12-1972

Study of the Discrepancies between Student Evaluations and Faculty Self-Perceptions of Instructional Procedures in Higher Education

Eugene W. Thompson
Western Michigan University

Follow this and additional works at: https://scholarworks.wmich.edu/dissertations

Recommended Citation
Thompson, Eugene W., "Study of the Discrepancies between Student Evaluations and Faculty Self-Perceptions of Instructional Procedures in Higher Education" (1972). Dissertations. 2974.
https://scholarworks.wmich.edu/dissertations/2974

This Dissertation-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Dissertations by an authorized administrator of ScholarWorks at WMU. For more information, please contact maira.bundza@wmich.edu.
A Study Of The Discrepancies Between
Student Evaluations And Faculty
Self-Perceptions Of Instructional
Procedures In Higher Education

by

Eugene W. Thompson

A Dissertation
Submitted to the
Faculty of The Graduate College
in partial fulfillment
of the
Degree of Doctor of Education

College of Education
Western Michigan University
Kalamazoo, Michigan
December 1972
ACKNOWLEDGEMENTS

Without the assistance of many individuals this study would not have been possible. It gives a great deal of pleasure to recognize the following persons and groups:

The participating faculty of the Teacher Education Department, College of Education, at Western Michigan University.

Dr. Dorothy McCuskey who has given so freely of her time and advice while serving as Chairperson of my committee.

Dr. Kenneth Engle who first encouraged me to begin work on the doctorate, and who has been a close friend constantly displaying empathic human regard.

Dr. Uldis Smidchens who provided continuous research and statistical advice.

Dr. John Nangle whose advice, comments, and editorial assistance were invaluable.

Dr. Charles Townsend who provided countless hours of computer instruction and programming.

Finally, recognition is due my wife Violet, son Jeffrey Sean, and daughter Elizabeth Mae. Each in their own way provided love, support, and understanding, and in reality made this project possible.

Eugene W. Thompson
INFORMATION TO USERS

This dissertation was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in “sectioning” the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from “photographs” if essential to the understanding of the dissertation. Silver prints of “photographs” may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106
A Xerox Education Company

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
THOMPSON, Eugene Wesley, 1941-
A STUDY OF THE DISCREPANCIES BETWEEN
STUDENT EVALUATIONS AND FACULTY SELF-
PERCEPTIONS OF INSTRUCTIONAL PROCEDURES
IN HIGHER EDUCATION.

Western Michigan University, Ed.D., 1972
Education, higher

University Microfilms, A XEROX Company, Ann Arbor, Michigan
PLEASE NOTE:

Some pages may have
indistinct print.
Filmed as received.

University Microfilms, A Xerox Education Company
<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>TABLE OF CONTENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>NEED FOR THE STUDY ............................................. 1</td>
</tr>
<tr>
<td></td>
<td>Introduction .................................................. 1</td>
</tr>
<tr>
<td></td>
<td>Purpose of the Study .......................................... 2</td>
</tr>
<tr>
<td></td>
<td>Significance of the Study ..................................... 3</td>
</tr>
<tr>
<td></td>
<td>Definition of Terms ............................................ 4</td>
</tr>
<tr>
<td></td>
<td>Organization of the Dissertation ............................ 6</td>
</tr>
<tr>
<td>II</td>
<td>REVIEW OF SELECTED RELATED LITERATURE ........................ 7</td>
</tr>
<tr>
<td></td>
<td>Trends and Findings in the Uses of College Student Ratings of Instructors ........................................ 7</td>
</tr>
<tr>
<td></td>
<td>Student Feedback As a Basis For Instructor Behavior Modification ............................................. 16</td>
</tr>
<tr>
<td></td>
<td>Self-Perceptions and Self-Ratings ........................... 20</td>
</tr>
<tr>
<td></td>
<td>Summary .......................................................... 25</td>
</tr>
<tr>
<td>III</td>
<td>METHODS, DESIGN, PROCEDURES ................................... 28</td>
</tr>
<tr>
<td></td>
<td>Population and Sample ......................................... 29</td>
</tr>
<tr>
<td></td>
<td>Instrumentation ................................................ 30</td>
</tr>
<tr>
<td></td>
<td>Design ............................................................ 35</td>
</tr>
<tr>
<td></td>
<td>Hypotheses ....................................................... 36</td>
</tr>
<tr>
<td></td>
<td>General Procedures ............................................. 37</td>
</tr>
<tr>
<td></td>
<td>Statistical Analyses ........................................... 40</td>
</tr>
<tr>
<td></td>
<td>Summary .......................................................... 43</td>
</tr>
<tr>
<td>IV</td>
<td>RESULTS .......................................................... 46</td>
</tr>
<tr>
<td></td>
<td>Samples Included in the Analyses ............................. 46</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
TABLE OF CONTENTS (continued)

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relationships of Student Ratings and Instructor Self-Perceptions</td>
<td>48</td>
</tr>
<tr>
<td>Summary</td>
<td>65</td>
</tr>
<tr>
<td>V SUMMARY, CONCLUSIONS, AND IMPLICATIONS</td>
<td>68</td>
</tr>
<tr>
<td>Summary</td>
<td>68</td>
</tr>
<tr>
<td>Conclusions</td>
<td>70</td>
</tr>
<tr>
<td>Implications</td>
<td>74</td>
</tr>
<tr>
<td>Recommendations</td>
<td>75</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>77</td>
</tr>
<tr>
<td>APPENDIX A</td>
<td>83</td>
</tr>
<tr>
<td>APPENDIX B</td>
<td>89</td>
</tr>
<tr>
<td>APPENDIX C</td>
<td>95</td>
</tr>
<tr>
<td>APPENDIX D</td>
<td>97</td>
</tr>
<tr>
<td>APPENDIX E</td>
<td>99</td>
</tr>
<tr>
<td>APPENDIX F</td>
<td>101</td>
</tr>
<tr>
<td>APPENDIX G</td>
<td>104</td>
</tr>
<tr>
<td>APPENDIX H</td>
<td>107</td>
</tr>
<tr>
<td>TABLE</td>
<td>PAGE</td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>1</td>
<td>Characteristics of Instructors Providing Self-Perceptions About Instructional Procedures</td>
</tr>
<tr>
<td>2</td>
<td>Characteristics of Students Responding to the Student Opinions About Instructional Procedures Questionnaire</td>
</tr>
<tr>
<td>3</td>
<td>Relationship Between Student-Ratings and Instructor Self-Perceptions Regarding Professional Competence</td>
</tr>
<tr>
<td>4</td>
<td>Relationship Between Student Ratings and Instructor Self-Perceptions Regarding Evaluation Procedures</td>
</tr>
<tr>
<td>5</td>
<td>Relationship Between Student Ratings and Instructor Self-Perceptions Regarding Student Centeredness</td>
</tr>
<tr>
<td>6</td>
<td>Pearson Product-Moment Correlations Between Student and Instructor Factor Ratings</td>
</tr>
<tr>
<td>7</td>
<td>Relationship Between Student GPA and the Congruence of Student Ratings and Faculty Self-Perceptions Regarding Professional Competence</td>
</tr>
<tr>
<td>8</td>
<td>Relationship Between Student GPA and the Congruence of Student Ratings and Faculty Self-Perceptions Regarding Evaluation Procedures</td>
</tr>
<tr>
<td>9</td>
<td>Relationship Between Student GPA and the Congruence of Student Ratings and Faculty Self-Perceptions Regarding Student Centeredness</td>
</tr>
<tr>
<td>10</td>
<td>Relationship Between Class Size and the Congruence of Student Ratings with Faculty Self-Perceptions Regarding Professional Competence</td>
</tr>
<tr>
<td>11</td>
<td>Relationship Between Class Size and the Congruence of Student Ratings with Faculty Self-Perceptions Regarding Evaluation Procedures</td>
</tr>
<tr>
<td>TABLE</td>
<td>PAGE</td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>12</td>
<td>58</td>
</tr>
<tr>
<td>13</td>
<td>59</td>
</tr>
<tr>
<td>14</td>
<td>60</td>
</tr>
<tr>
<td>15</td>
<td>61</td>
</tr>
<tr>
<td>16</td>
<td>62</td>
</tr>
<tr>
<td>17</td>
<td>63</td>
</tr>
<tr>
<td>18</td>
<td>64</td>
</tr>
</tbody>
</table>
CHAPTER I

NEED FOR THE STUDY

Introduction

The evaluation of teaching performance has, in recent years, assumed a high level of priority in the minds of many of those concerned with the educational process, whether they be professionals or concerned laymen. In this project interest is directed toward classroom behaviors and instructional procedures of higher educational instructors.

Traditionally, university instructors have worked under the philosophy of academic freedom. That is, the activities of the classrooms, within ethical boundaries, were basically of concern only to themselves. The responsibility for evaluation and improvement of the instructional process, then, rested essentially with the individual instructor. Some, generally working independently, included in their personal evaluation a solicitation of opinions from the students in their classes.

Of late this procedure of utilizing student opinions has gained in favor to the point that many noted institutions now make systematic surveys of student opinions regarding instructional procedures. The source of the activity, however, has varied from place to place as has the preparation of a questionnaire. In some cases the entire procedure, including creation of the document, has been
carried out by students who lacked measurement expertise, leading one to suspect the intent and reliability of the process. In others, the process originated and developed among faculty groups, generally departmental committees.

Assuming that development of such an instrument has taken place under the direction of persons possessing behavioral measurement skills, it would seem that there are two groups which could provide a reliable response. These groups or persons are those who would have first-hand knowledge of classroom interactions. First, of course, are the students, and second the instructors. Others may have formed opinions regarding an instructor's classroom behaviors but these opinions will be based on rumor, hearsay, and only rarely on first-hand knowledge.

Purpose of the Study

The purpose of this study was to examine the degree to which higher education instructors' self-perceptions of their instructional behaviors and procedures would be congruent with their students' assessments of these same behaviors and procedures.

It was believed by the writer that an assessment of behaviors by individuals actually involved in the interaction represented one realistic approach to providing an evaluation tool which could provide for behavior change and teaching improvement. To meet this goal of a realistic model, 58 instructors in the Teacher Education Department, College of Education at Western
Michigan University were requested to respond to an instrument designed to provide a rating of their self-perceived classroom behaviors and procedures. Similarly, over 3,000 students of participating instructors responded to a questionnaire which differed only by pronoun change in items of their perceptions of the instructor's behaviors and procedures. At the same time that these ratings were collected various personal characteristics of each respondent were also identified.

The examination of the degree of congruence between students and instructors was made by dichotomizing several variables of both student and instructor characteristics. These characteristics, it was hypothesized, would have an impact on the evaluation process. The resulting analyses serve as an indicator of the degree of congruence of student and self-evaluations as they compared on the basis of varying characteristics.

Significance of the Study

The significance of this study lies in the fact that it illustrates a little used tool for the improvement of instruction in higher education. Utilizing the only truly valid opinions regarding this process provides for considerable more openness and trust on the part of the participants than if the assessments of those not involved in the interaction were included.

Continued innovation is apparent in the use of higher education faculty self-perceptions of their instructional procedures.
The only previous effort encountered in this general area dealt with a military education setting.

Pressures from the patrons of higher education make it almost mandatory that instructors move to provide data regarding their effectiveness. Current moves by legislative bodies to regulate the number of class hours taught by community college teachers in a reported effort to improve the quality of instruction are already spreading to the universities. Instructors must begin to demonstrate their efforts to improve both the quality of instruction and its effectiveness. This study, then, provides a model by which certain instructional procedures can be assessed and provision made for behavior modification.

Definition of Terms

The use of the following terms in this study is intended to carry the attached definitions:

1. **Questionnaire or instrument** -- Data were collected from students by means of the Student Opinions About Instructional Procedures (SOAIP) (Appendix A) a questionnaire explained more fully in Chapter III. Only minor pronoun modification was made to this instrument in order to adapt it to a means for collecting faculty data.

2. **Student ratings (opinions)** -- These ratings or opinions given by students were in regard to questions posed about instructor behaviors and instructional procedures as
3. **Instructor self-perceptions** -- These self-ratings are also in response to the behaviors and procedures identified in the version of the instrument modified for faculty use (Appendix B).

4. **Classroom behaviors and procedures** -- The instructional behaviors and procedures referred to in this effort are those identified as relevant by faculty and students in the development of the questionnaire and are reflected in that instrument.

5. **Achievement levels** -- These levels are those reported in the student response to the questionnaire and are in the form of cumulative grade point averages. Higher achievement is referred to as category 3 (3.0-4.0), while lower achievement is considered to be in categories 0, 1, and 2 (under 1.0 through 2.99).

6. **Class size** -- Class size is stratified according to the reported numbers of student enrollment. This study will consider 1-30 students to be a small class, while one of 31 or over is large.

7. **Student enrollment status** -- This status for a course is reported by the student as being either elective or required.

8. **Teaching experience** -- These data are reported by the instructor, and refer to the number of years that faculty
member has taught in higher education. A dichotomy has been established with "less experience" being category 0 (less than five years), and "more" category 1 (five years or more).

