

1986

Student Characteristics And Student Evaluations Of College Teaching: A Within-class Analysis

Stephen William Erdle

Follow this and additional works at: <https://ir.lib.uwo.ca/digitizedtheses>

Recommended Citation

Erdle, Stephen William, "Student Characteristics And Student Evaluations Of College Teaching: A Within-class Analysis" (1986).
Digitized Theses. 1547.
<https://ir.lib.uwo.ca/digitizedtheses/1547>

This Dissertation is brought to you for free and open access by the Digitized Special Collections at Scholarship@Western. It has been accepted for inclusion in Digitized Theses by an authorized administrator of Scholarship@Western. For more information, please contact tadam@uwo.ca, wlsadmin@uwo.ca.



National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service

Services des thèses canadiennes

Ottawa, Canada
K1A 0N4

CANADIAN THESES

THÈSES CANADIENNES

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30.

**THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED**

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30.

**LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE**

STUDENT CHARACTERISTICS AND STUDENT EVALUATIONS OF
COLLEGE TEACHING: A WITHIN-CLASS ANALYSIS.

by

Stephen Erdle

Department of Psychology

Submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy

Faculty of Graduate Studies
The University of Western Ontario
London, Ontario
September 1986

© Stephen Erdle 1986

17	Correlations Between Student Personality Dimensions and Overall Effectiveness Ratings	58
18	Between-Class Differences in Mean Scores for Student Personality Dimensions	62
19	Correlations Between Perceived Student-Teacher Personality Difference Scores and Overall Effectiveness Ratings	63
20	Correlations Between Perceived Student-Teacher Personality Similarity and Overall Effectiveness Ratings	68

Abstract

Research examining the relation between student characteristics and individual student evaluations of college teaching has often pooled data across classes, thereby mixing between- and within-class variability and allowing the possibility of obscuring differential within-class relationships. To avoid these difficulties, the present study examined associations between student characteristics and evaluations of teaching within separate classes. Each of 278 students from six undergraduate classes completed a questionnaire assessing perceptions of instructional effectiveness and teacher personality, as well as their own personality, interest in the course, gender, university level, grades, perceptions of amount learned, likelihood of future course enrollment, and attitude toward university instruction. Averaged across classes, individual student evaluations were found to correlate with: other subjective indices of teaching effectiveness (ratings of amount learned and likelihood of future enrollment); perceived teacher nurturance, achievement, orderliness, and neuroticism (negatively related); perceived positive student-teacher dissimilarity in nurturance, extraversion, achievement, orderliness, and neuroticism; interest in the course; university level; and attitudes toward university instruction. Averaged across classes, however, individual student evaluations failed to correlate significantly with: perceived student-teacher

similarity in personality; students' university grade average; course grades; perceived grading leniency; gender; area of major; or student personality. The relationship between student personality and evaluations, however, was found to differ across classes for 3 of 5 dimensions, demonstrating the utility of within-class analyses. Moreover, these differences were found to be related to differences between classes in mean perceptions of relevant teacher personality traits. Certain types of students seemed to find certain types of instructors to be most effective; whereas other types of students seemed to find other types of instructors to be most effective. Findings were interpreted as evidence for the validity of student ratings of instructional effectiveness and discussed in terms of understanding and improving college instruction.

ACKNOWLEDGEMENTS

First, I would like to thank my thesis supervisor, Dr. J. P. Rushton, as well as Dr. H. G. Murray and Dr. R. C. Gardner for their constant advice and support throughout my doctoral studies. I also would like to thank the members of my thesis advisory committee and examining board, Dr. D. N. Jackson, Dr. W. Fisher, Dr. W. J. McKeachie, and Dr. N. C. Baird for their helpful comments and suggestions on earlier versions of this thesis. Finally, I must also thank all those teachers, students, friends, and family who either directly or indirectly contributed to this research, especially Josef and Margaret Erdle, my parents, without whom none of this would have been possible.

TABLE OF CONTENTS

	Page
CERTIFICATE OF EXAMINATION	11
ABSTRACT	111
ACKNOWLEDGEMENTS	v
TABLE OF CONTENTS	vi
LIST OF TABLES	vii
CHAPTER 1 - INTRODUCTION	1
1.1 Introduction	1
1.2 Reliability of class average student evaluations	3
1.3 Validity of class average student evaluations ...	6
1.4 Reliability of individual student evaluations ..	11
1.5 Validity of individual student evaluations	13
CHAPTER 2 - METHOD	23
2.1 Subjects	23
2.2 Measures	25
CHAPTER 3 - RESULTS	25
3.1 Construction of composite measures	25
3.2 Relation among teaching effectiveness measures .	41
3.3 Between-class differences in mean evaluations ..	44
3.4 Evaluations and perceived teacher personality ..	46
3.5 Between-class differences in teacher personality	49
3.6 Evaluations and student interest	51
3.7 Evaluations and student grades	53
3.8 Evaluations and student gender, university level, area of major, and attitude toward instruction .	55
3.9 Evaluations and student personality	57
3.10 Evaluations and perceived student-teacher personality dissimilarity	61
3.11 Evaluations and perceived student-teacher personality similarity	66
CHAPTER 4 - DISCUSSION	70
REFERENCES	88
APPENDIX 1. SURVEY OF STUDENT PERCEPTIONS OF UNIVERSITY INSTRUCTION	96
APPENDIX 2. FEEDBACK TO SUBJECTS	108
VITA	111

LIST OF TABLES

Table	Description	Page
1	Factor Analysis of Teaching Effectiveness Items	26
2	Internal Consistency of Composite Teaching Effectiveness Measures	29
3	Correlations Among Teaching Effectiveness Factor Scores	31
4	Factor Analysis of Student Personality Traits	33
5	Internal Consistency of Composite Student Personality Measures	35
6	Factor Analysis of Perceived Teacher Personality Traits	37
7	Internal Consistency of Composite Teacher Personality Measures	38
8	Factor Analysis of Student-Teacher Difference Scores for Personality Traits	40
9	Internal Consistency of Composite Student-Teacher Personality Difference Scores	42
10	Correlations Among Composite Teaching Effectiveness Measures	43
11	Between-Class Differences in Mean Scores for Composite Teaching Effectiveness Measures	45
12	Correlations Between Perceived Teacher Personality Dimensions and Overall Effectiveness Ratings	48
13	Between-Class Differences in Mean Scores for Perceived Teacher Personality Dimensions	50
14	Correlations Between Measures of Student Interest and Overall Effectiveness Ratings	52
15	Correlations Between Student Grades and Overall Effectiveness Ratings	54
16	Correlations Between Student Gender, University Level, Area of Major, Attitude Toward Instruction and Overall Effectiveness Ratings	56

17	Correlations Between Student Personality Dimensions and Overall Effectiveness Ratings	58
18	Between-Class Differences in Mean Scores for Student Personality Dimensions	62
19	Correlations Between Perceived Student-Teacher Personality Difference Scores and Overall Effectiveness Ratings	63
20	Correlations Between Perceived Student-Teacher Personality Similarity and Overall Effectiveness Ratings	68

Student Characteristics and Student Evaluations of
College Teaching: A Within-Class Analysis.

