
<td>S</td>
<td>'3-5 Finish Work Before Class Over'</td>
<td>798</td>
<td>3.66541</td>
<td>1.08730</td>
<td>12%</td>
</tr>
<tr>
<td>10</td>
<td>CC</td>
<td>'3-5 Makes Follow Rules'</td>
<td>799</td>
<td>4.65707</td>
<td>0.52821</td>
<td>7%</td>
</tr>
<tr>
<td>11</td>
<td>SR</td>
<td>'3-5 New Work Without Waiting'</td>
<td>799</td>
<td>3.70970</td>
<td>1.32809</td>
<td>10%</td>
</tr>
<tr>
<td>12</td>
<td>S</td>
<td>'3-5 Explains Lesson Clearly'</td>
<td>799</td>
<td>4.31246</td>
<td>0.50502</td>
<td>16%</td>
</tr>
<tr>
<td>13</td>
<td>S</td>
<td>'3-5 Teacher Knows Me Well'</td>
<td>796</td>
<td>4.32915</td>
<td>1.12031</td>
<td>11%</td>
</tr>
<tr>
<td>14</td>
<td>SR</td>
<td>'3-5 Work to Do Finished Before Over'</td>
<td>799</td>
<td>3.49186</td>
<td>1.65921</td>
<td>11%</td>
</tr>
<tr>
<td>15</td>
<td>SR</td>
<td>'3-5 Work at the Right Pace'</td>
<td>793</td>
<td>4.03258</td>
<td>1.40746</td>
<td>13%</td>
</tr>
<tr>
<td>16</td>
<td>S</td>
<td>'3-5 Tells New Things Learn in Lesson'</td>
<td>796</td>
<td>3.79377</td>
<td>1.44499</td>
<td>14%</td>
</tr>
<tr>
<td>17</td>
<td>S</td>
<td>'3-5 Explain Way Easy to Understand'</td>
<td>799</td>
<td>4.14216</td>
<td>1.17864</td>
<td>17%</td>
</tr>
<tr>
<td>18</td>
<td>CC</td>
<td>'3-5 Teacher Available to Help'</td>
<td>792</td>
<td>3.28027</td>
<td>1.29423</td>
<td>13%</td>
</tr>
<tr>
<td>19</td>
<td>S</td>
<td>'3-5 Variety Activities Resources'</td>
<td>794</td>
<td>3.81234</td>
<td>1.43723</td>
<td>23%</td>
</tr>
<tr>
<td>20</td>
<td>SR</td>
<td>'3-5 Teacher Well Prepared'</td>
<td>793</td>
<td>4.44248</td>
<td>0.99764</td>
<td>13%</td>
</tr>
</tbody>
</table>

**Cronbach Alpha Reliability Based on 12 Items with Discrimination > = 13% is** 0.863

**Description**

- **Descriptive** = an abbreviated label used to identify the specific student rating question.
- **Code** = S = original question (known to be a valid and reliable discriminating item) from the School Improvement Model (SIM).
- **SR** = revised question from the original SIM listing of discriminating items.
- **CC** = new questions developed by Cave Creek Schools
- **N** = total number of students who completed a teacher rating questionnaire.
- **Mean** = the arithmetic average of the total student responses for each item.
- **Variance** = the relationship of scores to a central value, such as the mean. Variance is defined as the sum of squared deviations around the mean. This statistic describes how similar or different the scores are, from the mean, for a given group.

**Item Discrimination** = A question is considered to be most effective when it has a high level of item discrimination. This means that the question is useful in separating high teacher performance from that of average and low performance. A percentage value of at least 13% equals discrimination (at the .05 level of significance) and a percentage value of at least 22% equals discrimination at the .01 level of significance.
### SPECIAL B. M. ELEMENTARY SCHOOL MENNE-TOLSKA