Organization of the Dissertation

The intent of Chapter I has been to present an overview of the study through a statement of the problem, review of the significance and purposes of the study, definitions of terms, and presentation of the organization of the dissertation. Chapter II contains the report of Rationale and Related Literature. Within this report may be found statements of previous studies and writings dealing with student evaluations of instructors, various investigations into self-perception, and means by which behavior is modified. Chapter III, Design of the Study, is a discussion of the procedures utilized, hypotheses and questions explored, sources of data, methods of data collection, instrumentation, and the methods of analyzing data. Chapter IV contains a report of the findings, and Chapter V a summary of the study, and a review of the findings, conclusions, and implications.
CHAPTER II

Review of Selected Related Literature

This review of the literature will survey selected previous efforts that are related to this study in the following areas:
(a) trends and findings in the uses of college student ratings of instructors, (b) uses of student feedback as a basis for instructor behavior modification, and (c) self-perceptions and self-ratings.

Trends and Findings In The Uses of College Student Ratings of Instructors

Eble (1970 p. 17) in a detailed report of a project designed to improve college teaching, lends considerable support to the uses of student evaluation processes for instructors. In support of this type of evaluation he cites the following arguments: (1) the chances are increased that excellence in teaching will be recognized, (2) greater student-teacher interaction may result, (3) the institution may consider its overall goals in light of this evaluation of teaching, (4) provision is thus made for the only direct information about faculty teaching, (5) a tangible sign is displayed by faculty of the need for student involvement in goal setting.

French (1957) found ten characteristics which were the high contributors to student judgments at the University of Washington.
They were as follows: (1) interprets abstract ideas clearly, (2) gets students interested, (3) has increased my skills in thinking, (4) has helped broaden my interests, (5) stresses important material, (6) makes good use of examples, (7) motivated me to do my best work, (8) inspires class confidence in his knowledge of the subject, (9) has given me new viewpoints, and (10) is clear and understandable in his explanations.

Similarly, Gadzella (1968) asked a randomly selected sample of students to indicate their views on an "ideal professor." Criteria listed as most important were knowledge of subject, interest in subject, flexibility, and preparation.

At the University of Michigan McKeachie (1969) attempted to determine the criteria by which students make their judgments. Utilizing student scores on a measurement of various aspects of "critical thinking," he and his colleagues found that those students who performed well on the test rated their teachers as "more effective" than did students who performed poorly. It is interesting to note, the items on which the "effective" teachers rated higher included the usual clear explanations, stimulating students' curiosity, and interesting presentations, as well as attention to student reactions, friendliness, permissiveness, and flexibility.

In an early study, Guthrie (1949) attempted to determine the correlation between student ratings and "faculty-jury" ratings of instructors using a nine item questionnaire dealing with general
professional contributions but not specific classroom behaviors. His data indicated that while the correlation of student ratings with other student ratings were "of the order of .89," and between .64 and .76 when faculty-juries were compared with each other, the correlation between students and faculty-juries was .48. The statement was made that a likely cause of this "radical difference" is the fact that students have sat through many hours with the instructor, while the faculty are highly dependent on personal acquaintance and student hearsay.

Maslow and Zimmerman (1956) began an investigation out of what they referred to as "skepticism with the common tendency to dichotomize teaching and research (creating)." Correlating student and colleague ratings on instructors as good teachers, healthy personalities, and "creativeness in their field" the authors found a high degree of validity regarding student judgments when faculty judgments were used as a criterion measure. A high correlation (r=.69) with faculty judgments of the same teachers indicated to them that "a faculty cannot take student judgment lightly without casting aspersion on its' own competence to judge."

Attacking an earlier position by McKeachie, Borgatta (1970) offered the opinion that student ratings tend to indicate if a teacher is being "negligent." He states that this kind of feedback may lead professors to pay more attention to their teaching skills, participation and interest, but that it may be more
related to student morale and their educational objectives and
course selections. His most emphatic points deal with the
"questionable notion" of using these ratings for administrative
purposes such as promotion or salary determinations.

By isolating characteristics of "effective" university
teachers Crawford and Bradshaw (1968) attempted to determine scale
values of these characteristics by: (1) faculty members, (2)
university administrators, and (3) students. Their questionnaire
was developed largely from student input of essential teacher
characteristics. Findings indicated that not only do various
levels of faculty view performance differently but so do students
as well as administrators. The observation was made that these
characteristics should be interpreted only "in light of who is
judging."

In a study carried out in the Department of Psychology at the
University of Pittsburgh, Bendig (1952) used a sample of 490
students and instructors to investigate the use of 14 scales deal­
ing with teacher classroom behaviors, attitudes, skills, etc. He
concluded that students can discriminate among instructors through
the use of such scales. Twelve scales were significant at .001,
and two at .05.

Working with a faculty sample of teaching fellows and
students in introductory psychology courses at the University of
Michigan, Isaacson and McKeachie (1963) used a peer rating form
along with a self-descriptive adjective checklist and a student
evaluation rating instrument. They offered the judgment at the end of their findings that there was "no hesitancy to say that a college teacher's possession of agreeableness, emotional stability, and enthusiasm as well as high cultural attainment work well for him." Other variables, however, such as warmth may be highly dependent on student needs.

Checking claims for concurrent validity of student ratings Costin (1966) found $r = .49$ (p < .01) between ratings assigned to 32 instructors of a cross-section of university courses by department chairmen and those assigned by students on an assessment of overall effectiveness. Low correlations were achieved, however, on the individual items of his scale. Reasoning that this result may be a function of the greater difficulty chairmen had in making judgments on individual items as compared to the "overall" category he concludes that the results do lend support to the claims of the validity of student ratings.

McKeachie, Lin, and Mann (1971) worked to determine the validity of student ratings in terms of the criteria of student achievement. Data varied with teachers of differing types of skills being more effective with different types of students. The authors expressed disappointment and guessed that the "major slippage in validity studies" is the differing goals of teachers and students. They remarked that teaching effectiveness is not a unitary concept but one involving a number of complex interactions.
Numerous efforts have been carried out to test the relationship between various variables of students on ratings of instructors. Bendig (1953) stated that previous works indicated little or no effect of achievement on evaluation but that his study indicated some doubt. He concluded that achievement does affect the ratings, but not to a degree that invalidates the continued use of the scales. The suggestion is made that while individual achievement may be ignored, the class means should be considered.

At the State College of Washington Downie (1952) worked with a sample of 300 students to investigate the relationship of student achievement, enrollment status, college rank, and the class size upon the assessment of four factors identified as: (1) instructional procedures, (2) exams, (3) cultural value, and (4) instructor-student relations. The data indicated that students with higher achievement rated instructors higher on all four factors. Those enrolled on an elective basis rated instructors higher on factor four. There was no difference between upper and lower division student ratings. Large classes, however, rated teachers lower on factor one, two, and three than did smaller classes. An additional independent variable investigated was instructor experience. No differences were found between those with experience above or below five years.

Working with 131 students at Clemson University, Caffrey (1969) found that course grades, overall GPA, and sex of student
were not of critical importance in determining a students' rating of his teacher. Nor were the personal qualities of the teacher except for a factor labeled "rapport." The author decided that personal bias was absent from the evaluation of teachers made by students in this study.

Reporting a study carried out in the School of Education at Colorado State University involving all faculty and all students, Rayder (1968) concluded that student ratings were not related to student sex, age, class level, major, or grade point average. He did find, however, that the student ratings were somewhat related (r=.30) to such instructor characteristics as sex, age, degree held, rank, and years of experience. Further, he was able to ascertain the high degree of accuracy with which students reported their grade point averages.

Similarly Voeks and French (1960) at the University of Washington concluded at the end of their study that "high ratings cannot be bought with high grades, nor lost with low." They noted that both when judging overall value as a teacher and when rating his skill in specific aspects, students were rarely influenced by grades they received. An editorial note by the authors observed that, "College students appear to have greater objectivity and less superficial value systems than we had realized."

Based upon a study conducted at Central Michigan University in 39 English, history, personnel, and speech classes, Weaver (1960) arrived at the opposite point of view. His material
indicated to him that student ratings were biased in the direction of the grades they expected to receive. He also found that students expecting to receive higher grades were generally more discriminating in their appraisal of instructors. His conclusion was that the relatively narrow dispersion of scores by students expecting lower grades suggested that their rating behavior conformed to a culturally determined norm.

Collecting data verbally from students, Russell and Bendig (1953) discovered that relative achievement does not generally influence overall ratings of specific aspects of the instructor and the course. These specific aspects were primarily content and orientation of lectures, assignments, grading systems, and sense of humor. All of these were rated higher by better achieving students.

In a report to the 178th Annual Convention of the American Psychological Association Cohen and Berger (1970) noted that most studies have not found significant correlations between class grades and student ratings of instructors within a given section. Their study examined the correlation between an instructor's mean student ratings on various dimensions of teaching effectiveness and subsequent achievement by the class on a standardized test of material covered in the course. They observed that specific dimensions underlying students' ratings of instructors are "predictive of achievement on a comprehensive exam." It was the student-centered factors (interest and interaction with instructor)
which manifested this relationship rather than those aspects of
the ratings which emphasized either course structure (organiza-
tion and difficulty) or instructor regard.

Eash and Bennett (1964) found little evidence showing differ-
ences in academic achievement between small and large classes.
Attitudes toward teaching and learning, however, were signifi-
cantly different between the two. Carried out at Ball State
University in general psychology classes the work concluded that:
(1) large lecture classes plus small recitation groups did
significantly better than did straight lecture classes on
objective tests, (2) students from lecture-recitation classes
made more appointments to see the instructor, (3) attitudes toward
additional course work in psychology were not affected, and (4)
overall there were no negative effects from large classes. While
student ratings per se were not a part of this study, the possibil-
ity of varying attitudes toward teaching and learning could well
affect these ratings.

Investigating numerous independent variables which could
affect student ratings, McKeachie (1969) cited evidence that these
ratings correlate well (.40 to .68) with ratings of the same
instructors made by alumni who had graduated ten years previously.
He observed that students do not rate teachers on their personality
but on how they are learning. Further findings indicated that
undergraduate students tend to rate higher. Sex of the instructor
had no effect on the ratings. Associate professors tended to be
rated higher than other ranks. The degree held by the instructor did have an impact with higher degrees receiving higher ratings. Older teachers tended to be rated lower. The evidence was mixed regarding the effects of class size and whether the student was enrolled on a required or elective basis. The author made the statement that while student reactions are valuable for improving teaching he doubted their validity for inter-instructor comparisons. He further observed that the fact that feedback or knowledge of results aids learning is a "psychological principle of long standing."

Student Feedback As a Basis For Instructor Behavior Modification

With the rise in the use of student ratings of faculty and of the use of others' perceptions as feedback has come additional work on the effects of feedback insofar as behavior change is concerned. As is often the case, the results of these studies have produced mixed evidence.

In an introduction to his study Miller (1971) observed that H. H. Remmers commented in 1958 that...

"No teacher has any choice as to whether he wishes to be judged by his students. The only choice he has is whether he wishes to know how he is judged and thus possibly capitalize on this feedback."

The purpose of Miller's study was to determine whether providing instructors with feedback from students had effects on subsequent ratings by students and on student achievement. In
addition he attempted to assess whether these efforts were a function of instructor attitudes toward the value of student ratings. Analysis of the data suggests that feedback from students does not seem to alter instructor behavior, and that ratings do not appear to be changed as a function of instructor attitudes toward ratings. In addition the ratings generally were not related to student achievement.

Working with sixth-grade teachers Gage, Runkle, and Chatterjee (1963) provided the teachers with student feedback regarding the need for behavior change. They were able to record significant change. They remarked that feedback not only produced change, it also improved the accuracy of teachers' perceptions of their pupils' opinions.

Tuckman and Oliver (1968) involved 286 teachers of vocational education at the high school and technical level. Some of the teachers received feedback from students, some from supervisors, some from both, and some none at all. Using a twelve week interval between pre and post testing they concluded that student feedback had a significant effect but that from supervisors did not.

At Western Michigan University, Bryan (1963) spent approximately 35 years attempting to determine the effectiveness of written student feedback to teachers. His data indicated that 57 percent of the teachers receiving this feedback made significant gains in student ratings, compared to 24 percent of the control group.
A theory was developed by Gage, Runkle, and Chatterjee (1960) which they feel explains why behavior change is likely to occur when persons are provided with feedback from others. The rationale is that feedback can create an imbalance in the individual's self-perception that he will attempt to correct. His most likely response would be to modify his behavior, or at least to attempt to modify others' perceptions of his behavior.

Comparing various types of feedback, Ryan (1966) worked with teacher trainees and also attempted to determine what types of teachers were more receptive to feedback. His analyses did not indicate that any type of feedback was able to effect significant behavior change.

In his effort Savage (1957) investigated the effects of student feedback on student teacher behavior. While his results failed to show change resulting from feedback, a very serious limitation of the study was the fact that pretesting occurred with the semester only five days old. Further, only 20 days after the student teachers were given feedback posttesting took place.

At the University of Chicago, Lauroesch, Pereira, and Ryan (1969) also used student feedback in the teaching of interns. The comparison was made between student feedback, feedback supplemented by an interview, and no feedback at all. Data analysis indicated that interns who received only written feedback were rated lower on subsequent measures than those who received either no feedback or feedback plus an interview. Interns who received
no feedback were rated higher than those who received any type of feedback. The authors arrived at the conclusion that written student feedback can be detrimental to young teachers over a relatively brief period of time (eight weeks) but that if a longer period of time were allowed in order to accept, process, and utilize the feedback, beneficial results might be obtained.