The evaluation of university teaching effectiveness has been a topic of continuing interest to both instructors and researchers in the field of higher education. As Glasman and Gmelch (1976) point out, evaluations of teachers provide: a basis for administrative decisions concerning faculty hiring, promotion, and salary; faculty with informative feedback; students with information useful in selecting future courses; and researchers with a measure of teaching effectiveness. Of course, evaluations of teachers should be used for these purposes only when they have been demonstrated to have been made with reliable and valid assessment instruments. Centra (1977) has observed that three methods of assessing teaching effectiveness have frequently been used in North American colleges and universities: student performance, ratings by colleagues or administrators, and ratings by students.

Although it makes intuitive sense that teaching should be evaluated in terms of student learning, and that students' performance would be a logical indicant of their learning, several practical difficulties arise in the use of student performance as a measure of effective instruction. Murray (1980) summarizes the problems as follows: comparison of teachers is possible only when

courses have multiple sections or when norms are available; student ability may confound the evaluations since random assignment of students to course sections is seldom practiced; and examinations may not provide a valid measure of all the types of learning that are likely to be affected by the teaching skill of the instructor. Thus, student performance is seldom used as a measure of teaching effectiveness.

Alternatively, it has been suggested that colleagues and administrators are best qualified to evaluate teaching, since they are experienced teachers themselves. As Murray (1980) points out, however, this measure also has several drawbacks. Colleague and administrator ratings can be: threatening and disruptive to faculty morale; unrepresentative of teaching effectiveness under typical circumstances; affected by non-instructional factors such as research productivity and scholarly reputation of the instructor being rated; and impractical because of the effort involved in obtaining reliable ratings. Partly as a result of factors such as these, colleague and administrator ratings also have failed to gain wide acceptance as a measure of instructional effectiveness.

The most commonly employed method of evaluating university teaching effectiveness is student ratings. It has been suggested that students may be the best judges of teaching effectiveness, since they observe the teachers on a frequent basis under natural conditions. The practical

advantages of student evaluations coupled with the limitations associated with alternative measures of teaching effectiveness has led to a widespread use of student ratings. Surveys have reported a marked increase in the use this measure of teaching effectiveness. Thus, although Mueller (1951) found that only 35% of American universities had formal student ratings of courses, Bejar (1975) reported an estimate of 68%. Even more recently, Seldin (1980) found that more than 95% of liberal arts colleges considered student ratings in the evaluation of teaching performance.

Given the widespread use of student evaluations for the assessment of university teaching effectiveness, considerable research has been generated addressing the adequacy of this measure. The reliability of student ratings has been examined across questionnaire items, time, types of courses, and raters, and both convergent and discriminant validity have been examined. Moreover, many of these issues have been investigated at two different levels of analysis: class mean ratings, and students' individual ratings.

Reliability of class average student evaluations.

Studies of the inter-item reliability of mean student ratings have consistently reported high internal consistency coefficients (Murray, 1980). For example, Gillmore (1972) obtained coefficients ranging from .80 to

.98, and Hoffman (1978) reported estimates ranging from .68 to .94.

Since teaching ability is considered to be a relatively stable trait, mean student ratings also should have high test-retest reliability. Several studies have computed estimates of temporal stability by comparing average ratings made by students in the middle and at the end of an academic term. These studies have consistently reported high reliability coefficients. For example, Centra (1973) reported reliabilities ranging from .55 to .83, and Murray (1972) obtained an estimate of .83. Since these investigations only demonstrate high test-retest reliability over a relatively short time span, several studies have examined the stability of average student ratings over the same course taught by the same instructor in two or more successive years. Such investigations report reliabilities as high as those obtained within a single term. Guthrie (1954) found test-retest reliabilities of .87 and .89 for mean ratings of teachers in successive years, and Murray (1972, 1978, 1986) reported estimates of .66, .73, and .86 for mean ratings of teachers in the same or similar course in successive years.

The consistency of mean student ratings across different types of courses taught by the same instructor has also been examined. Using a sample of 45 instructors who had taught two different courses in the same semester,

Hogan (1973) reported reliability estimates ranging from .14 to .47, with a median of .38. Bausell, Schwartz, and Purohit (1975) computed the reliability of mean ratings for: the same course taught by the same instructor (mean $r = .69$); different courses taught by the same instructor (mean $r = .33$); the same course taught by different instructors (mean r approximately 0 for items relating to instructional skill); and different courses taught by different instructors (mean $r = .07$). In an investigation which examined the consistency of mean student ratings across both time and course types, Seiler, Weybright, and Stang (1977) reported reliability estimates of .69 and .62 for ratings of the same instructor teaching the same course in the same or successive years, respectively, and estimates of .55 and .45 for ratings of the same instructor teaching different courses in the same or successive years. Finally, Murray, Rushton, and Paunonen (1986) reported higher consistency across undergraduate course types (mean $r = .66$) than across undergraduate vs. graduate course types (mean $r = .15$).

Thus, it seems that mean student ratings of instructors are consistent over the same or similar courses taught in the same or successive years, but less generalizable across different types of courses taught in the same or successive years. It may be possible to regard this lack of consistency across course types as systematic

variation rather than unreliability or error variance. For example, it has been suggested that teachers may be more effective in one type of course than another (Seiler et al, 1977; Murray et al, 1986).

The interrater reliability of mean student evaluations also has been examined. Estimates are typically obtained either by correlating averaged ratings made by odd and even numbered raters across teachers or by intraclass correlation procedures. Reviews have concluded that such interrater reliabilities are generally in the .70's, .80's, and .90's (Feldman, 1977; Marsh, 1984). For example, Hoffman (1978) obtained estimates ranging from .76 to .90, Doyle and Crichton (1978) reported coefficients ranging from .73 to .84, and Erdle and Murray (1986) obtained a coefficient of .92.

In sum, therefore, research examining the reliability of class average student evaluations of teaching effectiveness has shown them to be consistent across items, time, similar types of courses, and raters.

Validity of class average student evaluations.

If valid, class mean ratings of instructional effectiveness should have moderate correlations with theoretically relevant variables such as other measures of teaching effectiveness, and certain characteristics of instructors, but minimal correlations with extraneous variables such as characteristics of courses. Considerable research has been undertaken to investigate these

relationships.

From a meta-analysis of research concerning the relation between mean student ratings and student achievement, Gohen (1981) reported coefficients ranging from .43 to .50 for studies employing multisection courses with common exams. Similarly, in a multisection validity study, Murray (1983a) reported positive correlations between mean student evaluations and final exam grades ($r=.30$), ratings of amount learned ($r=.43$), and subsequent registration in junior courses ($r=.55$). Thus, mean student ratings correlate moderately with both objective and subjective indices of teaching effectiveness.

Theories of teaching such as Dunkin and Biddle's (1974) model suggest that teaching effectiveness should be influenced by characteristics of teachers such as classroom behavior and personality. Many investigators, therefore, have correlated student ratings with measures of teachers' classroom behavior and personality in order to examine the convergent validity of mean student evaluations.

Similar classroom teaching behaviors have been found to correlate with mean student evaluations of teaching effectiveness, regardless of whether students themselves or independent observers rate the behaviors (Solomon, 1966; Deshpande et al, 1970; Keaveny & McGann, 1978; Mintzes, 1979; Cranton & Hillgartner, 1981; Murray, 1983b;

Erdle, Murray, & Rushton, 1985, Erdle & Murray, 1986). These predictive teaching behaviors can be considered to belong to two general categories: those that capture the attention of students and encourage student-teacher rapport (charismatic behavior); and those which reflect organization, explanation, and clear exposition of material (organizational behavior). Dimensions reflecting these classroom behaviors typically correlate moderately with mean student ratings. For example, Erdle et al. (1985) reported correlations of .63 and .53 between evaluations and Charisma and Organization dimensions, respectively.