#### 1990 CAVE CREEK STUDENT RATINGS OF TEACHERS

<table>
<thead>
<tr>
<th>ITEM</th>
<th>CODE</th>
<th>DESCRIPTIVE</th>
<th>N</th>
<th>MEAN</th>
<th>VARIANCE</th>
<th>ITEM DISCRIMINATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>S</td>
<td>'F-S MAKES WORK INTERESTING'</td>
<td>407</td>
<td>3.76413</td>
<td>1.92471</td>
<td>12%</td>
</tr>
<tr>
<td>2</td>
<td>SR</td>
<td>'F-S SCHOOL DAY INTERESTING'</td>
<td>407</td>
<td>3.59214</td>
<td>1.65183</td>
<td>12%</td>
</tr>
<tr>
<td>3</td>
<td>SR</td>
<td>'F-S DO SAME THING EVERY DAY'</td>
<td>407</td>
<td>2.49877</td>
<td>2.20577</td>
<td>16%</td>
</tr>
<tr>
<td>4</td>
<td>S</td>
<td>'F-S DISCUSS ABOUT SUBJECT STUDIED'</td>
<td>406</td>
<td>3.95813</td>
<td>1.67597</td>
<td>9%</td>
</tr>
<tr>
<td>5</td>
<td>SR</td>
<td>'F-S TEACHER USUALLY PREPARED'</td>
<td>407</td>
<td>4.12032</td>
<td>1.9656</td>
<td>17%</td>
</tr>
<tr>
<td>6</td>
<td>S</td>
<td>'F-S MAKES FEEL GOOD / GOOD WORK'</td>
<td>407</td>
<td>3.68059</td>
<td>2.07488</td>
<td>6%</td>
</tr>
<tr>
<td>7</td>
<td>S</td>
<td>'F-S CAN GET HELP FROM TEACHER'</td>
<td>407</td>
<td>3.79115</td>
<td>1.59027</td>
<td>6%</td>
</tr>
<tr>
<td>8</td>
<td>S</td>
<td>'F-S TEACHER FAIR WITH EVERYBODY'</td>
<td>403</td>
<td>3.97284</td>
<td>1.59479</td>
<td>6%</td>
</tr>
<tr>
<td>9</td>
<td>CC</td>
<td>'F-S MAKES FOLLOW RULES'</td>
<td>407</td>
<td>4.36855</td>
<td>1.24009</td>
<td>10%</td>
</tr>
<tr>
<td>10</td>
<td>CC</td>
<td>'F-S TELLS ME I DO GOOD WORK'</td>
<td>406</td>
<td>3.56665</td>
<td>1.87324</td>
<td>7%</td>
</tr>
<tr>
<td>11</td>
<td>S</td>
<td>'F-S EXPLAINS LESSON CLEARLY'</td>
<td>406</td>
<td>4.09389</td>
<td>1.77779</td>
<td>6%</td>
</tr>
<tr>
<td>12</td>
<td>CC</td>
<td>'F-S TEACHER EASY TO UNDERSTAND'</td>
<td>404</td>
<td>3.84453</td>
<td>1.62494</td>
<td>12%</td>
</tr>
<tr>
<td>13</td>
<td>CC</td>
<td>'F-S TEACHER STAYS IN CLASSROOM'</td>
<td>406</td>
<td>4.09172</td>
<td>1.63033</td>
<td>12%</td>
</tr>
<tr>
<td>14</td>
<td>SR</td>
<td>'F-S TEACHER WORKS AT THE RIGHT PACE'</td>
<td>407</td>
<td>3.82801</td>
<td>1.78943</td>
<td>3%</td>
</tr>
<tr>
<td>15</td>
<td>S</td>
<td>'F-S TELLS NEW THINGS LEARN IN LESSON'</td>
<td>406</td>
<td>3.57882</td>
<td>1.99077</td>
<td>3%</td>
</tr>
<tr>
<td>16</td>
<td>S</td>
<td>'F-S EXPLAIN WAY EASY TO UNDERSTAND'</td>
<td>407</td>
<td>3.88482</td>
<td>1.52721</td>
<td>10%</td>
</tr>
<tr>
<td>17</td>
<td>S</td>
<td>'F-S TEACH AVAIL TO HELP'</td>
<td>405</td>
<td>3.54139</td>
<td>1.47379</td>
<td>7%</td>
</tr>
<tr>
<td>18</td>
<td>CC</td>
<td>'F-S TEACHER KNOW ABOUT LESSON TAUGHT'</td>
<td>405</td>
<td>4.42365</td>
<td>1.14569</td>
<td>9%</td>
</tr>
<tr>
<td>19</td>
<td>SR</td>
<td>'F-S TEACHER WELL PREPARED'</td>
<td>397</td>
<td>4.21159</td>
<td>1.42123</td>
<td>9%</td>
</tr>
</tbody>
</table>

**Cronbach Alpha Reliability Based On** 5 Items With Discrimination = 13% Is 0.701

**DESCRIPTIVE** = an abbreviated label used to identify the specific student rating question.

**CODE** = 8 = original question (known to be a valid and reliable discriminating item) from the School Improvement Model (SIM).

SR = revised question from the original SIM listing of discriminating items.

CC = new questions developed by Cave Creek Schools

N = total number of students who completed a teacher rating questionnaire.

**MEAN** = the arithmetic average of the total student responses for each item.

**VARIANCE** = the relationship of scores to a central value, such as the mean. Variance is defined as the sum of squared deviations around the mean. This statistic describes how similar or different the scores are, from the mean, for a given group.

**ITEM DISCRIMINATION** = A question is considered to be most effective when it has a high level of item discrimination. This means that the question is useful in separating high teacher performance from that of average and low performance. A percentage value of at least 13% equals discrimination (at the .05 level of significance) and a percentage value of at least 22% equals discrimination at the .01 level of significance.
**1990 CAVCreek Student Ratings of Teachers**