Clark (1970) compared the effects of written student feedback, interaction analysis feedback, research-based statements, and group guidance in modifying the image of high school teachers. His work led him to conclude that all types of feedback were more effective in modifying teacher image than no feedback at all. Written student feedback was the most effective.

Working with a sample of higher education professors, Wolthuis (1970) attempted to determine what effect written student feedback would have on college instructors. Two types of written feedback were compared. The first consisted of both positive and negative information while the second contained only positive information. His observation indicated negative change on post-testing. All instructors received lower ratings on the second evaluation. Those who received no feedback displayed the least negative change and those who received only positive feedback the greatest negative change. His final statements included a notation that apparently feedback of any type is less valuable than no feedback when the criterion for determining the value of feedback is student ratings of instructors. He also commented that the
reduced ratings could also have been a result of students' reactions to the rating scales, since in the period that the study was carried out students in some classes had completed as many as six different such instruments.

Self-Perceptions and Self-Ratings

While empirical evidence concerning self-evaluation in education is almost non-existent, numerous authors have commented upon the need for, and value of, constant self-appraisal for teachers.

In his report of The Project To Improve College Teaching Eble (1970, p. 11) commented that the peak of popularity has probably passed for the use of self-evaluation as a means of improving instruction. However, he observed that "self-evaluation which precisely sets forth a teacher's objectives and which with equal precision describes and analyzes how these objectives might be achieved can be a brilliant contribution to our knowledge about teaching." He continued with the thought that this was probably not a sound idea for departmental evaluations.

As part of an essay Gray (1969, p. 642) mused, "How can we honestly call ourselves objective scholars unless we are willing to investigate teaching and make a serious effort at self-analysis?"

In their discussion Cohen and Brawer (1969, p. 10) made the statement that self-evaluation is the most difficult and yet the
most rewarding. They note that it assumes a degree of maturity and that a definite need for objectivity exists in the process. Agreeing with Eble they state that, "with self-examination there needs to be a definite basis upon which one must structure goals and objectives which act as a criterion for careful scrutiny." In regard to the use of student rating forms, they suggest that the instructor not average the responses but rather note patterns of responses that cluster about particular strengths and weaknesses.

As a statement of belief Jarrett (1969) commented that:

Any one of us can, if he will, be a surprisingly detached observer of himself, a careful and systematic analyst of his own teaching acts, an objective judge of "how the class went," or whether indeed it went at all or just lay there, playing dead.

Largely studies dealing with self-estimate or self-evaluation have come from the areas of psychology or educational psychology. Only rarely has the setting been a classroom situation.

Stotland, Thorley, Thomas, Cohen, and Zander (1957) observed that the research shows that a person's evaluation of his performance does not depend upon his absolute level of achievement alone. What counts is the level of achievement relative to his aspiration or goal line. Placement of this goal line may be influenced by different scales of reference one of which may be pressure from other persons or groups to perform in a particular fashion on that activity. Another source may be pressure from the person's own expectations concerning his likely performance in that activity.
arising from past experience. Their study examines the effect of a specific level of achievement when (1) the goal is set by the group, and (2) the goal level is relative to the person's stabilized expectations about himself as represented by his self-esteem. Analyses of the data indicated that group expectations appear to have been more potent as a scale of reference in determining his evaluation of his performance.

Comparing self, peer, and experts' ratings of personality adjustment, Powell (1948) attempted to determine the relationship between self-insight into adjustment and the "real" nature of the person as shown from ratings of peers and of an expert. The author remarked that there is a lack of insight into self-adjustment by comparison of self-rating to peer and expert ratings. Little agreement existed between any two of the measures leading Powell to comment that no one source should be taken as adequate evidence on which to base a diagnosis.

In a study utilizing both rural and urban later-elementary school teachers, Amatora (1955) attempted to ascertain the relationships between self-ratings and ratings by peers for these teachers on a number of personality variables. Her highest correlations were found in traits such as persistence .54, thoughtfulness .47, common sense .45, and sense of humor .43. The lowest r's appeared on items such as intelligence .13, egotism .16, tolerance .18, and sincerity .21.

Studying the validity of self-estimate Shen (1925) compared
self and peer ratings on eight traits. His work indicated a reliability range from .62 on impulsiveness to .91 for scholarship. He observed that the inaccuracy of self-estimates is largely due to systematic error on the part of the individual. This systematic error is believed to be a tendency to over or underestimate himself in all of the traits according to "the kind of delusions he has about himself."

At Wheaton College, Furbay (1969) made a preliminary report of a project designed to determine how participation in self-assessment groups influences the behavior of teachers. Trends suggested by the early data were that participants in the groups tended to differ from the control group in that they placed a greater emphasis on academic excellence, creative thinking, individualized instruction, openness of communication, and an understanding of student misbehavior. The groups utilized feedback data of the teachers' own design. Thus it was based upon their own instructional objectives. Therefore, they identified and confronted their own discrepancies.

In a study conducted in senior level educational psychology classes at the University of Delaware, Jenkins and Deno (1969) investigated the effects of varying types of feedback upon student teacher self-evaluation. The treatment variations were either positive or negative feedback. Their data report led them to conclude that student classroom behavior has a powerful influence on self-evaluation by teachers, since significantly higher
self-evaluation scores were recorded for those receiving positive feedback.

Gorwin and Payne (1962) dealt with the concept of "cross-perceptions" between students' views of teachers and teachers' views of students. Students were asked to give their opinion of the teacher and also to predict what the teacher thought of himself. Teachers were asked to indicate their opinions of themselves as a teacher and to predict what the students thought of him. The results suggest that teachers expected students to see them as they saw themselves, and that students expected teachers to see themselves as they had seen them. Both were wrong. The correlations between students descriptions of teachers and the teachers self-descriptions was .19. The t-ratio between the mean of students opinions and the mean of self-descriptions was not significant at the .05 level.

With a paper designed to survey the means of self-evaluation of 608 social science instructors Simpson (1961) reported that some 17 different procedures had been utilized. By comparison he reported that instructors in education courses had indicated that they used five variations. Neither group mentioned using a comparison of student and self-evaluations.

In the only study discovered which investigated the concept of comparing student ratings and self-evaluations, Webb and Nolan (1955) worked at the Naval Air Technical Training School at Jacksonville, Florida. Their study was based on feelings
that "personal learning and improvement stems from an understand-
ing of one's own adequacies and inadequacies." It was felt that a
self-evaluation serves to focus the individual's attention on his
inadequacies and as a result he will be motivated to attempt to
correct them. They felt this must be carried out in a non-
threatening situation and in conjunction with other evaluative
procedures. Teachers in this study were largely not professional
teachers. Likewise, supervisors were not trained professional
supervisors. The data indicated a high relationship between
student ratings and teacher self-ratings ($r=.62$). There was little
relationship between either student or self-ratings with supervisor
ratings. The authors stated that supervisors based their ratings
on some factor other than those which were valid estimates of
teaching ability or that they were random invalid intuitions.
"Not significant but high correlations (-.25 and -.23)" indicated
that those teachers with higher measured intelligence and more educa-
tion appeared to be more self-critical.

Summary

In this chapter an attempt has been made to present findings
of selected related literature. This literature represented the
three major concerns or premises of this project. In order these
areas of concern are (1) findings in the uses of student ratings
of instructors, (2) uses of feedback as a basis for behavior
change, and (3) the general area of self-perceptions and
self-ratings.

While evidence is sometimes mixed, generally studies indicated that students can reliably and objectively provide feedback regarding the classroom performance of instructors. Again with mixed evidence it would appear that these ratings are largely free of bias introduced through various independent variables associated with students such as grade point averages, class size, college rank, and sex.

Considerable evidence exists to indicate that feedback can have a significant effect upon an individual's behavior pattern. Gage et al. (1960) have suggested an interesting theory that an individual is placed in a state of disequilibrium by feedback which is not congruent with his own self-perceptions. He then will move to either modify his behavior or at least to modify the perceptions of others opinions of his behavior.

Investigation into teacher self-evaluation is almost non-existent, although many authors have praised it as a method and called for extended use of the procedure. Recognition is made by several of these writers that this evaluation must be based upon specifically outlined objectives and conducted through sound measurement procedures. As a result much of the literature surveyed deals with psychological concepts rather than with classroom performance. With some degree of varying evidence it would appear that it is possible for an individual to objectively evaluate his own performance. The literature does indicate that certain
personal variables could affect this self-estimation.

No studies were located which considered the question of comparing student ratings of instructors with the self-ratings of the instructors in higher education. During a conversation with this writer regarding mutual work with the concept, Dwight Allen, Dean of the School of Education at the University of Massachusetts, observed that it was amazing that in the long history of educational research no one had apparently thought to attempt to measure this very simplistic concept. This study attempted to provide empirical data regarding one aspect of this type of evaluation.
CHAPTER III

Methods, Design, Procedures

Procedures for the operation of this entire study were worked out in advance with two members of the doctoral committee. These were the Chairperson, Dr. Dorothy McCuskey and Dr. Uldis Smidchens, both of whom are faculty members in the Teacher Education Department. This was done in order to make it possible for them to be able to issue any necessary assurances to their colleagues, since the topic of rating faculty can be a very sensitive area with many instructors. Public approval of the project by two respected members of the department helped to ease concerns.

In addition, the entire study was discussed in advance with the Acting Chairman of the department, Dr. Kenneth Dickie. Following minor modifications in the cover letter, his public approval was granted. This too, helped to ease faculty concerns.

Each member of the department was visited and his/her cooperation solicited. With the endorsement of the aforementioned persons, this was largely a task of arranging interviews and sitting down to explain the overall aspects of the study. Every attempt was made to provide for limited time demands upon instructors as they supplied data. It was estimated that the time involvement was probably five to ten minutes per class.

In this chapter, a discussion is made of the full procedures of this study. The population and sample is identified, the
instrumentation explained, research design, hypotheses, and general procedures are discussed in this section. The chapter concludes with a description of the statistical analyses.

Population and Sample

The population for this study was located in the Teacher Education Department, College of Education, at Western Michigan University. The group was made up of all faculty members within the department, excluding those assigned to the Directed Teaching Office and the Continuing Education Center. Also, all students enrolled in Teacher Education Department (TEED) class offerings during the Winter Semester, 1972 with the exceptions of the previously mentioned divisions, were part of the total population. These two divisions, Directed Teaching and Continuing Education, were eliminated because of the atypical nature of their functions within the department.

As mentioned previously, the 61 faculty members of the department who were teaching classes appropriate for the use of the Student Opinions About Instructional Procedures questionnaire were approached in individual conferences and their cooperation solicited. Fifty-eight instructors agreed to participate and to supply their self-perceptions of their classroom behaviors. Thus, the faculty sample of the study was made up of 95 percent of the members of the population. These instructors taught a combined total of 135 classes. Seventy-seven of these classes had 30 or fewer enrolled,
while 58 had over 30. A further characteristic of this sample to be noted, is that with the exception of two members new to the department, all participants had in the previous year received feedback data from students regarding their classroom procedures. Additional characteristics are presented in Table 1 located in Chapter IV.

The student sample was determined by the cooperation of the faculty, since it was necessary to obtain each instructor's permission to utilize student data about himself. Thus, the student sample was made up of students enrolled in both graduate and undergraduate classes of participating instructors. The size of this sample was somewhat in excess of 3,000 students. A further description of this sample is presented in Table 2 located in Chapter IV.

Instrumentation

Student Opinions About Instructional Procedures

The SOAIP questionnaire (Appendix A) served as the main instrument of this study. All data collected from students were by means of this questionnaire. Developed in the Teacher Education Department at Western Michigan University through the cooperative efforts of faculty and students, the instrument has been used for two years as a means whereby students may evaluate teaching proficiency. Often such instruments are notable for the lack of systematic procedures by which they are produced. The SOAIP, however, was developed by an Ad Hoc Committee of the department with the
speciality in educational measurement being offered by Dr. Uldis Smidchens, Director of the Center for Educational Research at the University.

Work was begun on the questionnaire in September of 1969 following a charge presented by the University Faculty Senate to all departments for the development of instruments "to be used by students for the evaluation of teaching proficiency." Following the charge by the Senate, the committee moved to establish the following objectives as the guidelines for development of the instrument:

1. The entire department should participate in the construction and approval of the evaluation instrument.

2. A representative sample of students should participate in the construction and approval of the evaluation instrument.

3. The evaluation form should be complete in its coverage of teaching activities and yet consume the least amount of class time possible in its completion.

4. The respondents should be free to rate only those activities which apply to the teaching-learning situation and of which he has personal knowledge.

5. The evaluation instrument should be first, designed to collect data which will facilitate the improvement of instruction and secondly, to fulfill an administrative need for ratings on which to substantiate promotions and salary increases.

Using these guidelines, similar evaluation instruments used at Indiana University, Michigan State University, University of Michigan, and Ohio University were examined. Considerable overlap in the criteria used was noted. It appeared that the criteria found in these instruments would fit into one of the following categories: (1) evaluation of student, (2) personal relationships.
between faculty and students, (3) professional competence of faculty, and (4) individual (personal and physical) characteristics of the faculty member.