Similar personality traits of teachers also have been found to correlate with class average ratings, regardless of whether independent observers or students themselves rated teacher personality (Costin & Grush, 1973; Sherman & Blackburn, 1975; Murray, 1975, 1978; Tomasco, 1980; Rushton, Murray, & Paunonen, 1983; Erdle et al, 1985, Murray, Rushton, & Paunonen, 1986; Feldman, 1986). Teacher personality traits such as leadership, sociability, supportiveness, objectivity, intelligence, orderliness, and ambition are typically found to correlate positively with mean ratings, whereas anxiety, authoritarianism, and defensiveness are found to correlate negatively. For example, using peer ratings of personality, Erdle et al (1985) found composite dimensions of Achievement and Interpersonal Orientation to correlate .51 and .45,

respectively, with mean effectiveness ratings.

If valid, however, mean student ratings should also correlate minimally with theoretically extraneous variables. Two general classes of variables have been examined: course characteristics and student characteristics. Low to moderate negative correlations have been reported between class size and mean ratings, and low to moderate positive correlations have been typically found between course level and mean ratings (Murray, 1980; Marsh, 1984). Thus, teachers of smaller, upper-level courses tend to receive higher ratings than those of larger, lower-level courses. Although correlations are often significant, they are considerably smaller than coefficients demonstrating convergent validity, typically accounting for less than 10 percent of the variance in mean ratings (Marsh, 1984). Moreover, such relationships may reflect actual differences in effectiveness rather than a bias in student ratings. It is possible that some charismatic behaviors such as encouraging participation may be more easily exhibited in smaller, senior level courses, and thus teachers are actually more effective in these classes.

Class average ratings of teaching effectiveness have also been correlated with class average levels of certain student characteristics. For example, ratings have typically been found to correlate moderately with

students' prior subject interest (Murray, 1980; Marsh, 1984). Thus, instructors of more popular courses tend to receive higher ratings than those of less popular ones. This finding, however, does not necessarily indicate a bias in mean student evaluations. It has been suggested that higher student interest in the course may create a more favorable environment which facilitates effective instruction (Marsh, 1984). Class average levels of student personality traits, on the other hand, have not been found to be associated with mean student evaluations of teaching, perhaps because different personality types occur in similar proportions in different classes, thereby minimizing variability in class average levels of student personality (Murray, 1980; Marsh, 1984). Of course, studies employing mean ratings to assess the relation between student characteristics and instructional effectiveness ratings actually treat such characteristics as those of courses rather than as those of students per se. Variability in levels of student characteristics is likely to be greater between individual students than between class averages. A more appropriate unit of analysis for the investigation of associations between student attributes and evaluations of teaching, therefore, is individual ratings within classes rather than class average ratings.

In sum, class average student ratings of teaching correlate moderately with relevant variables such as

objective and subjective measures of student learning, and teachers' personality traits and classroom behaviors. Mean ratings also correlate minimally with extraneous course characteristics. Thus, substantial evidence exists for both the convergent and discriminant validity of class average student ratings.

Reliability of individual student evaluations.

Research has also examined the reliability of individual student evaluations of teaching. In studies examining the internal consistency of individual ratings, Costin, Greenough, and Menges (1971) reported estimates ranging from .85 to .93, and Doyle (1975) reported coefficients ranging from .40 to .98, with a median of .85. In a study employing individual ratings to examine temporal stability across terms, Costin (1968) reported reliability coefficients ranging from .70 to .87. Thus, the internal consistency and test-retest reliabilities of individual ratings of teaching effectiveness approximate those of class mean ratings.

Although high interrater reliabilities have been reported for class average ratings, it is important to note that this does not necessarily indicate that students within classes give identical ratings to instructors. In fact, reviews of relevant research have concluded that interrater reliability estimates based on individual students' ratings are typically found to be in the .10's,

.20's and .30's (Feldman, 1977; Marsh, 1984). Clearly, individual students within classes vary in their ratings of instructors.

The rationale for averaging individual ratings over students in a class is that errors of observation will be reduced. That is, raters are believed to function in the same manner as items in conventional tests. There is some question, however, as to whether the observed variability within classes is simply random error variance or whether it is systematically related to attributes of student raters. Moreover, if individual ratings of effectiveness are related to student characteristics, there is some question as to whether such associations necessarily indicate a component of bias in student ratings. These issues have been considered to be within the domain of convergent-discriminant validity (Crittenden & Norr, 1973; Feldman, 1977; and Abrami, Perry, & Leventhal, 1982). Some writers have suggested that a valid assessment of teaching effectiveness should not be influenced by student characteristics since these are perceived as extraneous sources of error and bias (Abrami et al, 1982). On the other hand, some writers have argued that teachers may actually be more effective for some types of students than for others (Feldman, 1977; Marsh, 1984). Viewed in this manner, positive correlations between student attributes and effectiveness ratings could be interpreted as evidence for the validity of such ratings.

Validity of individual student evaluations.

Several studies have examined the relation between individual ratings and theoretically relevant variables. Positive correlations have been found between individual ratings and other measures of teaching effectiveness such as ratings of amount learned (Whitely & Doyle, 1979) and future course enrollment (Brunton & Crull, 1982) as well as between individual ratings and perceptions of teacher personality traits (Abrami et al, 1982). In these studies, however, ratings were pooled across classes, leaving open the possibility of mixing variability in student characteristics with variability in instructor characteristics and/or masking differential within-class relationships.

Similarly, although a considerable number of investigations have examined the association between student characteristics and individual student ratings of teaching effectiveness, Feldman (1977) and Abrami et al. (1982) have pointed out that the interpretation of results is problematic because inappropriate units of analyses are often used. Studies employing class average ratings clearly do not address the issue of within-class variability in ratings. Furthermore, such investigations mix variability in student ratings due to student characteristics with variability due to teacher and situational factors. Similarly, studies which use

individual ratings that have been pooled across classes as the units of analysis mix between-class and within-class variability in effectiveness ratings. In addition, pooling individual ratings will mask relationships between student characteristics and ratings which differ across types of courses and teachers.

Two procedures have been developed to attempt to control for variability due to teachers and situations in pooled ratings. Whitely and Doyle (1979) suggest the use of deviation scores as the units of analyses, and Abrami et al. (1982) advocate the use of hierarchical multiple regression to remove between-class variability before within-class variability is examined. Although these two approaches avoid the possibility of confounding variability due to student attributes with variability due to teacher and situational factors, they remain unable to control for the possibility of masking differential within-class relationships. Clearly, the most appropriate units of analysis are individual student ratings obtained from a single class for then no possibility would exist of either masking relationships or confounding between-class and within-class variability. Researchers, however, continue to examine the relationships between student characteristics and student ratings using pooled data (Hofman & Kremer, 1980; Abrami et al, 1982; Abrami & Mizener, 1985).

Studies have examined the relationships between

individual student evaluations and such student characteristics as interest in the course, gender, university level, attitude toward education, similarity to instructors, academic ability, and personality. Unfortunately, most investigations have employed pooled data.

Modest but significant correlations are typically reported between prior student interest in course material and individual effectiveness ratings (Feldman, 1977). For example, using pooled data, Whitely and Doyle (1979) obtained correlations of .06 and .26 between interest in course material and evaluations. Similarly, Howard and Schmeck (1979) found that student interest correlated from .08 to .26 with ratings of various teaching effectiveness items, in pooled data. It is interesting to note that assessments of prior interest obtained at the start of a term correlated .61 with those obtained at the end of a term, and that both measures showed similar correlations with effectiveness items.