<table>
<thead>
<tr>
<th>ITEM</th>
<th>CODE</th>
<th>DESCRIPTIVE</th>
<th>N</th>
<th>MEAN</th>
<th>VARIANCE</th>
<th>ITEM DISCRIMINATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>S</td>
<td>'6-8 MAKES CLASS WORK INTERESTING'</td>
<td>1623</td>
<td>3.43993</td>
<td>1.91367</td>
<td>31%</td>
</tr>
<tr>
<td>2</td>
<td>S</td>
<td>'6-8 TEACHER FAIR WITH ALL'</td>
<td>1625</td>
<td>3.71631</td>
<td>1.70844</td>
<td>25%</td>
</tr>
<tr>
<td>3</td>
<td>SR</td>
<td>'6-8 TEACHER MAINTAINS DISCIPLINE'</td>
<td>1622</td>
<td>4.07152</td>
<td>1.25382</td>
<td>13%</td>
</tr>
<tr>
<td>4</td>
<td>S</td>
<td>'6-8 TEACHER WELL-PREPARED'</td>
<td>1622</td>
<td>4.17201</td>
<td>1.22627</td>
<td>23%</td>
</tr>
<tr>
<td>5</td>
<td>SR</td>
<td>'6-8 GIVES ASSIGNMENTS RELATED SUBJ'</td>
<td>1620</td>
<td>4.29136</td>
<td>1.26943</td>
<td>28%</td>
</tr>
<tr>
<td>6</td>
<td>SR</td>
<td>'6-8 DISCUSS/SUMMARIZE LESSON'</td>
<td>1620</td>
<td>4.68877</td>
<td>1.73025</td>
<td>24%</td>
</tr>
<tr>
<td>7</td>
<td>SR</td>
<td>'6-8 DISCUSSIONS ON TOPIC OF LESSON'</td>
<td>1623</td>
<td>4.02403</td>
<td>1.29887</td>
<td>26%</td>
</tr>
<tr>
<td>8</td>
<td>CC</td>
<td>'6-8 LIKES WHEN WE ASK QUESTIONS'</td>
<td>1607</td>
<td>3.75233</td>
<td>1.40337</td>
<td>18%</td>
</tr>
<tr>
<td>9</td>
<td>SR</td>
<td>'6-8 MORE TIME TO WORK THAN NEED'</td>
<td>1615</td>
<td>2.92384</td>
<td>1.65426</td>
<td>18%</td>
</tr>
<tr>
<td>10</td>
<td>CC</td>
<td>'6-8 EXPLAINS LESSON AND WHY DO IT'</td>
<td>1620</td>
<td>3.98642</td>
<td>1.68080</td>
<td>19%</td>
</tr>
<tr>
<td>11</td>
<td>SR</td>
<td>'6-8 ASK QUESTIONS/UNDERSTAND TAUGHT'</td>
<td>1621</td>
<td>3.90253</td>
<td>1.53276</td>
<td>23%</td>
</tr>
<tr>
<td>12</td>
<td>S</td>
<td>'6-8 EXPLAINS IDEAS EASY UNDERSTAND'</td>
<td>1618</td>
<td>3.67182</td>
<td>1.67413</td>
<td>27%</td>
</tr>
<tr>
<td>13</td>
<td>S</td>
<td>'6-8 LOOKS AT WORK SEE WE UNDERSTAND'</td>
<td>1621</td>
<td>3.98112</td>
<td>1.77087</td>
<td>15%</td>
</tr>
<tr>
<td>14</td>
<td>S</td>
<td>'6-8 ALLOWS MORE THAN OTHER TEACHERS'</td>
<td>1610</td>
<td>3.63727</td>
<td>1.69824</td>
<td>22%</td>
</tr>
<tr>
<td>15</td>
<td>SR</td>
<td>'6-8 WORK TO DO IF FINISH CLASS OVER'</td>
<td>1619</td>
<td>2.90488</td>
<td>1.86371</td>
<td>15%</td>
</tr>
<tr>
<td>16</td>
<td>S</td>
<td>'6-8 MAKES MATERIALS/WORKSHEETS USE'</td>
<td>1618</td>
<td>3.64400</td>
<td>1.64706</td>
<td>21%</td>
</tr>
<tr>
<td>17</td>
<td>S</td>
<td>'6-8 GIVES TESTS AND QUIZES'</td>
<td>1617</td>
<td>3.98949</td>
<td>1.56804</td>
<td>36%</td>
</tr>
<tr>
<td>18</td>
<td>S</td>
<td>'6-8 RETURNS TESTS/ASSIGNMENTS QUICK'</td>
<td>1619</td>
<td>3.44232</td>
<td>1.72704</td>
<td>24%</td>
</tr>
<tr>
<td>19</td>
<td>S</td>
<td>'6-8 VARIETY OF ACTIVITIES/Resources'</td>
<td>1616</td>
<td>3.58136</td>
<td>1.69021</td>
<td>22%</td>
</tr>
<tr>
<td>20</td>
<td>S</td>
<td>'6-8 ENOUGH TIME TO DO OUR WORK'</td>
<td>1693</td>
<td>3.55744</td>
<td>1.82737</td>
<td>17%</td>
</tr>
</tbody>
</table>

**CRONBACH ALPHA RELIABILITY BASED ON 20 ITEMS WITH DISCRIMINATION >= 13% IS 0.935**

**ITEM DISCRIMINATION** = A question is considered to be most effective when it has a high level of item discrimination. This means that the question is useful in separating high teacher performance from that of average and low performance. A percentage value of at least 13% equals discrimination (at the .05 level of significance) and a percentage value of at least 22% equals discrimination at the .01 level of significance.

**DESCRIPTIVE** = an abbreviated label used to identify the specific student rating question.

**CODE** = S = original question (known to be a valid and reliable discriminating item) from the School Improvement Model (SIM).

**SR** = revised question from the original SIM listing of discriminating items.

**CC** = new questions developed by Cave Creek Schools.

**N** = total number of students who completed a teacher rating questionnaire.

**MEAN** = the arithmetic average of the total student responses for each item.

**VARIANCE** = the relationship of scores to a central value, such as the mean. Variance is defined as the sum of squared deviations around the mean. This statistic describes how similar or different the scores are, from the mean, for a given group.
## 1990 Cave Creek Student Ratings of Teachers