Under these categories, all statements from the instruments reviewed were then listed. These lists were reviewed and all redundant items deleted. From an original list of 140 criterion statements, a list of 61 relatively independent items was evolved. To involve faculty members and as many students as possible in the design, a plan was used whereby both students and faculty selected from the 61 statements, both those which were meaningful to the evaluation of teaching proficiency, and those which could be evaluated. The thought behind this plan was that "any instrument which contained statements which were viewed by the faculty and students as being meaningful to the evaluation of teaching proficiency and further, were seen by the faculty and students as being measurable criteria, would be acceptable and valid for the population involved." (Committee for the Evaluation of Teaching Proficiency, 1970).

These instruments of 61 items were then administered to both faculty and students. The faculty members were asked to respond first in terms of their perceptions of the students' ability to make the judgments necessary and secondly, in terms of their interpretations of the value of the characteristic to the evaluation of teaching. Students were also requested to react to the items in two ways. First, to their own ability to rate faculty on the item characteristic and secondly, to their perception of the value of the
item contents to the evaluation of teaching. A total of 61 faculty members were sent packages of material. Thirty-nine were returned representing a total of 923 usable student responses.

Responses of both faculty and students were analyzed to show percentages of both groups in each category on each item. The final instrument was developed by selecting only those items where there was at least 80 percent or greater agreement among both faculty and students on the importance of that item in the evaluation of teaching. The items selected were then rechecked in terms of the perceptions of faculty and students belief in the measurability of the item. Again, a minimum of 80 percent agreement was the cutoff point. Following this process, 21 items remained. Two were deleted because of redundancy. The remaining 19 items make up the evaluation form.

The instrument has been factor analyzed and three factors were identified which jointly account for 57 percent of the total variance. These factors have been labeled "Professional Competence, Evaluation Process, and Student Centeredness." In addition, a split-half reliability formula has been computed and has shown a Spearman-Brown r of .50 on the mean of the 19 items. While this value is somewhat low in view of previous studies reporting values in excess of .80 for similar instruments, it is felt from viewing the data that one reason for the lower value might lie with the scale used. The SOAIP utilizes a 5-point scale which describes the characteristic as almost never present, infrequently present,
frequently present, almost always present, or undecided. Research conducted under the direction of Roy C. Bryan at Western Michigan University by DuBois (1960 p. 28) indicates that a 4-point scale definitely tends to exhibit greater skewness than does a 5-point scale. The use of a 5-point scale produces a clear tendency toward a more normal frequency distribution. The use of the 5-point could have spread the scores out and possibly have produced the slightly lower reliability coefficient. Further, the instrument was adopted for use by the department with the understanding that it might be imperfect and "in reality represents only a beginning in the design of an evaluation instrument for the assessment of teaching proficiency."

Instructor Self-Perceptions About Instructional Procedures

The Instructor Self-Perception instrument (Appendix B) was used to record the instructors' self-perceptions about their instructional procedures. In the content of the individual 19 items, it is identical to the SOAIP. It differs only in that the pronouns have been changed from the third to the first person. Informational items about the respondent vary also. For example, instead of the grade point average item found in the student version, the instructor is asked to indicate the highest educational degree he holds. By substituting items such as these, additional information about instructors was obtained, and it also became possible through identical data formats to utilize some of the computer programs used
for the analysis of the SOAIP.

Design

The design of this study was to gather data from an existing field situation. Through analyzing the data by t-ratios and correlation coefficients it was believed that conclusions could be drawn regarding the relationships of the variables under consideration. The expectation was that through this method, some degree of uncertainty regarding these relationships, would be reduced.

The dependent variables are the ratings assigned to instructors by students through the use of the SOAIP. Similarly, the ratings instructors assigned to themselves through the faculty version of the instrument, are dependent variables.

Near the end of the Winter 1972 semester, students responded to the SOAIP in terms of their class instructor. At approximately the same time, participating instructors recorded their own perceptions of their classroom procedures. The data from both of these groups were examined to determine the nature of the relationship between the dependent variables (rating scores) and independent variables (grade point average, class size, course selection basis, and instructor teaching experience) under consideration. This examination of the independent-dependent variables was largely conducted through the use of discrepancy scores. These values were developed simply by comparing each student's responses to the responses presented by the class instructor.
Hypotheses

The general objectives of this study were to examine the nature of the relationship between student ratings of instructors and the self-assigned ratings by instructors for various student, instructor, and classroom characteristics. To the author's knowledge, no previous efforts have been made to study these relationships in a large teacher education department. Many writers have, however, reported the results of their findings regarding the nature of the relationship between student characteristics and the way they rate faculty.

The theoretical hypotheses of this study were:

$H_1$: A positive relationship exists between student ratings and instructor self-perceptions regarding "professional competence" behaviors.

$H_2$: A positive relationship exists between student ratings and instructor self-perceptions regarding "evaluation procedures."

$H_3$: A positive relationship exists between student ratings and instructor self-perceptions regarding "student centeredness" behaviors.

$H_4$: The degree of congruence between student ratings and instructor self-perceptions regarding "professional competence" behaviors will vary with the student's grade point average.

$H_5$: The degree of congruence between student ratings and instructor self-perceptions regarding "evaluation procedures" will vary with the student's grade point average.

$H_6$: The degree of congruence between student ratings and instructor self-perceptions regarding "student centeredness" behaviors will vary with the student's grade point average.
$H_7$: The degree of congruence between student ratings and instructor self-perceptions regarding "professional competence" behaviors will vary with the size of the class.

$H_8$: The degree of congruence between student ratings and instructor self-perceptions regarding "evaluation procedures" will vary with the size of the class.

$H_9$: The degree of congruence between student ratings and instructor self-perceptions regarding "student centeredness" behaviors will vary with the size of the class.

$H_{10}$: The degree of congruence between student ratings and instructor self-perceptions regarding "professional competence" behaviors will vary with the basis used by the student to select the class.

$H_{11}$: The degree of congruence between student ratings and instructor self-perceptions regarding "evaluation procedures" will vary with the basis used by the student to select the class.

$H_{12}$: The degree of congruence between student ratings and instructor self-perceptions regarding "student centeredness" behaviors will vary with the basis used by the student to select the class.

$H_{13}$: The degree of congruence between student ratings and instructor self-perceptions regarding "professional competence" behaviors will vary with the instructor's teaching experience.

$H_{14}$: The degree of congruence between student ratings and instructor self-perceptions regarding "evaluation procedures" will vary with the instructor's experience.

$H_{15}$: The degree of congruence between student ratings and instructor self-perceptions regarding "student centeredness" behaviors will vary with the instructor's experience.

General Procedures

The outline of this study was begun early in the Winter Semester.
of 1972. During this semester, numerous conferences were held with members of the doctoral committee and with the Acting Chairman of the Teacher Education Department. Following approval of the study by the committee and upon being granted permission by Dr. Dickie to approach department members to solicit their cooperation, an appointment was arranged in early April, with every member of the department included in the population. A letter (Appendix C), outlining the exact nature of the request being made, was prepared and handed to each instructor early in each interview. In addition, a one-page abstract (Appendix D), explaining in greater detail the precise nature of the study, was given to each faculty member. If consent was given to participate, the faculty member was provided with a copy of the instrument to which he/she was to respond and an answer sheet (Appendix E) for each class he/she was teaching. Self-perceptions were recorded by the instructors according to their own time considerations and returned to the investigator generally via a sealed inter-office envelope. Of the 61 members of the population under study, 58 (95 percent) elected to take part.

Concerning the faculty sample, it should be noted that virtually all members had previously received feedback data from student responses to the instrument being used. It is possible that the instructors' self-perceptions were influenced consciously or unconsciously by this feedback. The only exceptions to this limitation were the two first-year members in the department.

Further, in two situations, instructors were involved in
cooperative teaching efforts. Thus, the students were asked to assess behaviors and procedures on an individual instructor basis while two instructors were involved in the interaction.

The Teacher Education Department had systematized the procedures for collecting data from students. Responsibility for distributing materials and collecting data was supervised by the acting chairman, while analyzing data was carried out by the Center for Educational Research. Directions (Appendix F) for the distribution of questionnaires and answer sheets were sent out to all faculty early in March. Since a student from the class actually administered the instrument with the instructor absent, detailed instructions (Appendix G) were provided with each questionnaire in order to standardize the procedure from class to class. In addition, the investigator was a member of the staff of the Educational Research Office, making it possible to readily sort out the data provided by students regarding instructors who had elected to participate. By doing so, these instructors had given permission to utilize this data for the purposes of the study.

Concern was evidenced that students enrolled in more than one class with a given instructor could provide contaminated data. In order to determine the extent to which this type of enrollment was occurring, approximately 20 percent of the instructors were included in a sample which was asked how many dual-enrollees they had taught. Based on this feedback, it would appear that perhaps three percent of the total student sample were actually enrolled with the same
instructor in two classes. Ideally, these responses would have been discarded, but since the student feedback is completely anonymous this procedure was not possible.

In order to insure full protection of anonymity for every faculty member, each was assigned a code number prior to data collection. In this report, all instructors were viewed as a group or in subgroups making it impossible for any one person to be singled out. Following completion of this project, the code will be destroyed, finalizing the protection of all individuals who participated.

Each instructor who participated was also provided with feedback (Appendix H) regarding the discrepancies between the way he rated himself and the way he was viewed by his students. This feedback was provided for two reasons. First, it was an added inducement for faculty members to participate in the study and secondly, it was an attempt to instill the concept that this procedure is a workable and valuable one for the improvement of college teaching.

Statistical Analyses

The following discussion of procedures to be utilized in the analysis of data collected will be organized on the basis of each hypothesis. Each hypothesis will be itemized and the procedures to be utilized in answering the question discussed.

All analyses undertaken were on the basis of the three factors...
the instrument was measuring. As means, discrepancy scores, t-ratios and correlations are discussed, it should be remembered that these scores and values were computed on the basis of each factor. For example, a mean score for each student from smaller classes was determined for factor one, factor two, and factor three, as were the means for all other categories and subcategories.

**Hypotheses One Through Three**

The data collected to study Hypotheses One through Three were analyzed factor by factor, first through the use of an independent t-ratio and secondly, with a Pearson Product-Moment Correlation which measured the degree to which each instructor's self-perceptions varied with the ratings given him by each of his students. The mean for the student sample has been determined and compared to the mean of the 135 instructor responses for each of the three factors. Interest in the correlation displayed by this data varies with the significance of the differences between means. With highly significant differences, the interest in the correlation is lessened.

**Hypotheses Four Through Six**

In order to analyze data for Hypotheses Four through Six, which dealt with student achievement, the students in each class were divided into two groups. The first of these consisted of those reporting a GPA of 3.0 or higher, and the second, those reporting less than 3.0. Student mean factor scores were compared to the
faculty mean factor scores and, a discrepancy score computed for each student. These discrepancy scores represent the absolute difference between the faculty member and each of his students. Following computation of these scores for each group, a comparison of the mean discrepancy score for high achievers and low achievers was made through the use of an independent t-ratio.

**Hypotheses Seven Through Nine**

The independent variable, in the case of Hypotheses Seven through Nine was class size. Classes were categorized on the basis of size with those having 30 or fewer students making up one group, and those with over 30, the second group. Again, each student's mean factor scores were compared to his/her instructor's responses to develop a discrepancy score. The mean discrepancy score for classes of 30 or less was compared to the mean for those classes of over 30 by means of a t-ratio.

**Hypotheses Ten Through Twelve**

Procedures to study the questions posed by Hypotheses Ten through Twelve were virtually identical to those used for $H_4$-$H_6$. Students within each class were divided on the basis of those taking an elective versus a required course. Mean factor scores for students and faculty are compared and discrepancy scores determined on an individual student basis. The mean discrepancy score for the required-course group was compared to the mean for the elective-
course group with an independent t-ratio.

**Hypotheses Thirteen Through Fifteen**

Similarly, the analysis of data for Hypotheses Thirteen through Fifteen corresponds to procedures outlined for $H_7-H_9$. All classes have been divided into two groups based on the experience factor of the instructor. Students of those instructors having five or more years of high education teaching background comprised one group with the remainder making up the second. Discrepancy scores between student and faculty responses were developed and means compared for the two groups with an independent t-ratio.

All t-values were reported and described in terms of probability levels. Although the .05 level was utilized in order to decide whether the null hypothesis would be rejected, the reader was presented with the opportunity to view the precise probability at which any observed difference is likely to occur by chance.

**Summary**

This project is a field study concerned with the relationships between student's ratings of instructors' classroom behaviors and instructional procedures and the instructors' self-perceptions of these same behaviors and procedures. Certain characteristics of both students and faculty members were identified as independent variables and statistical comparisons made between the responses of the varying groups.
Fifty-eight instructors from the Teacher Education Department, College of Education at Western Michigan University elected to participate in the study. The sample represented 95 percent of the total population of faculty members. The student sample exceeded 3,000 and was found in the 135 classes taught by participating instructors.

The basic instrument used for data collection was the "Student Opinions About Instructional Procedures" questionnaire developed by a committee of the Teacher Education Department. Slight pronoun adjustment was made to the context of each item in order to adapt the questionnaire for the collection of faculty self-perceptions.

Near the end of the Winter Semester 1972, each student responded to the SOAIP concerning his opinions of his instructor's classroom behaviors. The procedures for the collection of these responses had previously been established by the Teacher Education Department. Participating instructors responded to an almost identical instrument in terms of the way they perceived their own classroom behaviors. In order to elicit this participation, each instructor was visited and the goals of the project explained to him/her. During this visit, the instructor was presented the questionnaire, if cooperation was indicated, and asked to record responses in a thoughtful manner.