Student gender generally has not been found to be related to ratings (Feldman, 1977; Marsh, 1984). Whitely and Doyle (1979), for example, reported a coefficient of .03. Individual student evaluations, however, have been found to correlate from .13 to .25 with items assessing student attitudes toward university instruction (Hofman and Kremer, 1980). Findings concerning the relation

between the university level of students and individual evaluations, on the other hand, have been inconsistent (Feldman, 1977). Some studies report no relation, others report positive correlations, and yet others report negative correlations.

Individual student ratings of teaching have been found to be positively related to student-teacher attitude similarity (Good & Good, 1973; Abrami & Mizener, 1985), and to positive personality dissimilarity (instructor being higher than the student) on traits relevant to effective instruction (Grush, Clore, and Costin, 1975). With respect to this latter finding, it was predicted on the basis of Clore and Byrne's (1974) reinforcement-affect model of attraction that students who were positively dissimilar from the instructor on relevant personality traits would give higher ratings than those who were similar to the instructor. In support of this prediction, students who perceived their instructor to be higher in Personal Relations and Ascendancy than themselves rated the instructor as more effective than those who perceived the instructor to be similar to themselves on these dimensions. Teacher Personal Relations and Ascendancy were considered to be relevant to effective instruction since student perceptions of these dimensions were positively related to ratings of instructional effectiveness. On the basis of these results, it might also be expected that students perceiving their instructor to be lower than

themselves on relevant traits would give lower ratings than those perceiving themselves to be similar to the instructor. This prediction, however, was not tested in the study by Grush et al.

Academic ability of students has not been found to be associated with individual ratings (Feldman, 1977; Whitely and Doyle, 1979). Studies concerning the relationship between students' course grades and individual evaluations, on the other hand, generally report correlations ranging from .10 to .40 (Feldman, 1976). It should be noted, however, that the relationship between grades and ratings was found to differ significantly across classes in a rare within-class study (Yonge and Sassenrath, 1968).

The relation between perceived grading leniency and individual evaluations has been examined in some studies by calculating the discrepancy between student grades and grade-point average. Students with discrepantly high grades have been found to rate teachers more favorably than those with discrepantly low grades (Bausell & Magdon, 1972). Discrepancy scores of this type, however, may not be an accurate measure of perceived leniency. A more direct measure, such as students' self-reported perceptions of grading leniency, may provide a clearer estimate of the relationship between this variable and effectiveness ratings.

Finally, student personality has been examined in relation to individual student evaluations. Although Follman (1976), after reviewing research concerning student personality traits, concluded that ratings were substantially influenced by these variables, more recent reviews have concluded that the literature reveals little evidence of significant relationships (Feldman, 1977; Abrami et al, 1982). For example, using ratings pooled across classes, Abrami et al (1982) found no relation between student personality and evaluations of teaching effectiveness. Yonge and Sassenrath (1968), however, examining within-class data, reported that relationships differed significantly across classes for some traits. Thinking Introversion, for example, correlated $-.26$, $.01$, and $.23$ with evaluations in three separate classes. Similarly, the relation between students' need for affiliation and grades was found to differ across classes by McKeachie, Lin, Milholland, and Isaacson (1966). In classes taught by instructors high in affiliation, students high in affiliation obtained higher grades than did those lower in affiliation. In classes taught by instructors low in affiliation, however, the reverse pattern was found. Thus, differential within-class relationships between student personality and evaluations may have been obscured in research employing pooled data to examine the relation between these variables.

In general, findings concerning the relation between

individual student evaluations of teaching and student characteristics have been inconsistent. One reason for this may be that most studies have used ratings pooled across classes which mixes between- and within-class variability and can mask differential within-class relationships. It has been suggested that the nature of relationships between student characteristics and evaluations may be affected by types of courses and teachers (Feldman, 1977). That is, certain types of students might find certain types of courses and teachers to be most effective; whereas other types of students might find other types of courses and teachers to be most effective. For example, extraverted students might find extraverted instructors and discussion-oriented courses to be most effective; whereas introverted students might find introverted instructors and lecture-oriented courses to be most effective. Given that the relationships between evaluations and both student personality and course grades have been found to differ significantly across classes (Yonge and Sassenrath, 1968), further research concerning within-class associations between student characteristics and evaluations of teaching is necessary.

The purpose of the present study, therefore, was to examine the relationships between a variety of student characteristics and student evaluations of teaching within separate classes, thereby avoiding the methodological

problems associated with the use of pooled ratings. On the basis of previous research, several specific predictions were made. First, it was expected that within-class variability in student evaluations of teaching effectiveness would be related to theoretically relevant variables. Thus, evaluations were predicted to be positively correlated with other subjective measures of teaching effectiveness (ratings of amount learned and likelihood of future course enrollment). Similarly, students' evaluations of teaching were expected to be positively related to their perceptions of teacher achievement, supportiveness, sociability, and orderliness, but negatively related to their perceptions of teacher anxiety, defensiveness, and authoritarianism. As well as providing evidence for the convergent validity of individual evaluations, support for this prediction would replicate results of studies using pooled ratings.

Second, it was predicted that associations between student personality traits and evaluations would vary across classes, accounting for the lack of consistent results in previous investigations using pooled data. Moreover, it was predicted that such differences would be related to differences between classes in relevant teacher personality traits. Thus, it was expected that a given student personality trait would be positively related to evaluations in classes taught by teachers perceived to have high levels of the trait, but would be negatively

related to evaluations in classes taught by teachers perceived to have low levels of the trait. Support for this hypothesis might be viewed as evidence for the validity of individual student ratings. That is, certain types of teachers may be effective for certain types of students; whereas other types of teachers are effective for other types of students.

Third, it was predicted that perceived student-teacher personality dissimilarity on traits relevant to teaching effectiveness would be related to student ratings within classes. For example, perceived teacher orderliness was predicted to be positively related to ratings of instructional effectiveness. Thus, it was expected that students perceiving their instructor to be more orderly than themselves would rate the instructor as more effective than would students perceiving the instructor to be less orderly than themselves. On the other hand, perceived teacher neuroticism was expected to be negatively related to ratings. Therefore, it was predicted that students perceiving their instructor to be more neurotic than themselves would rate the instructor as less effective than would students perceiving the instructor to be less neurotic than themselves. Support for this prediction would both replicate and extend research finding a relation between positive personality dissimilarity and evaluations in pooled ratings.

Finally, the within-class relationships between student instructional ratings and student gender, university level, attitude toward university instruction, prior interest in material, university grade average, and perceived grading leniency were examined to determine if associations varied across classes.

Method

Subjects.

The sample consisted of 278 students enrolled in six undergraduate psychology classes (N's of 79, 48, 33, 29, 25, and 64) taught by five different male professors at the University of Western Ontario. All classes were junior-level optional courses in three similar content areas (social, personality, and developmental psychology) with the exception of Class 6 (N=64) which was an introductory level course taught by the same instructor as Class 1 (N=79). Sample sizes for Classes 1, 3, and 4 were obtained by combining data from students in the same course^s taught by the same instructor in successive years. The percentages of students who completed the questionnaire in the six classes were 87, 98, 42, 51, 60, and 80, respectively.

Measures.