**Analysis based on 517 subjects in 21 groups.**

<table>
<thead>
<tr>
<th>Item</th>
<th>Code</th>
<th>Descriptive</th>
<th>N</th>
<th>Mean</th>
<th>Variance</th>
<th>Item Discrimination</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>S</td>
<td>'F-8 Teacher makes interesting'</td>
<td>516</td>
<td>3.22948</td>
<td>2.02910</td>
<td>26%</td>
</tr>
<tr>
<td>2</td>
<td>S</td>
<td>'F-8 Teacher fair with all'</td>
<td>517</td>
<td>3.30948</td>
<td>1.94678</td>
<td>16%</td>
</tr>
<tr>
<td>3</td>
<td>SR</td>
<td>'F-8 Teacher maintains discipline'</td>
<td>517</td>
<td>3.56673</td>
<td>1.85870</td>
<td>13%</td>
</tr>
<tr>
<td>4</td>
<td>S</td>
<td>'F-8 Teacher well-prepared'</td>
<td>514</td>
<td>3.75875</td>
<td>1.80172</td>
<td>21%</td>
</tr>
<tr>
<td>5</td>
<td>SR</td>
<td>'F-8 Gives assignments related subj'</td>
<td>512</td>
<td>3.43945</td>
<td>2.54712</td>
<td>11%</td>
</tr>
<tr>
<td>6</td>
<td>SR</td>
<td>'F-8 Discuss/Summarize lesson'</td>
<td>515</td>
<td>3.23301</td>
<td>2.10493</td>
<td>19%</td>
</tr>
<tr>
<td>7</td>
<td>SR</td>
<td>'F-8 Discussions on topic of lesson'</td>
<td>509</td>
<td>3.53438</td>
<td>1.96984</td>
<td>22%</td>
</tr>
<tr>
<td>8</td>
<td>CC</td>
<td>'F-8 Likes when we ask questions'</td>
<td>511</td>
<td>3.34834</td>
<td>1.80430</td>
<td>18%</td>
</tr>
<tr>
<td>9</td>
<td>CC</td>
<td>'F-8 Explains rules for behavior'</td>
<td>512</td>
<td>3.46289</td>
<td>2.01034</td>
<td>20%</td>
</tr>
<tr>
<td>10</td>
<td>CC</td>
<td>'F-8 Explains lesson and why do it'</td>
<td>514</td>
<td>3.50584</td>
<td>1.98927</td>
<td>21%</td>
</tr>
<tr>
<td>11</td>
<td>SR</td>
<td>'F-8 Ask questions/understand taught'</td>
<td>517</td>
<td>3.37718</td>
<td>2.07244</td>
<td>17%</td>
</tr>
<tr>
<td>12</td>
<td>S</td>
<td>'F-8 Explains ideas easy understand'</td>
<td>513</td>
<td>3.33138</td>
<td>1.91358</td>
<td>16%</td>
</tr>
<tr>
<td>13</td>
<td>S</td>
<td>'F-8 Looks at work see we understand'</td>
<td>512</td>
<td>3.25000</td>
<td>2.27734</td>
<td>15%</td>
</tr>
<tr>
<td>14</td>
<td>S</td>
<td>'F-8 Do same thing in class every day'</td>
<td>514</td>
<td>3.14008</td>
<td>2.14380</td>
<td>5%</td>
</tr>
<tr>
<td>15</td>
<td>CC</td>
<td>'F-8 Teacher easy to understand'</td>
<td>515</td>
<td>3.36505</td>
<td>2.04150</td>
<td>19%</td>
</tr>
<tr>
<td>16</td>
<td>CC</td>
<td>'F-8 Tests are fair'</td>
<td>516</td>
<td>3.40871</td>
<td>2.13705</td>
<td>23%</td>
</tr>
<tr>
<td>17</td>
<td>SR</td>
<td>'F-8 Gives tests and quizzes'</td>
<td>516</td>
<td>3.46899</td>
<td>1.87695</td>
<td>11%</td>
</tr>
<tr>
<td>18</td>
<td>S</td>
<td>'F-8 Returns tests/assignments quick'</td>
<td>516</td>
<td>3.21705</td>
<td>2.06527</td>
<td>11%</td>
</tr>
<tr>
<td>19</td>
<td>SR</td>
<td>'F-8 Variety of activities/resources'</td>
<td>514</td>
<td>3.50778</td>
<td>1.99702</td>
<td>23%</td>
</tr>
<tr>
<td>20</td>
<td>SR</td>
<td>'F-8 Expects best work I can'</td>
<td>503</td>
<td>3.78728</td>
<td>2.19928</td>
<td>21%</td>
</tr>
</tbody>
</table>

**Cronbach Alpha Reliability**

Based on 16 items with discrimination >= 13% is 0.950

**Descriptive**

- an abbreviated label used to identify the specific student rating question.

**Code**

- S = original question (known to be a valid and reliable discriminating item) from the School Improvement Model (SIM).
- SR = revised question from the original SIM listing of discriminating items.
- CC = new questions developed by Cave Creek Schools

**N**

- total number of students who completed a teacher rating questionnaire.

**Mean**

- the arithmetic average of the total student responses for each item.

**Variance**

- the relationship of scores to a central value, such as the mean. Variance is defined as the sum of squared deviations around the mean. This statistic describes how similar or different the scores are, from the mean, for a given group.

**Item Discrimination**

- A question is considered to be most effective when it has a high level of item discrimination. This means that the question is useful in separating high teacher performance from that of average and low performance. A percentage value of at least 13% equals discrimination (at the .05 level of significance) and a percentage value of at least 22% equals discrimination at the .01 level of significance.
1990 CAVE CREEK STUDENT RATINGS OF TEACHERS

ANALYSIS BASED ON 2551 SUBJECTS IN 139 GROUPS.