The data were statistically analyzed, basically by means of t-ratios and Pearson Product-Moment Correlations, to determine the nature of the relationship between the independent and dependent variables.
variables. Attention was focused through these analyses to the discrepancies between student and faculty responses, values of the t-ratios were described in terms of the probability level at which they could occur by chance.
CHAPTER IV

Results

The data collected through the means described in Chapter III were analyzed by viewing the differences between various group means and Pearson Product-Moment Correlations. These findings are presented in the following manner: First, characteristics of the samples included in the analyses are described and secondly, the results of the analyses are displayed following the outline of the hypotheses. Statistically significant findings are denoted by an asterisk. Each table is followed by a brief discussion of the findings.

Samples Included in the Analyses

Instructor Population and Sample

Sixty-one instructors taught classes in the Teacher Education Department which were appropriate for the use of the SOAIP. Of this population, 58 or 95 percent elected to participate in the study. The characteristics of this sample are outlined in Table 1. Concerning those who did not participate: all three were males; one was a full professor, one an associate professor, and one an assistant professor; one held a doctorate, and two a masters degree; all three were full-time instructors; and all had taught in higher education five or more years.
TABLE 1
Characteristics of Instructors Providing
Self-Perceptions About Instructional Procedures

<table>
<thead>
<tr>
<th>Total Number of Instructors</th>
<th>58</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank of Instructor</td>
<td></td>
</tr>
<tr>
<td>Instructors</td>
<td>2</td>
</tr>
<tr>
<td>Assistant Professors</td>
<td>22</td>
</tr>
<tr>
<td>Associate Professors</td>
<td>18</td>
</tr>
<tr>
<td>Professors</td>
<td>16</td>
</tr>
<tr>
<td>Total Number of Classes</td>
<td>135</td>
</tr>
<tr>
<td>Sex of Instructors</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>38</td>
</tr>
<tr>
<td>Female</td>
<td>20</td>
</tr>
<tr>
<td>Instructor Employment Status</td>
<td></td>
</tr>
<tr>
<td>Part-time</td>
<td>5</td>
</tr>
<tr>
<td>Full-time</td>
<td>53</td>
</tr>
<tr>
<td>Degree Held by Instructor</td>
<td></td>
</tr>
<tr>
<td>MA - MS</td>
<td>18</td>
</tr>
<tr>
<td>Ed.S.</td>
<td>1</td>
</tr>
<tr>
<td>Ed.D. - Ph.D.</td>
<td>39</td>
</tr>
<tr>
<td>Instructor Teaching Experience</td>
<td></td>
</tr>
<tr>
<td>Less than Five Years</td>
<td>16</td>
</tr>
<tr>
<td>Five Years or More</td>
<td>42</td>
</tr>
</tbody>
</table>

Student Sample

The student sample included in this study was made up of approximately 96 percent of the total population of students enrolled in TEED classes. The following table details numerically characteristics of this group. Without the cooperation of all instructors, however, it was not possible to ascertain the characteristics of those members of the student population not included in the sample.
TABLE 2

Characteristics of Students Responding to the Student Opinions About Instructional Procedures Questionnaire

<table>
<thead>
<tr>
<th>Total Number of Students</th>
<th>Student Enrollment Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>3048</td>
<td>Part-time 592</td>
</tr>
<tr>
<td></td>
<td>Full-time 2424</td>
</tr>
<tr>
<td></td>
<td>Status Not Identified 32</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sex of Students</th>
<th>Classification of Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>Freshmen 48</td>
</tr>
<tr>
<td>Female</td>
<td>Sophomores 721</td>
</tr>
<tr>
<td>No Sex Identified</td>
<td>Juniors 1055</td>
</tr>
<tr>
<td></td>
<td>Seniors 485</td>
</tr>
<tr>
<td></td>
<td>Graduates 638</td>
</tr>
<tr>
<td></td>
<td>Non-Degree 75</td>
</tr>
<tr>
<td></td>
<td>No Class Identified 26</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Level Planning to Teach</th>
<th>Reported Grade Point Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elementary 1356</td>
<td>Less than 1.0 6</td>
</tr>
<tr>
<td>Secondary 924</td>
<td>1.0 - 1.99 47</td>
</tr>
<tr>
<td>Post-Secondary 193</td>
<td>2.0 - 2.99 1260</td>
</tr>
<tr>
<td>Not Teaching But</td>
<td>3.0 - 4.0 1588</td>
</tr>
<tr>
<td>Education 222</td>
<td>GPA Not Established 114</td>
</tr>
<tr>
<td>Not Education 243</td>
<td>GPA Not Reported 33</td>
</tr>
<tr>
<td>Level not Identified 110</td>
<td></td>
</tr>
</tbody>
</table>

Relationships of Student-Ratings and Instructor Self-Perceptions

This section of the chapter is organized according to the statement of the null hypotheses. Each null hypothesis will be presented and the pertinent statistical findings indicated in table.
form. Accompanying each null hypothesis and table will be a discussion of the findings.

Ho: No differences exist between mean student ratings and mean instructor self-perceptions of classroom behaviors and procedures regarding the factor professional competence.

TABLE 3
Relationship Between Student-Ratings and Instructor Self-Perceptions Regarding Professional Competence

<table>
<thead>
<tr>
<th>Student Responses</th>
<th>Instructor Responses</th>
<th>df</th>
<th>t</th>
<th>p</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>N:</td>
<td>3024</td>
<td>135</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M:</td>
<td>2.656</td>
<td>2.663</td>
<td>3157</td>
<td>.172</td>
<td>.86</td>
</tr>
<tr>
<td>SD:</td>
<td>.498</td>
<td>.314</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results of the analysis presented in Table 3 indicate that no significant difference exists between student ratings of instructors' professional competence and the way that instructors view themselves regarding this factor. Since significant differences were not demonstrated by the data, the null hypothesis is not rejected.

Attempting to assess the degree to which the ordered pairs of student and faculty responses vary together, the value of a Pearson Product-Moment Correlation was determined. This value indicates
that virtually no relationship existed between the two variables.

**H0₂**: No difference exists between mean student ratings and mean instructor self-perceptions of classroom behaviors and procedures regarding the factor evaluation procedures.

**TABLE 4**

<table>
<thead>
<tr>
<th></th>
<th>Student Responses</th>
<th>Instructor Responses</th>
<th>df</th>
<th>t</th>
<th>p</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>N:</td>
<td>2836</td>
<td>133</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M:</td>
<td>2.426</td>
<td>2.532</td>
<td>2967</td>
<td>1.726</td>
<td>.08</td>
<td>.05</td>
</tr>
<tr>
<td>SD:</td>
<td>.701</td>
<td>.591</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Significant differences do not exist between the way students rated instructors regarding their evaluation procedures and the way instructors perceived their own behaviors. As can be seen from the information in Table 4 the .05 level of significance was not achieved, no rejection of the null hypothesis is confirmed.

In an effort to determine the degree to which student and instructor ratings of evaluation behaviors vary together, a Pearson Product-Moment Correlation was computed. The value of this *r* indicated that the two variables display almost no tendency to vary together. That is, it is virtually impossible to predict instructor
self-ratings from student ratings.

$H_03$: No difference exists between mean student ratings and mean instructor self-perceptions of class behaviors and procedures regarding the factor student centeredness.

**TABLE 5**

Relationship Between Student Ratings and Instructor Self-Perceptions Regarding Student Centeredness

<table>
<thead>
<tr>
<th>Student Responses</th>
<th>Instructor Responses</th>
<th>df</th>
<th>t</th>
<th>p</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>N: 3025</td>
<td>135</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M: 2.606</td>
<td>2.655</td>
<td>3158</td>
<td>1.011</td>
<td>.31</td>
<td>.12</td>
</tr>
<tr>
<td>SD: .554</td>
<td>.277</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From the results of the data presented in Table 5, it is apparent that no difference exists between the manner in which students rate instructors regarding student centeredness and the way instructors rate themselves. These differences could occur by chance, 31 times in 100. This probability level failed to approach the level of significance necessary to reject the null hypothesis. The Pearson Product-Moment $r$ computed to determine the strength of the relationship between the two ratings was so low as to indicate that very little association does exist between the two.

As no significant differences and little tendency to vary together between the variables was displayed, interest increased in
the effort to learn more about the manner in which students were
rating faculty and faculty were rating themselves. In the following
table additional correlational values of interest are presented
regarding these relationships. In this table, the symbol S refers
to students while I, indicates instructors. The letter F indicates
factor. The numeral following this designation indicates the factor
to which reference is being made according to the following code:
(1) Professional Competence, (2) Evaluation Procedures, and (3)
Student Centeredness.

\[
\begin{array}{cccccc}
\text{S} & \text{F}_1 & \text{F}_2 & \text{F}_3 & \text{I} & \text{F}_2 \\
\hline
\text{SF}_1 & 1.000 & .526 & .785 & \text{IF}_1 & \text{IF}_2 \\
\text{SF}_2 & 1.000 & .546 & & \\
\text{SF}_3 & & 1.000 & & \\
\text{IF}_1 & 1.000 & .234 & .584 & \\
\text{IF}_2 & 1.000 & .217 & & \\
\text{IF}_3 & & 1.000 & & \\
\end{array}
\]

Based on the findings reported in Table 6, it would appear that
the student factor ratings show a fairly consistent tendency to
vary together. Apparently, students tended to rate their instructors
in a like manner concerning all classroom behaviors. In other words, an instructor rated high on professional competence is quite likely ($r=.78$) to be rated high on student centeredness, etc.

Instructors, on the other hand, while showing some of the same tendencies, appeared to discriminate more highly between the behavioral factors. Again, it would seem that instructors rated themselves high ($r=.58$) on student centeredness. These correlations, however, generally do not approach the magnitude displayed by the student data.

$H_0_4$: No difference exists between the mean discrepancy score of ratings assigned by higher achieving students versus instructors' self-perceptions, and the mean discrepancy score of ratings assigned by lower achieving students versus instructors' self-perceptions regarding the factor professional competence.

### TABLE 7

Relationship Between Student GPA and the Congruence of Student Ratings and Faculty Self-Perceptions Regarding Professional Competence

<table>
<thead>
<tr>
<th></th>
<th>Discrepancy of High Achiever vs. Instructor</th>
<th>Discrepancy of Low Achiever vs. Instructor</th>
<th>df</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>N:</td>
<td>1576</td>
<td>1303</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M:</td>
<td>.371</td>
<td>.427</td>
<td>2777</td>
<td>3.602</td>
<td>.001*</td>
</tr>
<tr>
<td>SD:</td>
<td>.391</td>
<td>.436</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
From the data presented in Table 7 it is apparent that differences which exist concerning the level of discrepancy between faculty and high achievers and the level of discrepancy between faculty and low achievers concerning professional competence were significant. The null hypothesis is rejected since a relationship is indicated between student GPA and the congruence between student ratings and faculty self-perceptions. High achiever's ratings are significantly more congruent with the levels self-assigned by faculty than are those ratings assigned by low-achieving students in terms of professional competence.

Ho5: No difference exists between the mean discrepancy score of ratings assigned by higher achieving students versus instructors' self-perceptions, and the mean discrepancy score of ratings assigned by lower achieving students versus instructors' self-perceptions regarding the factor evaluation procedures.

| TABLE 8 | Relationship Between Student GPA and the Congruence of Student Ratings and Faculty Self-Perceptions Regarding Evaluation Procedures |
| Discrepancy of High Achiever vs. Instructor | Discrepancy of Low Achiever vs. Instructor | df | t  | p  |
| N: 1459 | 1229 |  |
| M: .612 | .647 | 2686 | 1.420 | .15 |
| SD: .637 | .671 |  |

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
From the results of the data presented in Table 8, it is apparent that no significant relationship exists between the student's GPA and the degree of congruence between his/her ratings of the instructor and the instructor's self-ratings regarding evaluation procedures. The null hypothesis is not rejected.

\textbf{H}_0^2: \text{No difference exists between the mean discrepancy score of ratings assigned by higher achieving students versus instructors' self-perceptions, and the mean discrepancy score of ratings assigned by lower achieving students versus instructors' self-perceptions regarding the factor student centeredness.}

\begin{table}
\centering
\caption{Relationship Between Student GPA and the Congruence of Student Ratings and Faculty Self-Perceptions Regarding Student Centeredness}
\begin{tabular}{llllll}
\hline
 & Discrepancy of & Discrepancy of & df & t & p \\
 & High Achiever & Low Achiever & & & \\
 & vs. Instructor & vs. Instructor & & & \\
\hline
 N: & 1581 & 1298 & & & \\
 M: & .404 & .432 & 2878 & 1.813 & .07 \\
 SD: & .412 & .438 & & & \\
\hline
\end{tabular}
\end{table}

According to the findings indicated in Table 9, no significant differences exist regarding the levels of congruence displayed between higher achievers versus instructors and low achievers versus instructors. However, it would seem high achievers do tend to reach
a higher level of agreement with instructors regarding the evaluation of student centeredness behaviors than do low achievers. These differences could occur approximately seven times in 100. This level is not sufficient to reject the null hypothesis.

$H_0$: No difference exists between the mean discrepancy score of ratings assigned by students in larger classes versus instructors' self-perceptions, and the mean discrepancy score of ratings assigned by students in smaller classes versus instructors' self-perceptions regarding the factor professional competence.