All subjects completed a questionnaire entitled "Survey of Student Perceptions of University Instruction" (see Appendix 1) during the last 3 weeks of classes. The first part of the questionnaire measured teaching effectiveness. Forty items were adapted from formal, end-of-course student evaluations of instructors used in North American colleges and universities (Endeavor Instructional Rating Card: Frey, 1978; The Student Instructional Report: Contra, 1973). Students rated these items on 7-point scales. In the next part of the survey students were asked

c

to rate first themselves and then their instructor on 30 personality trait adjectives on 7-point frequency-of-occurrence scales. These trait adjectives were adapted from the Personality Research Form (PRF; Jackson, 1974), but also included additional personality dimensions found to be useful in studies addressing the relation between teacher personality and instructional effectiveness. Finally, students were asked to indicate their gender, university level, major, attitude toward university instruction, prior and current interest in the course, amount learned, likelihood of future course enrollment, overall university grade average, and perceptions of grading leniency. For students who had given informed consent (see Appendix 1), course grades were obtained from instructors at the completion of the course. Feedback was mailed to subjects six months after participation in the study (see Appendix 2).

Results

Construction of composite teaching effectiveness measures.

A preliminary factor analysis of instructional effectiveness items was undertaken to identify underlying dimensions of teaching effectiveness. Student ratings for the 40 instructional effectiveness items were submitted to a principal components factor analysis. A scree test of eigenvalues (see Cattell, 1966) yielded a four factor structure accounting for 47.5 percent of the total variance in student ratings. Eigenvalues for the four factors were 12.22, 2.66, 2.07, and 2.03. The factor structure was rotated to a varimax criterion to facilitate interpretation. As can be seen in Table 1, 35 of 40 items had loadings of .45 or more on one of the 4 factors and no item loaded on more than one factor. Inspection of the rotated factor structure led to the interpretation of the factors as "Organization", "Rapport", "Availability", and "Difficulty", respectively.

Insert Table 1 About Here

Factor scores for these dimensions were computed by averaging all items loading higher than .45 on a given factor. As seen in Table 2, Cronbach alpha coefficients computed within each class ranged from .89 to .95, with a mean of .92, for Organization; .80 to .89, with a mean of .86, for Rapport; .74 to .79, with a mean of .63, for Availability; and .15 to .74, with a mean of .57, for

TABLE 1. LOADINGS OF INSTRUCTIONAL EFFECTIVENESS ITEMS FROM A PRINCIPAL-COMPONENTS FACTOR ANALYSIS, ROTATED TO A VARIMAX CRITERION.

INSTRUCTIONAL EFFECTIVENESS ITEM	FACTOR			
	I	II	III	IV
1. Overall, the instructor was effective in organizing and explaining course material.	.79	.30	.12	.04
2. Overall, the instructor was effective.	.78	.34	.06	.11
3. The instructor presented material clearly and summarized major points.	.74	.14	.20	-.16
4. The instructor was well prepared for class.	.73	.10	.15	.08
5. The instructor presented material in an interesting and relevant manner.	.72	.34	-.08	-.03
6. The instructor explained difficult concepts in a clear, straight-forward way.	.69	.16	.25	.02
7. The instructor summarized or emphasized major points.	.67	.11	.17	-.13
8. Lectures were in agreement with course objectives.	.66	.01	.23	.03
9. Overall, the course was valuable.	.63	.34	.16	.16
10. The instructor was enthusiastic about course material.	.63	.23	-.04	-.02
11. The instructor used class time well.	.62	.32	.07	.28
12. The instructor accomplished course objectives.	.59	.12	.35	.13
13. Each class was carefully planned in advance.	.59	.19	.06	.15
14. Overall, the lectures were valuable.	.58	.44	-.08	.21
15. The course has increased my knowledge and competence in this area.	.55	.23	.08	.22
16. Course objectives were clearly defined.	.48	.16	.23	.11

TABLE 1 (cont.)

INSTRUCTIONAL EFFECTIVENESS ITEM	FACTOR			
	I	II	III	IV
17. The instructor encouraged class participation.	.16	.77	.01	-.07
18. The student felt free to ask questions or give opinions.	.19	.70	.14	-.05
19. Class discussion was welcome in this course.	.17	.70	.06	-.12
20. Overall, class discussions were valuable.	.21	.67	-.12	-.02
21. Overall, the instructor was in tune (had rapport) with the students.	.34	.66	.14	.04
22. The instructor encouraged students to think.	.27	.60	.05	.07
23. The instructor raised challenging questions.	.24	.54	-.10	.08
24. Exams reflected important aspects of course.	.17	.52	.30	.30
25. The instructor was open to other viewpoints.	.18	.50	.41	.09
26. The instructor made helpful comments on papers or exams.	.11	.48	.29	.23
27. The instructor was concerned with student progress.	.16	.47	.35 ^t	.18
28. Overall, the exams were valuable.	.16	.46	.28	.41
29. The instructor was available for student consultation.	.29	.20	.68	.09
30. Overall, the textbook was valuable.	.10	-.14	.59	-.04
31. The student was able to get personal help.	.21	.37	.59	.02
32. The instructor informed students as to how they would be evaluated.	.18	.14	.46	-.17

TABLE 1 (cont.)

INSTRUCTIONAL EFFECTIVENESS ITEM	FACTOR			
	I	II	III	IV
33. The student had to work hard in this course.	.15	.04	.04	.76
34. The course was difficult.	-.06	.07	.00	.73
35. The pace of the course was too fast.	.05	-.19	-.20	.54
36. The grading accurately reflected the student's performance.	.23	.33	.17	.30
37. Lectures were too repetitive of the textbook.	-.35	-.34	.43	-.03
38. The instructor knew when students didn't understand.	.23	.44	.10	-.12
39. The scope of the course was too limited.	-.38	-.21	.07	-.03
40. Overall, other readings were valuable.	.12	.22	.01	.17

TABLE 2: CRONBACH ALPHA COEFFICIENTS FOR MEASURES OF INSTRUCTIONAL EFFECTIVENESS IN EACH OF SIX CLASSES.

MEASURE OF INSTRUCTIONAL EFFECTIVENESS	CLASS						MEAN
	1 N=79	2 N=48	3 N=33	4 N=29	5 N=25	6 N=64	
ORGANIZATION (16-ITEMS)	.94	.90	.89	.91	.92	.95	.92
RAPPORT (12-ITEMS)	.88	.89	.80	.85	.88	.84	.86
AVAILABILITY (4-ITEMS)	.65	.77	.24	.68	.79	.64	.63
DIFFICULTY (3-ITEMS)	.50	.66	.15	.66	.74	.71	.57
OVERALL EFFECTIVENESS (32-ITEMS)	.93	.94	.89	.94	.95	.94	.93
AMOUNT LEARNED (2-ITEMS)	.80	.74	.75	.57	.92	.49	.71
FUTURE ENROLLMENT (2-ITEMS)	.59	.63	.72	.46	.02	.47	.48

Difficulty.

Insert Table 2 About Here

Table 3 shows that moderate positive correlations were found among Organization, Rapport, and Availability factor scores within each class. Averaged across classes, Organization correlated .66 and .52 with Rapport and Availability, respectively, and Rapport and Availability correlated .42. Factor scores for Difficulty, however, failed to correlate significantly with scores for Organization, Rapport, and Availability (average r 's = .03, .06, and .08, respectively). Given this pattern of intercorrelations, an overall effectiveness measure was computed by averaging the 32 items comprising the Organization, Rapport, and Availability dimensions. Alpha coefficients for this measure (see Table 2) ranged from .89 to .95 across the six classes, with a mean of .92, indicating considerable unidimensionality for this dimension within each class.