<table>
<thead>
<tr>
<th>ITEM</th>
<th>CODE</th>
<th>DESCRIPTIVE</th>
<th>N</th>
<th>MEAN</th>
<th>VARIANCE</th>
<th>DISCRIMINATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>S</td>
<td>'9-12 MAKES CLASS WORK INTERESTING'</td>
<td>2496</td>
<td>3.30329</td>
<td>1.91563</td>
<td>41%</td>
</tr>
<tr>
<td>2</td>
<td>S</td>
<td>'9-12 ASK QUESTIONS UNDERSTAND TAUGHT'</td>
<td>2497</td>
<td>3.71005</td>
<td>1.64280</td>
<td>33%</td>
</tr>
<tr>
<td>3</td>
<td>SR</td>
<td>'9-12 ASSIGNMENTS RELATED TO SUBJECT'</td>
<td>2473</td>
<td>4.01001</td>
<td>1.61574</td>
<td>37%</td>
</tr>
<tr>
<td>4</td>
<td>SR</td>
<td>'9-12 DISCUSS/SUMMARIZE EACH LESSON'</td>
<td>2490</td>
<td>3.70602</td>
<td>1.64290</td>
<td>31%</td>
</tr>
<tr>
<td>5</td>
<td>SR</td>
<td>'9-12 TELLS WHAT LEARNED TO LEARN NEW'</td>
<td>2499</td>
<td>4.39095</td>
<td>1.87493</td>
<td>31%</td>
</tr>
<tr>
<td>6</td>
<td>SR</td>
<td>'9-12 TEACHER MAINTAINS DISCIPLINE'</td>
<td>2495</td>
<td>3.78277</td>
<td>1.77325</td>
<td>40%</td>
</tr>
<tr>
<td>7</td>
<td>SR</td>
<td>'9-12 RETURNS TEST/ASSIGNMENTS QUICKLY'</td>
<td>2482</td>
<td>3.37631</td>
<td>1.93736</td>
<td>32%</td>
</tr>
<tr>
<td>8</td>
<td>SR</td>
<td>'9-12 GIVES FEEDBACK ABOUT PERFORMANCE'</td>
<td>2487</td>
<td>3.47165</td>
<td>1.76106</td>
<td>26%</td>
</tr>
<tr>
<td>9</td>
<td>S</td>
<td>'9-12 KNOWS MORE THAN OTHER TEACHERS'</td>
<td>2485</td>
<td>3.78551</td>
<td>1.79941</td>
<td>33%</td>
</tr>
<tr>
<td>10</td>
<td>S</td>
<td>'9-12 HOMEWORK HELPS ME LEARN'</td>
<td>2456</td>
<td>3.59731</td>
<td>1.88141</td>
<td>31%</td>
</tr>
<tr>
<td>11</td>
<td>SR</td>
<td>'9-12 MAKES MATERIALS/WORKSHEETS'</td>
<td>2470</td>
<td>3.61012</td>
<td>1.86597</td>
<td>36%</td>
</tr>
<tr>
<td>12</td>
<td>SR</td>
<td>'9-12 USE VARIETY ACTIVITIES/RESOURCES'</td>
<td>2452</td>
<td>3.42863</td>
<td>1.90478</td>
<td>36%</td>
</tr>
<tr>
<td>13</td>
<td>SR</td>
<td>'9-12 FILMS/VIDEOTAPES HELP US LEARN'</td>
<td>2405</td>
<td>3.19293</td>
<td>2.36527</td>
<td>35%</td>
</tr>
<tr>
<td>14</td>
<td>SR</td>
<td>'9-12 TELLS LIBRARY/MEDIA MATERIALS'</td>
<td>2436</td>
<td>3.03777</td>
<td>2.36311</td>
<td>35%</td>
</tr>
<tr>
<td>15</td>
<td>S</td>
<td>'9-12 TEACHER WELL ORGANIZED'</td>
<td>2481</td>
<td>3.60419</td>
<td>1.53608</td>
<td>37%</td>
</tr>
<tr>
<td>16</td>
<td>CC</td>
<td>'9-12 TEACHER LIKES WHEN ASK QUESTIONS'</td>
<td>2462</td>
<td>3.72653</td>
<td>1.75153</td>
<td>34%</td>
</tr>
<tr>
<td>17</td>
<td>SR</td>
<td>'9-12 DIFFERENT GROUPS DEPEND ACTIVITY'</td>
<td>2468</td>
<td>3.32455</td>
<td>2.14304</td>
<td>35%</td>
</tr>
<tr>
<td>18</td>
<td>S</td>
<td>'9-12 LOOK PROBLEMS/NEW WAYS TO SOLVE'</td>
<td>2462</td>
<td>3.33712</td>
<td>1.99195</td>
<td>33%</td>
</tr>
<tr>
<td>19</td>
<td>CC</td>
<td>'9-12 AVAILABLE CLASS/OTHER TIMES'</td>
<td>2472</td>
<td>2.70024</td>
<td>1.81184</td>
<td>37%</td>
</tr>
<tr>
<td>20</td>
<td>S</td>
<td>'9-12 LOOKS AT WORK/SEE IF UNDERSTAND'</td>
<td>2427</td>
<td>3.84508</td>
<td>1.84497</td>
<td>32%</td>
</tr>
</tbody>
</table>

CRONBACH ALPHA RELIABILITY BASED ON 20 ITEMS WITH DISCRIMINATION >= 13% IS 0.971

DESCRIPTIVE = an abbreviated label used to identify the specific student rating question.
CODE = S = original question (known to be a valid and reliable discriminating item) from the School Improvement Model (SIM).
SR= revised question from the original SIM listing of discriminating items.
CC= new questions developed by Cave Creek Schools
N = total number of students who completed a teacher rating questionnaire.
MEAN = the arithmetic average of the total student responses for each item.
VARIANCE = the relationship of scores to a central value, such as the mean. Variance is defined as the sum of squared deviations around the mean. This statistic describes how similar or different the scores are, from the mean, for a given group.
ITEM DISCRIMINATION = A question is considered to be most effective when it has a high level of item discrimination. This means that the question is useful in separating high teacher performance from that of average and low performance. A percentage value of at least 13% equals discrimination (at the .05 level of significance) and a percentage value of at least 22% equals discrimination at the .01 level of significance.
APPENDIX O.