**TABLE 10**

Relationship Between Class Size and the Congruence of Student Ratings with Faculty Self-Perceptions Regarding Professional Competence

<table>
<thead>
<tr>
<th>Discrepancy of</th>
<th>Discrepancy of</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Class</td>
<td>Large Class</td>
</tr>
<tr>
<td>Students vs. Instructor</td>
<td>Students vs. Instructor</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>df</th>
<th>$t$</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>3022</td>
<td>3.767</td>
<td>.001*</td>
</tr>
</tbody>
</table>

Differences are apparent, in Table 10, in the levels of agreement between the ratings of students from small classes with their instructors and those from large classes with their instructors regarding professional competence. The analysis indicates that these

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
differences in congruence could occur by chance only one time in more than 1000. Therefore, the null hypothesis is rejected. A relationship does exist between class size and these levels of congruence. Interesting to note, however, the students from larger classes displayed a higher level of congruence with their instructor.

H₀: No difference exists between the mean discrepancy score of ratings assigned by students in larger classes versus instructors' self-perceptions, and the mean discrepancy score of ratings assigned by students in smaller classes versus instructors' self-perceptions regarding the factor evaluation procedures.

TABLE 11

Relationship Between Class Size and the Congruence of Student Ratings with Faculty Self-Perceptions Regarding Evaluation Procedures

| Discrepancy of | Discrepancy of |
| Small Class    | Large Class    |
| Students vs.   | Students vs.   |
| Instructor     | Instructor     |
| df             | t              | p       |
|----------------|----------------|
| N: 1241        | 1575           |
| M: .620        | .631           |
| SD: .651       | .654           |
| 2814           | .417           |
| .67            |                |

The results displayed in Table 11 indicate that no relationship was observed between the size of the class and the degree of congruence between student ratings and faculty self-ratings regarding
evaluation procedures. The reported differences could occur 67 times in 100. The null hypothesis was not rejected.

\[ H_0: \] No difference exists between the mean discrepancy score of ratings assigned by students in larger classes versus instructors' self-perceptions, and the mean discrepancy score of ratings assigned by students in smaller classes versus instructors' self-perceptions regarding the factor student centeredness.

**TABLE 12**

<table>
<thead>
<tr>
<th>Discrepancy of</th>
<th>Discrepancy of</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Class Students vs. Instructor</td>
<td>Large Class Students vs. Instructor</td>
</tr>
<tr>
<td>df</td>
<td>t</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>N: 1340</td>
<td>1685</td>
</tr>
<tr>
<td>M: .450</td>
<td>.386</td>
</tr>
<tr>
<td>SD: .430</td>
<td>.412</td>
</tr>
<tr>
<td>3023</td>
<td>4.120</td>
</tr>
</tbody>
</table>

According to the results of the analysis of these data shown in Table 12, a significant relationship does exist between the size of the class and the degree of congruence between student ratings and faculty self-perceptions regarding student centeredness behaviors. The degree of congruence was greater between students from large classes and their instructors than between those involved in smaller classes.
classes. The findings caused the null hypothesis to be rejected.

\[ \text{Ho}_{10}: \text{No difference exists between the mean discrepancy score of ratings assigned by students enrolled in required classes versus instructors' self-perceptions, and the mean discrepancy score of ratings assigned by students enrolled in elective classes versus instructors' self-perceptions regarding the factor professional competence.} \]

**TABLE 13**

**Relationship Between Class Selection and the Congruence of Student Ratings with Faculty Self-Perceptions Regarding Professional Competence**

<table>
<thead>
<tr>
<th>Discrepancy of Elective Class Students vs. Instructor</th>
<th>Discrepancy of Required Class Students vs. Instructor</th>
<th>df</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>N: 585</td>
<td>2401</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M: .418</td>
<td>.386</td>
<td>2984</td>
<td>1.699</td>
<td>.08</td>
</tr>
<tr>
<td>SD: .458</td>
<td>.458</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results of the analysis shown in Table 13 indicate that no significant relationship exists between the type of class selection and the degree of congruence concerning student ratings and instructor self-ratings regarding professional competence. The observed results could occur eight times in 100. This observed statistical significance level is not sufficient to reject the null hypothesis.
Ho11: No difference exists between the mean discrepancy score of ratings assigned by students enrolled in required classes versus instructors' self-perceptions, and the mean discrepancy score of ratings assigned by students enrolled in elective classes versus instructors' self-perceptions regarding the factor evaluation procedures.

**TABLE 14**

Relationship Between Class Selection and the Congruence of Student Ratings with Faculty Self-Perceptions Regarding Evaluation Procedures

<table>
<thead>
<tr>
<th>Discrepancy of Elective Class</th>
<th>Discrepancy of Required Class</th>
<th>df</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students vs. Instructor</td>
<td>Students vs. Instructor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N: 532</td>
<td>2247</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M: .586</td>
<td>.632</td>
<td>2777</td>
<td>1.462</td>
<td>.14</td>
</tr>
<tr>
<td>SD: .602</td>
<td>.660</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Examination of the results shown in Table 14 indicates that no significant relationship exists between the basis for class selection and the degree of congruence of student ratings and faculty self-perceptions concerning behaviors in the area of evaluation procedures. The observed relationship could have occurred 14 times out of 100. On this basis of these findings, the null hypothesis is not rejected.
Ho_{12}: No difference exists between the mean discrepancy score of ratings assigned by students enrolled in required classes versus instructors' self-perceptions, and the mean discrepancy score of ratings assigned by students enrolled in elective classes versus instructors' self-perceptions regarding the factor student centeredness.

**TABLE 15**

Relationship Between Class Selection and the Congruence of Student Ratings with Faculty Self-Perceptions Regarding Student Centeredness

<table>
<thead>
<tr>
<th></th>
<th>Discrepancy of Elective Class Students vs. Instructor</th>
<th>Discrepancy of Required Class Students vs. Instructor</th>
<th>df</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>N:</strong></td>
<td>581</td>
<td>2405</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>M:</strong></td>
<td>.443</td>
<td>.407</td>
<td>2985</td>
<td>1.826</td>
<td>.07</td>
</tr>
<tr>
<td><strong>SD:</strong></td>
<td>.415</td>
<td>.423</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The analysis presented in Table 15 indicates that no relationship exists between type of class selection and the degree of congruence displayed by student ratings and faculty self-perceptions regarding student centeredness behaviors. These differences could be observed seven times in 100. This level of significance is not sufficient to reject the null hypothesis.
**Ho13:** No difference exists between the mean discrepancy score of ratings assigned by students versus the self-perceptions of instructors with more experience and the mean discrepancy score of ratings assigned by students versus the self-perceptions of instructors with less teaching experience regarding the factor professional competence.

**TABLE 16**

Relationship Between Instructor Experience and the Congruence of Student Ratings with Faculty Self-Perceptions Regarding Professional Competence

<table>
<thead>
<tr>
<th></th>
<th>Discrepancy of High Experience Instructors vs. Students</th>
<th>Discrepancy of Low Experience Instructors vs. Students</th>
<th>df</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>N:</td>
<td>2180</td>
<td>815</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M:</td>
<td>.401</td>
<td>.371</td>
<td>2993</td>
<td>1.77</td>
<td>.08</td>
</tr>
<tr>
<td>SD:</td>
<td>.437</td>
<td>.351</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

According to the results of the analysis presented in Table 16 no relationship exists between instructors' experience and the degree of congruence between student ratings and instructor self-ratings regarding professional competence behaviors. The statistical strength of this observation is insufficient to reject the null hypothesis since it could occur by chance eight times in 100.
H0: No difference exists between the mean discrepancy score of ratings assigned by students versus the self-perceptions of instructors with more experience and the mean discrepancy score of ratings assigned by students versus the self-perceptions of instructors with less teaching experience regarding the factor evaluation procedures.

### TABLE 17

<table>
<thead>
<tr>
<th>Discrepancy of High Experience Instructors vs. Students</th>
<th>Discrepancy of Low Experience Instructors vs. Students</th>
<th>df</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>N: 2009</td>
<td>778</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M: .670</td>
<td>.508</td>
<td>2785</td>
<td>5.898</td>
<td>.001*</td>
</tr>
<tr>
<td>SD: .682</td>
<td>.563</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

An examination of the results of the data shown in Table 17 appears to indicate a very strong relationship between instructor experience and the degree of congruence between his self-perceptions and students ratings of his behaviors regarding evaluation procedures. Again, it seems that instructors with lower experience are more in agreement with their self-perceptions measured against student ratings than are instructors with greater experience. The observed differences could occur only once in more than 1,000. Therefore, the null hypothesis is rejected.
Ho\textsubscript{16}: No difference exists between the mean discrepancy score of ratings assigned by students versus the self-perceptions of instructors with more experience and the mean discrepancy score of ratings assigned by students versus the self-perceptions of instructors with less teaching experience regarding the factor student centeredness.

TABLE 18

Relationship Between Instructor Experience and the Congruence of Student Ratings with Faculty Self-Perceptions Regarding Student Centeredness

<table>
<thead>
<tr>
<th>Discrepancy of High Experience Instructors vs. Students</th>
<th>Discrepancy of Low Experience Instructors vs. Students</th>
<th>df</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>N: 2182</td>
<td>814</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M: .433</td>
<td>.363</td>
<td>2994</td>
<td>4.024</td>
<td>.001*</td>
</tr>
<tr>
<td>SD: .428</td>
<td>.406</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Based on the results presented in Table 18, it is apparent that a strong relationship does exist between instructor experience and the degree of congruence between his self-perceptions and student ratings of his behaviors regarding student centeredness. Once more, the self-ratings of instructors with lower experience were much more in agreement with student ratings of student centered behaviors than were those with greater experience. These results could be observed less than one time in more than 1,000. The null hypothesis is rejected on the basis of these findings.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Summary

As has often been the case, the findings of this study present varying evidence. Not only have the various independent variables appeared to affect the discrepancy between student ratings and instructors' self-perceptions, but considerable differences were also noted according to the type of behavior being assessed.

Interestingly, no significant differences were noted in Tables 3, 4 and 5 between student ratings and instructor self-perceptions. However, while students and faculty were in close agreement regarding the assessment of Professional Competence and Student Centeredness behaviors, there was some discrepancy on Evaluation Procedures. The $t$ value of this difference was significant at .08. Further, the two sets of ratings displayed little tendency to vary together. On no occasion did the correlations reach a magnitude worthy of note.

Student grade point average appeared to have little effect on the congruence between student ratings and the instructor's self-perceptions regarding his classroom behaviors. Only on the assessment of behaviors regarding Professional Competence was the variance between discrepancy scores at a statistically significant level. Perceptions of Evaluation Procedures and Student Centeredness varied, but failed to achieve enough significance to reject the null hypotheses.

The size of the class in which the instructional interaction took place appeared to have a highly significant impact on the
congruence of student assessments and faculty self-perceptions regarding Professional Competence and Student Centeredness behaviors. In regard to both factors, the differences between the small and large classes were of such a size as to indicate chance occurrence only one time in more than 1,000. No such effect was noted on the assessment of Evaluation Procedures. Instructors' and student ratings displayed considerable congruence on the variable class size while measuring these behaviors.

No significant differences occurred in the discrepancy scores between students and instructors according to whether the student was enrolled on an elective or required basis. The differences noted in these discrepancy scores could have occurred by probability in the following order: Professional Competence .08, Evaluation Procedures .14, and Student Centeredness .07.

The independent variable, teaching experience related to instructors, also appeared to have considerable impact on the discrepancy levels between students and faculty as they perceived at least two types of behaviors.

Highly statistically significant differences in the discrepancies between students versus instructors with high experience and students versus instructors with low experience regarding Evaluation Procedure and Student Centeredness behaviors were evident in the findings. While not of statistical magnitude to reject the null hypothesis, the differences observed in the assessment of Professional Competence behaviors could have occurred with a
probability of only eight in 100.

The evidence was inconsistent, but it does appear that certain characteristics of students and instructors alter the manner in which students rate and faculty self-perceive their classroom behaviors. Viewing the results of the data analyses, a tendency is noted for the differences in discrepancies between student and faculty assessments to be greatest on the measurement of Professional Competence and Student Centeredness. Yet, the findings show in the analyses for null hypotheses one through three in which faculty and student characteristics are disregarded, that the differences between ratings are very slight.
CHAPTER V

SUMMARY, CONCLUSIONS, AND IMPLICATIONS

Summary

A review of selected literature indicated mixed and sometimes conflicting evidence regarding the various aspects of student assessment of faculty teaching behaviors. Generally, it is concluded, however, that students can reliably and with validity measure these behaviors. Several writers have suggested the need for, and value of, self-assessment by higher education instructors. Some have indicated that they believed this method to be the only assessment which could provide for the improvement of teaching. However, there have been no reported studies which investigated the question of agreement between student ratings of faculty classroom behaviors and the way faculty members themselves perceive these behaviors.

The purpose of this study was to investigate, in a field situation, the nature of the relationships between student ratings of faculty classroom behaviors and the reported self-perceptions regarding these behaviors by the faculty members. Two types of data were collected. Students responded to the Student Opinions About Instructional Procedures in terms of the Classroom behaviors of their instructors. Faculty members recorded their self-perceptions of these behaviors on a virtually identical instrument.
The responses recorded were developed into factor scores for the three factors which the instrument displays. These factors were entitled: (1) Professional Competence, (2) Evaluation Procedures, and (3) Student Centeredness. Since the main interest of this study was the degree of congruence between student ratings and faculty self-perceptions, following overall comparisons, discrepancy scores were generated by comparing each instructor's factor scores to each of his students. This process produced a factor discrepancy value for each factor for each pairing of student and instructor.