Insert Table 3 About Here

In addition to formal instructional effectiveness items, students provided ratings concerning two other aspects of teaching effectiveness: amount learned and likelihood of future enrollment. Two items concerning learning, "Compared to other courses, how much do you feel

TABLE 3: INTERCORRELATIONS AMONG SCORES FOR FOUR INSTRUCTIONAL EFFECTIVENESS FACTORS, IN SIX CLASSES.

PAIR OF FACTORS CORRELATED	CLASS						MEAN
	1 N=79	2 N=48	3 N=33	4 N=29	5 N=25	6 N=64	
DIFFICULTY X ORGANIZATION	.16	-.12	-.27	.12	.03	.09	.03
DIFFICULTY X RAPPORT	.17	-.00	-.13	.13	-.06	.07	.06
DIFFICULTY X AVAILABILITY	.22*	.15	.07	-.14	-.22	.07	.08
ORGANIZATION X RAPPORT	.56*	.75*	.62*	.76*	.76*	.62*	.66*
ORGANIZATION X AVAILABILITY	.42*	.42*	.22	.68*	.64*	.69*	.52*
RAPPORT X AVAILABILITY	.28*	.61*	.17	.57*	.71*	.38*	.42*

* $p < .05$.

you have learned from this course?" and "Compared to other instructors, how much do you feel you have learned from your instructor, independent of course readings and assignments?", were aggregated to provide a single measure of amount learned. As seen in Table 2, alpha-coefficients computed within each class for this composite measure ranged from .49 to .92, with a mean of .71.

Similarly, two items concerning future enrollment, "Would you take another course in this area taught by your instructor, if not restricted by degree requirements?" and "Would you take another course in this area taught by a different instructor, if not restricted by degree requirements?", were summed to yield a single measure of the likelihood of future enrollment. Table 2 shows that alpha coefficients for this composite measure ranged from .02 to .72, with a mean of .48.

Construction of composite measures of student personality.

Student ratings of the 30 personality trait adjectives were submitted to a principal components factor analysis, to identify underlying personality dimensions. A scree test of eigenvalues yielded a five factor structure accounting for 50.5 percent of the total variance in student ratings. Eigenvalues for the five factors were 4.94, 4.11, 2.39, 2.09, and 1.61. The factor structure was rotated to a varimax criterion to facilitate interpretation. As can be seen in Table 4, 28 of 30 traits had loadings of .45 or more on one of the 5 factors and

TABLE 4: LOADINGS OF STUDENT PERSONALITY TRAITS FROM A PRINCIPAL-COMPONENTS FACTOR ANALYSIS, ROTATED TO A VARIMAX CRITERION

PERSONALITY TRAIT	I	II	Factor III	IV	V
1. Independent	.72	-.10	.09	.08	.06
2. Clever	.69	.07	.23	-.01	.15
3. Dependent	-.61	.19	.19	-.07	-.08
4. Ambitious	.59	.20	-.08	.16	-.29
5. Enduring	.54	.12	.05	.00	-.22
6. Dull	-.49	.22	-.33	-.13	.04
7. Yielding	-.49	.24	.06	-.44	-.10
8. Apathetic	-.47	.14	-.12	-.22	.40
9. Leadership	.46	.05	-.12	.55	.01
10. Defensive	.01	.68	.07	-.18	.04
11. Anxious	-.22	.67	.03	-.20	.10
12. Relaxed	.14	.61	.22	.03	-.01
13. Impatient	.10	.60	-.10	-.00	.31
14. Aggression	.20	.55	-.28	.31	.25
15. Authoritarian	-.05	.51	-.25	.15	-.24
16. Cautious	-.14	.46	.12	-.42	-.33
17. Supportive	-.03	.06	.69	.23	-.20
18. Thoughtless	-.12	.20	-.62	.05	.18
19. Modest	.01	.01	.59	-.09	-.13
20. Objective	.23	-.10	.55	-.12	.11
21. Arrogant	.06	.39	-.55	.11	.03
22. Subjective	.00	.31	-.51	-.03	-.12
23. Sociable	.08	-.09	.12	.85	-.02
24. Reserved	-.15	.19	.22	-.76	-.11
25. Open	-.07	.00	.38	.58	.12
9. Leadership	.46	.05	-.12	.55	.01
26. Pacifistic	-.08	-.25	.38	-.50	-.13
27. Disorganized	-.13	.16	-.10	-.01	.81
28. Orderly	.13	-.10	.13	-.09	-.79
29. Liberal	.41	-.22	.36	-.09	.42
30. Adventurous	.28	.00	-.12	.30	.43

only one trait loaded on more than one factor (i.e., Leadership). Inspection of the rotated factor structure led to the interpretation of the factors as "Achievement", "Neuroticism", "Nurturance", "Extraversion", and "Orderliness", respectively.

 Insert Table 4 About Here

Factor scores for these dimensions were computed by averaging all traits loading higher than .45 on a given factor. Negatively loaded traits were coded in reverse direction, and factor scores for Orderliness were reversed so that high scores reflected high levels of Orderliness. Table 5 shows that alpha coefficients computed within each class ranged from .76 to .87, with a mean of .80, for Achievement; .44 to .82, with a mean of .69, for Nurturance; .61 to .77, with a mean of .71, for Neuroticism; .63 to .77, with a mean of .70, for Extraversion; and .77 to .91, with a mean of .86, for Orderliness. These values indicate considerable unidimensionality for these five student personality measures within each class.

 Insert Table 5 About Here

Construction of composite measures of teacher personality.

Student ratings of 30 trait adjectives concerning teacher personality also were submitted to a principal

TABLE 5: CRONBACH ALPHA COEFFICIENTS FOR MEASURES OF STUDENT PERSONALITY IN EACH OF SIX COLLEGE CLASSES.

MEASURE OF STUDENT PERSONALITY	CLASS						MEAN
	1 N=79	2 N=48	3 N=33	4 N=29	5 N=25	6 N=64	
ACHIEVEMENT (9-ITEMS)	.78	.81	.80	.78	.87	.76	.80
NURTURANCE (6-ITEMS)	.75	.78	.76	.44	.82	.60	.69
NEUROTICISM (7-ITEMS)	.69	.71	.77	.61	.76	.74	.71
EXTRAVERSION (5-ITEMS)	.75	.63	.74	.67	.66	.77	.70
ORDERLINESS (2-ITEMS)	.85	.91	.86	.89	.90	.77	.86

components factor analysis, to identify underlying personality dimensions. A scree test of eigenvalues yielded a four factor structure accounting for 51.8 percent of the total variance in student ratings of teacher personality. Eigenvalues for the four factors were 8.03, 4.00, 1.85, and 1.66. Again, the factor structure was rotated to a varimax criterion to facilitate interpretation. As can be seen in Table 6, 26 of 30 traits had loadings of .50 or more on one of the 4 factors and no trait loaded on more than one factor. Inspection of the rotated factor structure led to the interpretation of the factors as "Nurturance", "Neuroticism", "Achievement", and "Orderliness", respectively.