1991 CAVE CREEK STUDENT RATINGS OF TEACHERS
ANALYSIS AND SUMMARY
<table>
<thead>
<tr>
<th>ITEM</th>
<th>N</th>
<th>MEAN</th>
<th>VARIANCE</th>
<th>SS TOTAL</th>
<th>SS WITHIN</th>
<th>SS BETWEEN</th>
<th>ITEM DISCRIMINATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>797</td>
<td>1.50314</td>
<td>0.40808</td>
<td>325.24216</td>
<td>285.49896</td>
<td>39.74320</td>
<td>12%</td>
</tr>
<tr>
<td>2</td>
<td>789</td>
<td>1.36795</td>
<td>0.34653</td>
<td>273.40938</td>
<td>248.86222</td>
<td>24.54716</td>
<td>9%</td>
</tr>
<tr>
<td>3</td>
<td>798</td>
<td>1.16416</td>
<td>0.19977</td>
<td>127.49499</td>
<td>118.77443</td>
<td>8.71956</td>
<td>7%</td>
</tr>
<tr>
<td>4</td>
<td>792</td>
<td>1.27525</td>
<td>0.27525</td>
<td>217.99495</td>
<td>195.16486</td>
<td>22.83016</td>
<td>9%</td>
</tr>
<tr>
<td>5</td>
<td>796</td>
<td>2.36307</td>
<td>0.57296</td>
<td>415.07412</td>
<td>409.34329</td>
<td>5.73092</td>
<td>9%</td>
</tr>
<tr>
<td>6</td>
<td>798</td>
<td>1.71679</td>
<td>0.64912</td>
<td>517.99499</td>
<td>415.16486</td>
<td>102.83016</td>
<td>32%</td>
</tr>
<tr>
<td>7</td>
<td>798</td>
<td>1.25063</td>
<td>0.25548</td>
<td>203.87469</td>
<td>184.56343</td>
<td>19.31126</td>
<td>9%</td>
</tr>
<tr>
<td>8</td>
<td>785</td>
<td>1.13758</td>
<td>0.18744</td>
<td>147.14140</td>
<td>140.39221</td>
<td>6.74920</td>
<td>5%</td>
</tr>
<tr>
<td>9</td>
<td>777</td>
<td>1.76286</td>
<td>0.52720</td>
<td>420.18068</td>
<td>299.26182</td>
<td>120.91886</td>
<td>22%</td>
</tr>
<tr>
<td>10</td>
<td>798</td>
<td>1.46992</td>
<td>0.64759</td>
<td>516.77820</td>
<td>468.07078</td>
<td>48.70742</td>
<td>9%</td>
</tr>
<tr>
<td>11</td>
<td>794</td>
<td>1.28212</td>
<td>0.37129</td>
<td>294.80605</td>
<td>273.24491</td>
<td>21.56114</td>
<td>7%</td>
</tr>
<tr>
<td>12</td>
<td>794</td>
<td>1.27230</td>
<td>0.26158</td>
<td>207.69395</td>
<td>190.95410</td>
<td>16.73985</td>
<td>9%</td>
</tr>
<tr>
<td>13</td>
<td>793</td>
<td>1.36444</td>
<td>0.35773</td>
<td>283.67718</td>
<td>256.13894</td>
<td>27.53824</td>
<td>10%</td>
</tr>
<tr>
<td>14</td>
<td>795</td>
<td>1.39371</td>
<td>0.37958</td>
<td>301.76855</td>
<td>264.89866</td>
<td>36.86989</td>
<td>12%</td>
</tr>
<tr>
<td>15</td>
<td>793</td>
<td>1.18537</td>
<td>0.20649</td>
<td>163.75032</td>
<td>152.00037</td>
<td>11.74995</td>
<td>7%</td>
</tr>
<tr>
<td>16</td>
<td>798</td>
<td>1.44486</td>
<td>0.43242</td>
<td>345.07393</td>
<td>312.46462</td>
<td>32.60911</td>
<td>9%</td>
</tr>
<tr>
<td>17</td>
<td>795</td>
<td>1.52453</td>
<td>0.45569</td>
<td>362.27170</td>
<td>312.24517</td>
<td>50.02653</td>
<td>14%</td>
</tr>
<tr>
<td>18</td>
<td>796</td>
<td>1.35343</td>
<td>0.38121</td>
<td>303.44095</td>
<td>274.69478</td>
<td>28.74617</td>
<td>9%</td>
</tr>
<tr>
<td>19</td>
<td>796</td>
<td>1.67085</td>
<td>0.53739</td>
<td>427.76382</td>
<td>357.72913</td>
<td>70.03469</td>
<td>16%</td>
</tr>
<tr>
<td>20</td>
<td>793</td>
<td>1.27869</td>
<td>0.28173</td>
<td>223.40984</td>
<td>195.58040</td>
<td>27.81800</td>
<td>12%</td>
</tr>
</tbody>
</table>

Cronbach alpha reliability based on 4 items with discrimination >= 13% is 0.739
### Analysis Based on 1497 Subjects in 55 Groups

<table>
<thead>
<tr>
<th>ITEM</th>
<th>N</th>
<th>MEAN</th>
<th>VARIANCE</th>
<th>SS TOTAL</th>
<th>SS WITHIN</th>
<th>SS BETWEEN</th>
<th>ITEM DISCRIMINATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1491</td>
<td>3.84306</td>
<td>1.34894</td>
<td>2011.27565</td>
<td>1533.29414</td>
<td>477.98151</td>
<td>24%</td>
</tr>
<tr>
<td>2</td>
<td>1494</td>
<td>3.65328</td>
<td>1.42999</td>
<td>2136.39893</td>
<td>1746.11325</td>
<td>390.28568</td>
<td>18%</td>
</tr>
<tr>
<td>3</td>
<td>1489</td>
<td>3.68905</td>
<td>1.78847</td>
<td>2663.03156</td>
<td>2155.98734</td>
<td>507.04423</td>
<td>17%</td>
</tr>
<tr>
<td>4</td>
<td>1489</td>
<td>3.59033</td>
<td>1.98664</td>
<td>2958.10074</td>
<td>1651.26589</td>
<td>1306.81485</td>
<td>44%</td>
</tr>
<tr>
<td>5</td>
<td>1490</td>
<td>4.18289</td>
<td>1.07894</td>
<td>1607.61342</td>
<td>1429.40028</td>
<td>178.21314</td>
<td>11%</td>
</tr>
<tr>
<td>6</td>
<td>1488</td>
<td>3.49866</td>
<td>1.78494</td>
<td>2655.99731</td>
<td>2110.49819</td>
<td>545.49913</td>
<td>21%</td>
</tr>
<tr>
<td>7</td>
<td>1490</td>
<td>4.05034</td>
<td>1.40753</td>
<td>2097.22483</td>
<td>1721.91137</td>
<td>375.31346</td>
<td>18%</td>
</tr>
<tr>
<td>8</td>
<td>1487</td>
<td>4.20377</td>
<td>1.16964</td>
<td>1739.25891</td>
<td>1354.00752</td>
<td>385.25139</td>
<td>22%</td>
</tr>
<tr>
<td>9</td>
<td>1481</td>
<td>3.77245</td>
<td>1.38036</td>
<td>2044.31600</td>
<td>1758.64228</td>
<td>285.67373</td>
<td>14%</td>
</tr>
<tr>
<td>10</td>
<td>1486</td>
<td>3.30754</td>
<td>2.01107</td>
<td>2988.45599</td>
<td>2449.23471</td>
<td>539.22087</td>
<td>19%</td>
</tr>
<tr>
<td>11</td>
<td>1487</td>
<td>3.83457</td>
<td>1.53282</td>
<td>2279.30330</td>
<td>1837.20930</td>
<td>442.09400</td>
<td>19%</td>
</tr>
<tr>
<td>12</td>
<td>1485</td>
<td>4.25522</td>
<td>1.03587</td>
<td>1538.27205</td>
<td>1180.62782</td>
<td>357.64424</td>
<td>23%</td>
</tr>
<tr>
<td>13</td>
<td>1490</td>
<td>4.13020</td>
<td>1.48372</td>
<td>2210.74094</td>
<td>1710.96009</td>
<td>499.78085</td>
<td>23%</td>
</tr>
<tr>
<td>14</td>
<td>1489</td>
<td>4.10678</td>
<td>1.21157</td>
<td>1804.02149</td>
<td>1404.98985</td>
<td>399.03164</td>
<td>22%</td>
</tr>
<tr>
<td>15</td>
<td>1489</td>
<td>4.05104</td>
<td>1.38087</td>
<td>2056.12089</td>
<td>1671.92221</td>
<td>384.19668</td>
<td>19%</td>
</tr>
<tr>
<td>16</td>
<td>1478</td>
<td>3.84844</td>
<td>1.40070</td>
<td>2082.05142</td>
<td>1729.50129</td>
<td>352.55013</td>
<td>17%</td>
</tr>
<tr>
<td>17</td>
<td>1480</td>
<td>4.09797</td>
<td>1.19783</td>
<td>1772.79392</td>
<td>1370.11667</td>
<td>402.67725</td>
<td>23%</td>
</tr>
<tr>
<td>18</td>
<td>1476</td>
<td>3.88279</td>
<td>1.49442</td>
<td>2208.72290</td>
<td>1764.59126</td>
<td>444.13174</td>
<td>20%</td>
</tr>
<tr>
<td>19</td>
<td>1468</td>
<td>3.93188</td>
<td>1.44768</td>
<td>2125.18801</td>
<td>1775.67643</td>
<td>349.51158</td>
<td>16%</td>
</tr>
<tr>
<td>20</td>
<td>1413</td>
<td>4.35173</td>
<td>1.15937</td>
<td>1638.18825</td>
<td>1241.54437</td>
<td>396.64386</td>
<td>24%</td>
</tr>
</tbody>
</table>