These data were analyzed using $t$ tests on the basis of the dichotomies created by the independent variables. The independent variables were as follows: (1) student grade point average, (2) size of class, (3) types of class selection, and (4) amount of instructor experience. As $t$ values were developed, they were described by the precise probability levels at which they could occur by chance. Although all $t$ values were described by probabilities, the .05 level was necessary in order to reject the null hypotheses.

Fifty-eight instructors in the Teacher Education Department, College of Education at Western Michigan University elected to participate in the study. This sample represented 95 percent of the total faculty population. Three thousand forty-eight students in 135 classes represented the student sample. These students were in Teacher Education Department classes taught by participating instructors, and made up approximately 96 percent of the total.
student population enrolled in Teacher Education Department classes.

The basic instrument used in the study was the Student Opinion About Instructional Procedures (SOAIP) developed earlier by an Ad Hoc Committee of the Teacher Education Department. The instrument was altered only in pronoun form to adapt it for use as a device to record faculty self-perceptions. A factor analysis of this instrument identified three main factors which together accounted for 57 percent of the variance. These three factors, which formed the basis of the data analyses of this study, have been labeled professional competence, evaluation procedures, and student centeredness.

Conclusions

Discussion of the results of the analyses were organized according to the independent variables which were investigated. In order of examination, these variables were: (1) students and faculty, (2) grade point average, (3) class size, (4) types of class selection, and (5) instructor experience.

Student Ratings - Faculty Self-Perceptions

On two of the behavioral factors under examination, students rated instructors and instructors rated themselves in a very similar manner. These two factors were Professional Competence and Student Centeredness. It would appear from these findings that the two groups involved perceive these groups of behaviors in a like manner, and that faculty members can self-evaluate these behaviors accurately.
when student opinions are used as the basic criteria.

The differences between the opinions of the two groups in regard to Evaluation Procedures were not great enough to accept the hypothesis. However, there does seem to be some difference in the perceptions of students and faculty as this set of behaviors is considered. If student ratings are used as the validity check, instructors apparently do not do as well in evaluating this area of their classroom behaviors.

Grade Point Average

The analyses for the reported grade point average and following independent variables were carried out on the discrepancies between instructor self-perceptions and student ratings. These discrepancy scores were determined on the basis of the absolute differences between the opinions and were not concerned with the direction, but with the magnitude of these differences.

Based on the results of the analysis regarding perceptions of Professional Competence behavior, it would appear that higher achieving students and instructors are more congruent in their opinions than are lower achieving students and instructors. As a faculty member attempts to compare his self-evaluations with the opinions of his students, he should apparently be cognizant of their levels of achievement.

Differences were apparent in the analyses concerned with Evaluation Procedures and Student Centeredness, but were not great enough
to accept the hypotheses. Higher achieving students and instructors were again more congruent in their opinions than were lower achieving students and instructors, but not at the level displayed in regard to Professional Competence.

Class Size

The number of students enrolled in the class appeared to have some impact on the discrepancies between student rating and instructor self-perceptions as Professional Competence and Student Centeredness behaviors are concerned. However, there was apparently no effect on the discrepancy concerning Evaluation Procedures.

Interestingly, students from large classes and their instructors appeared to be more congruent in their opinions than those involved in small class interaction. Again, as faculty attempt to self-evaluate and to check their evaluation against student opinions concerning Professional Competence and Student Centeredness behaviors, they apparently should take the size of the class into consideration.

Type of Class Selection

The effect of the type of class selection (elective/required) on the discrepancies between student-ratings and instructor self-perceptions appeared to vary with the behavior set being considered. Discrepancies between students in elective classes and their instructors were not of a size to permit acceptance of the hypotheses.
A higher degree of congruence appeared to exist between the opinion of students in required classes and instructors than between students in elective classes and instructors on Professional Competence and Student Centeredness behaviors. The reverse was true regarding Evaluation Procedures. In this case, students in elective classes were more congruent in their opinions with their instructors.

Similarly, while the differences observed were not great enough to accept the hypothesis, instructors checking their self-evaluations might want to consider the type of class selection under which their students were enrolled.

**Instructor Experience**

The effect of the level of instructor experience upon the discrepancies between student and instructor opinions, appears to be considerable. Differences noted in regard to behaviors concerned with Evaluation Procedures and Student Centeredness indicated that the hypotheses should be accepted. In both instances, opinions of instructors with lower experience were more congruent with student ratings than were the opinions of instructors with higher experience. Differences regarding Professional Competence behaviors did not permit the acceptance of the hypothesis. However, opinions of instructors with lower experience appeared to be more congruent with their students' opinions than were those instructors with more experience.

When working with self-evaluation procedures, it appears that
instructors should also keep their own experience level in mind. This is true at least if student opinions are to be held as the validity check.

Implications

The initial thought behind this study was to attempt to provide a tool to ultimately be used in the improvement of teaching in higher education. Through comparing the perceptions of those intently involved in the interaction of teaching and learning, it was believed that the feasibility of self-evaluation could be demonstrated. Based on the findings of the data analyses, it has been shown that by and large, instructors can accurately self-evaluate their behaviors when student ratings are used as the validity check.

Student ratings of instructors in higher education have been reported to be both reliable and valid. In the opinion of this writer, these types of assessments combined with faculty self-ratings represent the only valid measures of the classroom interaction as far as the improvement of teaching is concerned.

Beyond the presentation of a little used evaluation tool this study also attempted to investigate additional variables in order to extend awareness regarding the concept. In order to achieve this goal, various characteristics of students and faculty were treated as independent variables and used to create dichotomies upon which the analyses were based. These findings extended knowledge of the various facets of the concept considerably by revealing that

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
certain characteristics of both groups are areas to be treated with concern as this type of self-evaluation was conducted. For example, it was shown that the instructor's teaching experience could have a highly significant impact upon the degree of congruence between his and his students' ratings.

This tool is a very simple extension of the existing use of student feedback. The additional self-evaluation can provide a focus point for instructors upon discrepancies between their own and their students' perceptions. These discrepancies may or may not actually describe shortcomings in the instructor's behavior; that is for him/her to decide. The psychology literature, however, indicates that behavior is generally changed as the need for change is seen. The assessment process which is the basis of this study, increases the opportunity for instructors to make the decision regarding the modification of classroom behaviors.

Recommendations

In view of the relatively high correlations between factors as presented in Table 6 it would seem wise to factor analyze once more the SOAIP using the data collected in this investigation. There is some question as to whether correlations of this magnitude should occur given the results of previous factor analyses. It may be that the instrument does not measure the three factors to the degree that had been believed.

Certainly there are other characteristics of both students and
faculty that could be explored in a replication of this study. It is possible that the analyses of instructor experience might actually have been measuring an age factor. Additional work would clarify the answers. Student and faculty characteristics such as sex, lifestyle, and personality dimensions are all factors which could conceivably have an impact upon the evaluation process. In addition it is possible that a study of the manner in which graduate and undergraduate students view the behaviors of instructors would reveal data which might further clarify this evaluation procedure. It is strongly felt that additional work in this area is necessary as we begin to convert the knowledge gained into projects designed to assist in the improvement of teaching in higher education.
REFERENCES


Bendig, A. W. The use of student rating scales in the evaluation of instructors in introductory psychology. *Journal of Educational Psychology*, 1952, 43, 167-175.


Bendig, A. W. Relations of level of course achievement of students or instructor, and course ratings in introductory psychology. *Educational and Psychological Measurement*, 1953, 13, 437-448.


Daw, R. W., George, N. L. Effect of feedback from principals to teachers. *Journal of Educational Psychology*, 1967, 58, 181-188.


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


Hudelson, E. The validity of student rating of instructors. *School and Society*, 1951, 73, 265-266.


Russell, H. E., Bendig, A. W. Investigation of the relations of student ratings of psychology instructors to their course achievement when academic aptitude is controlled. Educational and Psychological Measurement, 1953, 13, 626-635.


INSTRUCTIONS:

You should have the following materials in order to complete this instrument.

1. This booklet containing instructions and items to which you will have to respond.

2. Answer sheet on which your responses will have to be recorded.

3. A number two lead pencil (or softer pencil) which is to be used for recording your responses on the answer sheet.

The most important overall purpose of this questionnaire is to improve instruction. The responses to this questionnaire or summary data of the responses will be made available to the instructor of this class and the chairman of the department in which your instructor is working. To insure that your identity remains anonymous, please DO NOT place your name or any other identifying marks on the response sheet.

Please, do not make any marks in this booklet.

DIRECTIONS:

A. Write in the box after "Course Number" the identification number of this course, e.g., A. Course Number. Do not make any marks outside of the box.

B. Write on the answer sheet in the box after "Instructor's Name" the name of the instructor in this course, e.g., B. Instructor's Name.

If there are several instructors for this course, fill out a separate answer sheet for each instructor.

Do not make any other marks on the answer sheet except the ones necessary to respond to any item.
GENERAL INFORMATION (items 1-11)

For each of the following eleven items, pick the response appropriate to you or to the present situation. Note the code to the left of the appropriate response. Locate the item number on the answer sheet and shade in the area between the two dotted lines which correspond to this code. Only one response may be chosen for each item. (The outline of the rectangular box in the upper right hand corner of the answer sheet has no special meaning.)

If you are unable to respond to an item, you should make no mark after that item's number on the answer sheet.

Example:

If the present calendar year is 1970, the code for this year is 0 as is indicated in item 1 in this booklet and this should be recorded on the answer sheet in the following manner:

1. 0 = 0 = 1 = 2 = 3 = 4 = 5 = 6 = 7 = 8 = 9 =

1. The present calendar year is:

0. 1970
1. 1971
2. 1972
3. 1973
4. 1974

2. The present semester (or session) is:

0. Winter (January-April)
1. Spring (April-June)
2. Summer (June-August)
3. Fall (August-December)

3. The size of this class is:

0. under 10
1. 11-30
2. 31-50
3. over 50

4. I am a:

0. freshman
1. sophomore
2. junior
3. senior
4. graduate student in a masters degree program
5. graduate student in a doctoral program
6. student in a non-degree program

5. I am:
   0. planning to teach at the elementary school level
   1. planning to teach at the junior high school level (or middle school level)
   2. planning to teach at the high school level
   3. planning to teach at above high school levels
   4. planning to work at educational institutions but not teach
   5. not planning to work for educational institutions

6. This course is:
   0. a required course for me (I had no choice but take this course)
   1. an elective (there is at least one other course in the University I could have taken in place of this one)

7. I am a:
   0. male
   1. female

8. I am a:
   0. part-time student
   1. full-time student

9. My undergraduate or graduate grade point average to date is:
   0. under 1.0
   1. 1.0 - 1.999
   2. 2.0 - 2.999
   3. 3.0 - 4.0
   4. not established

CRITERIA FOR INSTRUCTION: (items 10-28)

After carefully reading each of the nineteen items on the next page, rate your instructor or the instructional environment on the characteristic described in the statement using the following categories and codes.

0. The characteristic described is almost never present.
1. The characteristic described is infrequently present.
2. The characteristic described is frequently present.
3. The characteristic described is almost always present.
4. I am undecided.

If an item does not apply to your situation, you should not respond to that item. Response 4: "I am undecided" should not be used for non-applicability of an item.

Then, on the response sheet, darken the area between the pair of dotted lines that corresponds to the code (0,1,2,3,4) representing the category you have selected for that item. Do not make more than one mark for each item.

EXAMPLE:

0. The instructor is on time for class.

If, from your experience, you can recall a few times that the instructor was late, you would probably mark the response sheet to represent that the characteristic described was frequently present.

0. =0= =1= =2= =3= =4= =5= =6= =7= =8= =9=

Complete the following items:

10. The instructor evaluates my achievement fairly.
11. The instructor responds to my papers with comments.
12. The instructor uses tests which require understanding of the material.
13. The test content is consistent with the materials in the course.
14. The instructor listens with concern to my views and ideas.
15. The instructor encourages questions.
16. The instructor creates an atmosphere of respect for viewpoints and opinions of students.
17. The instructor recognizes me as an individual.
18. The instructor has shown a willingness to help me after class.
19. The instructor is helpful when I have difficulty.
20. The instructor appears sensitive to my feelings and problems.
21. I can easily arrange outside appointments with the instructor.
22. The instructor is patient with me.

23. The instructor shows a willingness to try new instructional procedures.

24. The volume of the instructor's speech is loud enough for me to hear him.

25. For my purposes the instructor offers satisfactory explanations and illustrations for major points.

26. The instructor provides for active class participation.

27. The instructor demonstrates an understanding of contemporary developments in his field.

28. The instructor appears sincerely interested in teaching this course.

Please, return this booklet and the answer sheet(s) to the individual administering this instrument.
INSTRUCTOR SELF-PERCEPTIONS ABOUT INSTRUCTIONAL PROCEDURES

College of Education
Western Michigan University

INSTRUCTIONS:

You should have the following materials in order to complete this instrument.

1. **This booklet** containing instructions and items to which you will have to respond.

2. **Answer sheet** on which your responses will have to be recorded.

3. **A number two lead pencil** (or softer pencil) which is to be used for recording your responses on the answer sheet.