 Insert Table 6 About Here

Factor scores for these dimensions were computed by averaging all traits loading higher than .50 on a given factor. Negatively loaded traits were coded in reverse direction, and factor scores for Nurturance were reversed so that high scores reflected high levels of Nurturance. As seen in Table 7, alpha coefficients computed within each class ranged from .80 to .90, with a mean of .85, for Nurturance; .69 to .89, with a mean of .78, for Neuroticism; .63 to .92, with a mean of .75, for Achievement; and .58 to .90, with a mean of .73, for Orderliness. These values indicate substantial unidimensionality for the four teacher personality

TABLE 6: LOADINGS OF TEACHER PERSONALITY TRAITS FROM A PRINCIPAL-
COMPONENTS FACTOR ANALYSIS, ROTATED TO A VARIMAX CRITERION.

PERSONALITY TRAIT	FACTOR			
	I	II	III	IV
1. Arrogant	.80	.09	.11	-.02
2. Thoughtless	.75	.14	-.12	-.12
3. Modest	-.73	.09	.09	.11
4. Aggressive	.69	.21	.25	-.17
5. Supportive	-.65	-.00	.36	.04
6. Defensive	.63	.35	.03	-.18
7. Pacifistic	-.57	.19	.14	.29
8. Liberal	-.56	-.06	.45	-.06
9. Subjective	.54	.22	.02	.12
10. Relaxed	-.53	-.15	.13	.47
11. Objective	-.52	.07	.29	.07
12. Yielding	-.01	.79	-.19	-.01
13. Cautious	.10	.75	-.09	-.02
14. Reserved	-.09	.65	-.39	.11
15. Dull	.12	.61	-.21	-.22
16. Dependent	.14	.60	-.23	-.14
17. Anxious	.33	.54	-.07	-.29
18. Ambitious	.16	-.27	.66	.24
19. Enduring	-.14	-.18	.66	.18
20. Clever	-.18	-.23	.64	.33
21. Independent	.04	-.44	.60	.17
22. Leadership	.01	-.47	.54	.21
23. Sociable	-.19	-.47	.50	-.05
24. Adventurous	.07	-.27	.50	-.30
25. Orderly	-.14	-.10	.23	.76
26. Disorganized	.11	.32	-.08	-.67
27. Apathetic	.24	.45	-.31	-.24
28. Open	-.32	-.03	.46	-.30
29. Impatient	.44	.30	.14	-.40
30. Authoritarian	.36	.17	.09	.27

TABLE 7: CRONBACH ALPHA COEFFICIENTS FOR MEASURES OF PERCEIVED
TEACHER PERSONALITY IN EACH OF SIX COLLEGE CLASSES.

MEASURE OF PERCEIVED TEACHER PERSONALITY	CLASS						MEAN
	1 N=79	2 N=48	3 N=33	4 N=29	5 N=25	6 N=64	
NURTURANCE (11-ITEMS)	.88	.86	.80	.84	.90	.82	.85
NEUROTICISM (6-ITEMS)	.69	.77	.82	.80	.89	.71	.78
ACHIEVEMENT (7-ITEMS)	.73	.76	.76	.63	.92	.89	.75
ORDERLINESS (2-ITEMS)	.68	.58	.68	.84	.90	.67	.73

measures within each class.

 Insert Table 7 About Here

Construction of composite measures of perceived personality
 dissimilarity.

Difference scores between students' self and teacher ratings of 30 personality trait adjectives were submitted to a principal components factor analysis, to identify underlying dimensions of personality similarity. A scree test of eigenvalues yielded a five factor structure accounting for 59.6 percent of the total variance in difference scores. Eigenvalues for the five factors were 6.66, 5.28, 2.33, 1.94, and 1.67. Again, the factor structure was rotated to a varimax criterion to facilitate interpretation. Table 8 shows that 25 of 30 traits had loadings of .50 or more on one of the 5 factors and only one trait loaded on more than one factor (i.e., Leadership). Inspection of the rotated factor structure led to the interpretation of the factors as "Nurturance", "Neuroticism", "Achievement", "Orderliness", and "Extraversion", respectively.

 Insert Table 8 About Here

Factor scores for these measures were computed by averaging all traits loading higher than .50 on a given factor. Negatively loaded traits were coded in reverse

TABLE 8: LOADINGS OF STUDENT-TEACHER DIFFERENCE SCORES FOR
PERSONALITY TRAITS FROM A PRINCIPAL-COMPONENTS FACTOR
ANALYSIS, ROTATED TO A VARIMAX CRITERION.

PERSONALITY TRAIT	I	II	FACTOR III	IV	V
1. Arrogant	-.83	.05	.12	.13	-.08
2. Modest	.74	-.02	-.04	-.10	.14
3. Thoughtless	-.68	-.12	-.16	.07	.26
4. Supportive	.67	.04	.15	-.15	-.29
5. Objective	.66	.14	.19	-.07	.03
6. Authoritarian	-.65	-.06	.02	.01	.14
7. Liberal	.62	-.22	.24	.10	-.12
8. Aggressive	-.58	-.14	.32	.33	-.14
9. Pacifistic	.51	.12	.06	-.37	.32
10. Yielding	.14	.74	-.36	.06	.11
11. Cautious	.12	.72	.22	.05	.42
12. Dependent	.04	.68	-.37	.20	-.16
13. Adventurous	-.01	-.66	.14	.32	-.13
14. Independent	.06	-.65	.41	-.05	.02
15. Anxious	-.20	.58	.14	.47	.16
16. Leadership	-.14	-.56	.53	-.16	-.19
17. Ambitious	-.25	-.18	.76	-.06	-.20
18. Endurance	.14	-.11	.75	-.17	.01
19. Clever	.17	-.24	.69	.05	-.12
16. Leadership	-.14	-.56	.53	-.16	-.19
20. Disorganized	.15	.00	-.23	.82	.11
21. Orderly	-.02	.08	.25	-.80	.03
22. Impatient	-.35	.13	.04	.54	.01
23. Sociable	.06	-.42	.24	-.03	-.77
24. Open	.26	.06	.10	-.04	-.73
25. Reserved	.15	.47	-.21	.02	.67
26. Defensive	-.42	.24	.16	.46	.30
27. Relaxed	.37	-.31	.02	-.49	-.01
28. Subjective	-.48	.06	.10	-.12	.02
29. Apathetic	-.06	.08	-.48	.35	.44
30. Dull	-.20	.42	-.42	.07	.17

direction, and factor scores for Extraversion and Orderliness were reversed. As seen in Table 9, alpha coefficients computed within each class ranged from .63 to .82, with a mean of .75, for Nurturance; .60 to .89, with a mean of .76, for Neuroticism; .59 to .85, with a mean of .68, for Achievement; .38 to .81, with a mean of .65, for Orderliness; and .41 to .81, with a mean of .64, for Extraversion. These values indicate substantial unidimensionality for the five perceived personality similarity measures within each class.

 Insert Table 9 About Here

Relation among measures of teaching effectiveness.

To determine whether within-class variability in student ratings of instructional effectiveness was related to variability in other subjective indices of teaching effectiveness, scores for the composite measures of Overall Effectiveness, Difficulty, Amount Learned, and Likelihood of Future Enrollment were intercorrelated separately within each of six classes. Table 10 shows that 16 of the 18 correlations among Overall Effectiveness, Amount Learned, and Likelihood of Future Enrollment in the six classes were significant. Averaged across classes, Overall Effectiveness correlated .58 and .50 with Amount Learned and Likelihood of Future Enrollment, respectively, and these latter two variables correlated .47. As can also be seen in Table 10, however, none of the 18 correlations

TABLE 9: CRONBACH ALPHA COEFFICIENTS FOR MEASURES OF PERCEIVED PERSONALITY DISSIMILARITY IN EACH OF SIX COLLEGE CLASSES.