**Cronbach Alpha Reliability Based on 19 Items with Discrimination >= 13% is 0.937**
OMATANI DAMS DATA 3 MENNE-TOLSMA SPRING 1991

ANALYSIS BASED ON 2159 SUBJECTS IN 98 GROUPS

<table>
<thead>
<tr>
<th>ITEM</th>
<th>N</th>
<th>MEAN</th>
<th>VARIANCE</th>
<th>SS TOTAL</th>
<th>SS WITHIN</th>
<th>SS BETWEEN</th>
<th>ITEM DISCRIMINATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2146</td>
<td>3.56151</td>
<td>1.71406</td>
<td>3678.38071</td>
<td>2740.85642</td>
<td>937.52429</td>
<td>25%</td>
</tr>
<tr>
<td>2</td>
<td>2146</td>
<td>3.81407</td>
<td>1.54092</td>
<td>3306.81500</td>
<td>2690.55166</td>
<td>616.26335</td>
<td>19%</td>
</tr>
<tr>
<td>3</td>
<td>2148</td>
<td>4.14851</td>
<td>1.23726</td>
<td>2657.62523</td>
<td>2147.69346</td>
<td>507.92978</td>
<td>19%</td>
</tr>
<tr>
<td>4</td>
<td>2142</td>
<td>4.16760</td>
<td>1.20487</td>
<td>2580.83147</td>
<td>2124.69213</td>
<td>456.13933</td>
<td>18%</td>
</tr>
<tr>
<td>5</td>
<td>2123</td>
<td>4.23457</td>
<td>1.36372</td>
<td>2895.16229</td>
<td>2194.98562</td>
<td>700.19667</td>
<td>24%</td>
</tr>
<tr>
<td>6</td>
<td>2139</td>
<td>3.75877</td>
<td>1.63323</td>
<td>3493.52314</td>
<td>2832.16471</td>
<td>661.35843</td>
<td>21%</td>
</tr>
<tr>
<td>7</td>
<td>2134</td>
<td>4.04405</td>
<td>1.30171</td>
<td>2777.85942</td>
<td>2306.87651</td>
<td>470.98091</td>
<td>17%</td>
</tr>
<tr>
<td>8</td>
<td>2120</td>
<td>3.80283</td>
<td>1.46961</td>
<td>3115.58302</td>
<td>2651.25601</td>
<td>464.32701</td>
<td>15%</td>
</tr>
<tr>
<td>9</td>
<td>2122</td>
<td>2.95335</td>
<td>1.98133</td>
<td>4204.36124</td>
<td>3406.35864</td>
<td>798.02261</td>
<td>12%</td>
</tr>
<tr>
<td>10</td>
<td>2139</td>
<td>3.64843</td>
<td>1.70903</td>
<td>3655.62225</td>
<td>2959.47896</td>
<td>696.14330</td>
<td>19%</td>
</tr>
<tr>
<td>11</td>
<td>2144</td>
<td>3.88117</td>
<td>1.57255</td>
<td>3374.69944</td>
<td>2752.79603</td>
<td>621.90341</td>
<td>18%</td>
</tr>
<tr>
<td>12</td>
<td>2132</td>
<td>3.67999</td>
<td>1.62056</td>
<td>3455.04081</td>
<td>2776.98034</td>
<td>678.06047</td>
<td>20%</td>
</tr>
<tr>
<td>13</td>
<td>2136</td>
<td>3.68773</td>
<td>1.79434</td>
<td>3832.71863</td>
<td>2907.71533</td>
<td>925.00310</td>
<td>24%</td>
</tr>
<tr>
<td>14</td>
<td>2122</td>
<td>3.85297</td>
<td>1.57762</td>
<td>3348.12630</td>
<td>2729.59360</td>
<td>618.53269</td>
<td>18%</td>
</tr>
<tr>
<td>15</td>
<td>2125</td>
<td>3.15341</td>
<td>1.98964</td>
<td>4227.98776</td>
<td>3321.46962</td>
<td>906.51814</td>
<td>21%</td>
</tr>
<tr>
<td>16</td>
<td>2123</td>
<td>3.56288</td>
<td>1.76183</td>
<td>3740.35516</td>
<td>2857.61667</td>
<td>882.73849</td>
<td>24%</td>
</tr>
<tr>
<td>17</td>
<td>2122</td>
<td>3.97926</td>
<td>1.56036</td>
<td>3311.08765</td>
<td>2193.70664</td>
<td>1117.38102</td>
<td>34%</td>
</tr>
<tr>
<td>18</td>
<td>2113</td>
<td>3.52769</td>
<td>1.83560</td>
<td>3878.63038</td>
<td>2974.81895</td>
<td>903.81143</td>
<td>23%</td>
</tr>
<tr>
<td>19</td>
<td>2123</td>
<td>3.70089</td>
<td>1.68020</td>
<td>3567.06830</td>
<td>2826.88075</td>
<td>740.18755</td>
<td>21%</td>
</tr>
<tr>
<td>20</td>
<td>2080</td>
<td>3.65240</td>
<td>1.93447</td>
<td>4023.68798</td>
<td>3320.07040</td>
<td>703.61758</td>
<td>17%</td>
</tr>
</tbody>
</table>