The most important purpose of this questionnaire is to improve instruction. Neither the responses to this questionnaire nor summary data of the responses will be available to anyone other than the investigator of this study.

Please, do not make any marks in this booklet.

DIRECTIONS:

A. Write in the box after "Course Number" the identification number and section of this course. Do not make any marks outside of the box. 

   TEED 250 CA

B. Write your name on the answer sheet in the box after "instructor's name."

   Les Armstrong

Do not make any other marks on the answer sheet except the ones necessary to respond to an item.
GENERAL INFORMATION (items 1-9)

For each of the following nine items, pick the response appropriate to you or to the present situation. Note the code to the left of the appropriate response. Locate the item number on the answer sheet and shade in the area between the two dotted lines which correspond to this code. Only one response may be chosen for each item. (The outline of the rectangular box in the upper right hand corner of the answer sheet has no special meaning.)

If you are unable to respond to an item, you should make no mark after that item's number on the answer sheet.

Example:

If the present calendar year is 1972, the code for this year is 2 as is indicated in item 1 in this booklet and this should be recorded on the answer sheet in the following manner:

1. =0= =1= =2= =3= =4= =5= =6= =7= =8= =9=

1. The present calendar year is:

0. 1970
1. 1971
2. 1972
3. 1973
4. 1974

2. The present semester (or session) is:

0. Winter (January-April)
1. Spring (April-June)
2. Summer (June-August)
3. Fall (August-December)

3. The size of this class is:

0. under 10
1. 11-30
2. 31-50
3. over 50

4. The institutional rank I hold is:

0. Instructor
1. Assistant Professor
2. Associate Professor
3. Professor

5. The highest degree I hold is:
   0. BA-BS
   1. MA-MS
   2. Ed.S.
   3. Ed.D.-Ph.D.

6. In rank ordering courses I prefer to teach this course is my:
   0. First choice
   1. Second choice
   2. Third choice

7. I am a:
   0. Male
   1. Female

8. I am a:
   0. Part-time instructor
   1. Full-time instructor

9. The number of years I have taught in higher education is:
   0. Less than five years
   1. Five years or more

CRITERIA FOR INSTRUCTION: (item 10-28)

After carefully reading each of the nineteen items on the next page, rate yourself or your instructional environment on the characteristic described in the statement using the following categories and codes.

0. The characteristic described is almost never present.
1. The characteristic described is infrequently present.
2. The characteristic described is frequently present.
3. The characteristic described is almost always present.
4. I am undecided.

If an item does not apply to your situation, you should not respond to that item. Response 4: "I am undecided" should not be used for non-applicability of an item.

Then, on the response sheet, darken the area between the pair of dotted lines that corresponds to the code (0, 1, 2, 3, 4) representing
the category you have selected for that item. Do not make more than one mark for each item.

Example:

0. I am on time for class.

If, from your experience, you can recall a few times that you were late, you would probably mark the response sheet to represent that the characteristic described was frequently present.

0. =θ= =½= =2= =3= =4= =5= =6= =7= =8= =9=

Complete the following items:

10. I evaluate student achievement fairly.
11. I respond to student papers with comments.
12. I use tests which require understanding of the material.
13. The test content is consistent with the materials in the course.
14. I listen with concern to student views and ideas.
15. I encourage questions.
16. I create an atmosphere of respect for viewpoints and opinions of students.
17. I recognize students as individuals.
18. I have shown a willingness to help students after class.
19. I am helpful when students have difficulty.
20. I have displayed sensitivity to student feelings and problems.
21. Students can easily arrange outside appointments with me.
22. I am patient with students.
23. I have shown a willingness to try new instructional procedures.
24. The volume of my speech is loud enough for students to hear me.
25. I offer satisfactory explanations and illustrations for major points.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
26. I provide for active class participation.

27. I demonstrate an understanding of contemporary development.

28. I have displayed sincere interest in teaching this course.

Please, return this booklet and the answer sheet(s) to the individual administering this instrument.
March 29, 1972

Dear,

Work is about to begin on the most comprehensive study ever carried out in the area of faculty self-evaluation at Western Michigan University. The plan is to supplement the student evaluation of instructional procedures with a study of instructor perceptions. This study calls for a comparison of the way you perceive your own instructional procedures to the way they are perceived by your students. Your participation will strengthen any conclusions drawn from this research.

I would estimate that the amount of your time required will be less than five minutes for each class for which you are responsible. Enclosed in the envelope is a brief abstract of the study plus the questionnaire to which I am asking you to respond. In addition, you will find one answer sheet for each class you are teaching. Would you please carefully and thoughtfully complete an answer sheet for each of your classes. Your response to this request will be viewed as permission to utilize student data about your procedures for the purposes of this study. All data will be processed by myself, and at all times your anonymity will be protected. Your self-perceptions will be available to no one other than myself.

This research has been designed in cooperation with my doctoral committee chaired by Dr. Dorothy McCuskey; Dr. Kenneth Dickie, Acting Chairman of the Teacher Education Department; and Dr. Uldis Smidchens, Director of the Center for Educational Research.

Please return your responses to me either through the Teacher Education Office or via inter-office mails as soon as possible, but no later than April 12. In closing, please accept my deepest appreciation for your cooperation. If you would like copies of the data concerning yourself and your students please send me a note and I shall be more than happy to provide you with this information.

Sincerely,

Eugene W. Thompson
Graduate Associate
Center for Educational Research
3425 Sangren 3-1645
Home Phone: 345-4043

Approved for Distribution

Dorothy McCuskey
Uldis Smidchens
ABSTRACT

Problem and Objectives

Evaluation and improvement of teaching in higher education especially in teacher education continues to draw the interest of various groups concerned with this area. This study is based in the belief that the only groups in a position to form reliable opinions regarding teaching procedures are the student-consumers and the instructors themselves. In utilizing these opinions the assumption is that instructors in higher education can effectively evaluate their own behaviors without management-type intervention. It is felt that systematic self evaluations matched against those of student-consumers will provide comparative data which can provide the basis for change. The intent of this study then, is to investigate the relationships between student opinions and instructors' self-perceptions of classroom behaviors.

Procedures

The instructors selected to participate in this study will be from the Teacher Education Department at Western Michigan University. The data to be utilized will be that provided by student responses to the annual department administration of the Student Opinions About Instructional Procedures. In addition, faculty members will record their self-perceptions of these same procedures on a slightly modified version of the same instrument. Several other variables such as the degree of the instructor, years taught, etc. will also be examined. Since the questionnaire has been shown to be essentially a three factor instrument, data from each responding group (students-instructors) will be analyzed using a t-ratio on the basis of each factor. Pearson Product-Moment Correlations will provide a secondary analysis of the degree to which these evaluations vary together.

Significance

To date no major efforts have been discovered which have considered the utilization of instructor self-evaluation. An examination of the discrepancies will serve to focus attention upon variance in opinion and as a result provide the basis for any desired change. This type of assessment upon the part of those intimately involved will provide additional protection for the basic constructs of classroom freedom and responsibility.
ANSWER SHEET
Student Opinions
About Instructional Procedures

A. Course number and section

B. Instructors name

C. General information (items 1-9)
For each item darken in the area that corresponds to the digit in front of the response appropriate to your situation. Mark One response for each item.

D. Information related to instructional procedures (items 10-28).
After reading the item in the booklet rate your instructor or the instructional environment on the characteristic described in the statement. Use the following categories and codes:

0 - THE CHARACTERISTIC DESCRIBED IS ALMOST NEVER PRESENT.
1 - THE CHARACTERISTIC DESCRIBED IS INFREQUENTLY PRESENT.
2 - THE CHARACTERISTIC DESCRIBED IS FREQUENTLY PRESENT.
3 - THE CHARACTERISTIC DESCRIBED IS ALMOST ALWAYS PRESENT.
4 - I AM UNDECIDED.

If an item does not apply to your situation you should not respond to that item. Response 4: "I AM UNDECIDED" should not be used for nonapplicability of an item. Darken in the area between the pair of dotted lines that corresponds to the code (0,1,2,3, or 4) of the category you have chosen. Do not mark more than one response for each item.
March 1, 1972

Attention: Faculty, Teacher Education Department

At the end of the Winter Semester, all faculty members should plan to have the "Student Opinions About Instructional Procedures" Instrument administered to their classes. Materials will be available on April 3, 1972 in the Teacher Education Office, 2112 Sangren. Those individuals whose classes will be terminated prior to that time, should make arrangements to pick up their materials at an earlier date.

Please follow these instructions in administering the evaluation:

1. Identify a student in each of your classes who will assume responsibility for:
   a. stopping at the Teacher Education Office, 2112 Sangren, to pick up the Evaluation package
   b. supervising the administration of the Questionnaire to the class
   c. collecting and returning the completed Questionnaire to 2112 Sangren

2. Allow approximately twenty minutes of class time for the evaluation.

3. The faculty member should leave the room during the evaluation and is encouraged not to make any comments to the class about the evaluation instrument. The designated student should do the explaining.

4. The evaluation packages will be available any time after Monday, April 3, 1972, between the hours of 8:00-12:00 noon, and 1:00-5:00 P.M. If your class is held in the evenings or Saturdays and it is inconvenient for the designated student to pick them up during the above mentioned hours, the faculty member may pick the materials up for the student.

5. Give the student the instruction sheet enclosed with this memorandum in advance to the administration of the instrument.

6. The student should return all materials (booklets and answer sheets) to the Teacher Education Office as soon as possible.
7. Please return the attached information forms to the Teacher Education Office before the "Student Opinions About Instructional Procedures" Instrument is administered in your classes.

8. Your cooperation in completing this activity will be appreciated.

This information will be processed as soon as possible and you will receive two copies of summary information about your student responses. One of these copies, if you desire, should be returned to the Department Chairman.

Kenneth E. Dickie
Acting Chairman
INSTRUCTIONS
TO THE INDIVIDUAL ADMINISTERING THE
"STUDENT OPINIONS ABOUT INSTRUCTIONAL PROCEDURES" INSTRUMENT

A. The number in the upper right-hand corner of the envelope indi­
cates the number of instruction booklets with answer sheets in
the envelope. If you need more instruction booklets and
answer sheets, request them from the secretary in Room 2112
Sangren.

B. The instructor(s) should not be present during the adminis­
tration of this instrument.

C. Indicate on black board three items to be recorded on answer
sheet in this manner:

1. =8= =4= =8= =8= (Calendar Year)
2. =8= =4= =8= =8= (Semester
c. =8= =4= =8= =8= (Class enrollment—Check with
instructor)

D. Read the following statement to the group of individuals to
whom you are administering this instrument:

1. If several instructors are teaching this section, you should
complete a separate answer sheet for each instructor.

2. Read carefully and follow the instructions in the booklet
which you should have at this time.

3. Use #2 pencils only to record your response.

4. You should have one answer sheet for each instructor teach­
ing this section. In most cases there is only one instructor
teaching a section. Therefore, in most cases you should
have only one answer sheet.

5. After you have responded to all items, the answer sheet(s)
and the instruction booklets should be returned to the
individual administering this instrument.

6. Please use comment sheet if you have any specific statements
to make about this course or instructor not covered in the
instrument provided. The instructor will not receive the
comments until sometime after the completion of the semester.
E. If there are several instructors teaching this section, separate the answer sheet according to the instructor who is being evaluated. Use a separate envelope for each instructor. Be sure that the correct name is on the envelope.

F. To return the completed answer sheets, unused sheets, and the instruction booklets either a) return personally to Teacher Education Office, 2112 Sangren between 8-12 and 1-5 or mail to:

Ms. Linda Schmitt
Graduate Assistant
Department of Teacher Education
2112 Sangren Hall
Western Michigan University
Kalamazoo, Michigan 49001
Dear TEED Faculty Member:

Enclosed are data sheets for each of your Winter 1972 classes which responded to the Student Opinions About Instructional Procedures Questionnaire and for which you provided self-perceptions.

As you might be aware, the questionnaire displays three major factors which have been labeled: (1) Professional Competency, (2) Evaluation Procedures, and (3) Student Centeredness. The scores presented in your sheets are mean scores for these three factors. The three categories presented are (1) Instructor Means (your self-perceptions), (2) Departmental Means (your colleagues' self-perceptions), and (3) Student Means (your students' ratings).

The rating scale used reads as follows:

0. The characteristic described is almost never present.
1. The characteristic described is infrequently present.
2. The characteristic described is frequently present.
3. The characteristic described is almost always present.

I do sincerely appreciate the outstanding cooperation given me by yourself and your departmental colleagues. Without your assistance this study, of course, could not have been completed. With considerable faculty help, the project is progressing nicely and should be completed in the very near future.

Sincerely,

Eugene W. Thompson
Graduate Associate
Center for Educational Research
College of Education
Western Michigan University
## SAMPLE FEEDBACK DATA

<table>
<thead>
<tr>
<th>Instructor ID 999</th>
<th>Course Number 250</th>
<th>Section RA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Professional Competency</td>
<td>Evaluation Process</td>
</tr>
<tr>
<td>Instructor Means</td>
<td>2.71</td>
<td>3.00</td>
</tr>
<tr>
<td>Departmental Means</td>
<td>2.66</td>
<td>2.53</td>
</tr>
<tr>
<td>Student Means</td>
<td>2.91</td>
<td>2.83</td>
</tr>
</tbody>
</table>