CRONBACH ALPHA RELIABILITY BASED ON 20 ITEMS WITH DISCRIMINATION >= 13% IS 0.948
<table>
<thead>
<tr>
<th>ITEM</th>
<th>N</th>
<th>MEAN</th>
<th>VARIANCE</th>
<th>SS TOTAL</th>
<th>SS WITHIN</th>
<th>SS BETWEEN</th>
<th>ITEM DISCRIMINATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2071</td>
<td>3.53066</td>
<td>1.53443</td>
<td>3177.80299</td>
<td>2321.27607</td>
<td>856.52692</td>
<td>27%</td>
</tr>
<tr>
<td>2</td>
<td>2073</td>
<td>4.09793</td>
<td>1.12452</td>
<td>2331.12108</td>
<td>1964.86525</td>
<td>366.25583</td>
<td>16%</td>
</tr>
<tr>
<td>3</td>
<td>2064</td>
<td>4.43798</td>
<td>0.90201</td>
<td>1872.02620</td>
<td>1613.98252</td>
<td>258.03769</td>
<td>14%</td>
</tr>
<tr>
<td>4</td>
<td>2063</td>
<td>4.05332</td>
<td>1.15663</td>
<td>2386.13476</td>
<td>2092.00903</td>
<td>294.12573</td>
<td>13%</td>
</tr>
<tr>
<td>5</td>
<td>2065</td>
<td>3.65860</td>
<td>1.55463</td>
<td>3210.30953</td>
<td>2591.89868</td>
<td>618.40385</td>
<td>19%</td>
</tr>
<tr>
<td>6</td>
<td>2068</td>
<td>4.13540</td>
<td>1.05420</td>
<td>2180.98977</td>
<td>1883.79565</td>
<td>297.29412</td>
<td>14%</td>
</tr>
<tr>
<td>7</td>
<td>2066</td>
<td>3.62246</td>
<td>1.59706</td>
<td>3299.51791</td>
<td>2700.99375</td>
<td>598.52415</td>
<td>18%</td>
</tr>
<tr>
<td>8</td>
<td>2064</td>
<td>3.78052</td>
<td>1.41355</td>
<td>2917.57703</td>
<td>2544.67904</td>
<td>372.89800</td>
<td>13%</td>
</tr>
<tr>
<td>9</td>
<td>2035</td>
<td>4.05207</td>
<td>1.34569</td>
<td>2736.47862</td>
<td>2172.59273</td>
<td>565.88589</td>
<td>21%</td>
</tr>
<tr>
<td>10</td>
<td>2037</td>
<td>3.87923</td>
<td>1.52984</td>
<td>3116.29161</td>
<td>2662.10431</td>
<td>454.18729</td>
<td>15%</td>
</tr>
<tr>
<td>11</td>
<td>2050</td>
<td>3.94439</td>
<td>1.37252</td>
<td>2813.66049</td>
<td>2435.34437</td>
<td>378.31612</td>
<td>13%</td>
</tr>
<tr>
<td>12</td>
<td>2039</td>
<td>3.66797</td>
<td>1.59010</td>
<td>3242.21873</td>
<td>2482.54649</td>
<td>759.67225</td>
<td>23%</td>
</tr>
<tr>
<td>13</td>
<td>1948</td>
<td>3.52669</td>
<td>2.30062</td>
<td>4481.61191</td>
<td>3050.62242</td>
<td>1430.98948</td>
<td>32%</td>
</tr>
<tr>
<td>14</td>
<td>1980</td>
<td>3.15556</td>
<td>2.23540</td>
<td>4426.08889</td>
<td>3359.27569</td>
<td>1066.81320</td>
<td>24%</td>
</tr>
<tr>
<td>15</td>
<td>2031</td>
<td>3.90205</td>
<td>1.30951</td>
<td>2688.80497</td>
<td>2270.61054</td>
<td>418.19443</td>
<td>15%</td>
</tr>
<tr>
<td>16</td>
<td>2034</td>
<td>4.04625</td>
<td>1.29338</td>
<td>2656.60613</td>
<td>2215.35973</td>
<td>441.24641</td>
<td>17%</td>
</tr>
<tr>
<td>17</td>
<td>2040</td>
<td>3.70588</td>
<td>1.73997</td>
<td>3549.52941</td>
<td>2761.88103</td>
<td>787.64838</td>
<td>22%</td>
</tr>
<tr>
<td>18</td>
<td>2040</td>
<td>3.72270</td>
<td>1.48146</td>
<td>3022.17559</td>
<td>2489.37112</td>
<td>532.80446</td>
<td>18%</td>
</tr>
<tr>
<td>19</td>
<td>2049</td>
<td>4.07565</td>
<td>1.24904</td>
<td>2559.27477</td>
<td>2129.23339</td>
<td>430.04137</td>
<td>17%</td>
</tr>
<tr>
<td>20</td>
<td>2028</td>
<td>4.05769</td>
<td>1.37685</td>
<td>2792.25000</td>
<td>2261.56894</td>
<td>530.68106</td>
<td>19%</td>
</tr>
</tbody>
</table>

Cronbach Alpha Reliability based on 19 Items with Discrimination >= 13% is 0.950