MEASURE OF PERCEIVED PERSONALITY DISSIMILARITY	CLASS						MEAN
	1 N=79	2 N=48	3 N=33	4 N=29	5 N=25	6 N=64	
NURTURANCE (9-ITEMS)	.82	.74	.74	.63	.78	.81	.75
NEUROTICISM (7-ITEMS)	.75	.76	.73	.89	.84	.60	.76
ACHIEVEMENT (4-ITEMS)	.63	.59	.67	.61	.85	.74	.68
ORDERLINESS (3-ITEMS)	.71	.72	.62	.66	.81	.38	.65
EXTRAVERSION (3-ITEMS)	.59	.68	.78	.41	.59	.81	.64

TABLE 10: INTERCORRELATIONS AMONG COMPOSITE MEASURES OF TEACHING EFFECTIVENESS, FOR SIX CLASSES.

PAIR OF MEASURES OF TEACHING EFFECTIVENESS	CLASS						MEAN
	1 N=79	2 N=48	3 N=33	4 N=29	5 N=25	6 N=64	
OVERALL EFFECTIVENESS WITH DIFFICULTY	.23	-.02	-.02	.06	-.03	.10	.09
AMOUNT LEARNED WITH DIFFICULTY	.17	-.05	.08	.22	-.04	.10	.09
FUTURE ENROLLMENT WITH DIFFICULTY	.09	-.09	-.12	.15	-.17	-.06	-.02
OVERALL EFFECTIVENESS WITH AMOUNT LEARNED	.72*	.61*	.24	.56*	.59*	.52*	.58*
OVERALL EFFECTIVENESS WITH FUTURE ENROLLMENT	.60*	.68*	.42*	.44*	.46*	.26*	.50*
AMOUNT LEARNED WITH FUTURE ENROLLMENT	.54*	.54*	.22	.51*	.60*	.38*	.47*

* p < .05.

between these 3 measures of teaching effectiveness and Difficulty were significant the six classes. Averaged over classes, Difficulty correlated .09, .09, and -.02 with Overall Effectiveness, Amount Learned, and Likelihood of Future Enrollment, respectively.

Insert Table 10 About Here

Between-class differences in mean ratings for composite measures of teaching effectiveness.

Table 11 shows the pattern of differences in class average ratings for the composite measures of teaching effectiveness to be quite similar for Overall Effectiveness, Amount Learned, and Availability. Single-factor analyses of variance followed by post-hoc tests (Tukey's Honestly Significant Difference) indicated that the instructor of Class 1 and Class 6 received significantly lower mean ratings for Overall Effectiveness than the instructor of Class 3 (5.27 and 5.38 vs. 5.94). For Amount Learned, instructors of Classes 1 and 5 received significantly lower mean ratings than the instructor of Class 2 (4.60 and 4.70 vs 5.47), and the mean rating of Class 1 was significantly lower than that of Class 3 (4.60 vs. 5.35). Finally, for Likelihood of Future Enrollment, the mean rating of Class 5 (3.36) was significantly lower than that of Class 3 (4.00). In general, the instructors of Classes 2, 3, and 4 were among

TABLE 11: MEAN FREQUENCY ESTIMATES FOR COMPOSITE MEASURES OF
TEACHING EFFECTIVENESS, FOR SIX CLASSES.

MEASURE OF TEACHING EFFECTIVENESS	CLASS					
	1 N=79	2 N=48	3 N=33	4 N=29	5 N=25	6 N=64
OVERALL EFFECTIVENESS	5.27 ^a	5.64 ^{ab}	5.94 ^b	5.46 ^{ab}	5.61 ^{ab}	5.38 ^a
AMOUNT LEARNED	4.60 ^a	5.47 ^c	5.35 ^{bc}	5.22 ^{abc}	4.70 ^{ab}	5.06 ^{abc}
LIKELIHOOD OF FUTURE ENROLLMENT	3.72 ^{ab}	3.83 ^{ab}	4.00 ^b	3.91 ^{ab}	3.36 ^a	3.75 ^{ab}

NOTE. DIFFERENT ROW SUBSCRIPTS DENOTE MEAN DIFFERENCES SIGNIFICANT AT
THE .05 LEVEL.

SCALES FOR OVERALL EFFECTIVENESS AND AMOUNT LEARNED RANGED
FROM 1 TO 7, AND THE SCALE FOR LIKELIHOOD OF FUTURE
ENROLLMENT RANGED FROM 1 TO 5.

the most highly rated instructors for all three measures of teaching effectiveness, whereas the instructors of Classes 1, 5, and 6 were among the least highly rated on all three measures.

 Insert Table 11 About Here

Relation between perceptions of teacher personality and student instructional effectiveness ratings.

To test the hypothesis that within-class variability in student ratings of instructional effectiveness was related to variability in students' perceptions of their teacher's personality, the four composite measures of teacher personality were correlated with scores for Overall Effectiveness separately for each class. These coefficients are presented in Table 12. The correlation between student perceptions of teacher Nurturance and ratings of Overall Effectiveness was found to be significantly higher in Class 5 than Classes 1, 3, 4, and 6 (.83 vs. .38, .57, .44, and .53, respectively, and significantly higher in Class 2 (.71) than in Class 1 (.38). These differences, however, were only ones of magnitude. Teacher Nurturance was significantly related to ratings of Overall Effectiveness in each class, with an average correlation of .56 across classes. The correlation between perceived teacher Achievement and Overall Effectiveness was found to be significantly higher in Class 5 (.73) than in Class 2 (.39). Again, however,

teacher Achievement was significantly related to Overall Effectiveness in each class, with an average correlation of .51. The correlation between perceived teacher Neuroticism and Overall Effectiveness was found to be significantly higher in Classes 3, 4, and 5 than in Classes 1 and 6 (-.67, -.57, and -.62 vs. .03 and -.10), and significantly higher in Classes 3 and 5 than Class 2 (-.67 and -.62 vs. -.19). Perceived teacher Neuroticism was significantly related to Overall Effectiveness in 3 of the 6 classes, with a significant average correlation across all classes of -.26. Finally, teacher Orderliness was found to be more strongly related to Overall Effectiveness in Classes 4 and 5 than in Class 2 (.48 and .55 vs .06). Perceived teacher Orderliness was significantly related to Overall Effectiveness in 3 of the 6 classes, with a significant average correlation across classes of .26.

In general, therefore, despite some differences between classes in the magnitude of relationships, student perceptions of teacher Nurturance, Achievement, and Orderliness were found to be positively and significantly related to ratings of Overall Effectiveness; whereas perceptions of teacher Neuroticism were negatively related to ratings of Overall Effectiveness. Multiple regression analyses were conducted separately for each class with the four composite measures of perceived teacher personality as predictors and Overall Effectiveness ratings as the

TABLE 12: CORRELATIONS AND BETA WEIGHTS (IN PARENTHESES) BETWEEN
FOUR TEACHER PERSONALITY DIMENSIONS AND OVERALL
EFFECTIVENESS RATINGS IN SIX CLASSES.

PERSONALITY DIMENSION	CLASS						MEAN
	1 N=79	2 N=48	3 N=33	4 N=29	5 N=25	6 N=64	
NURTURANCE	.38 ^a (.37 [*])	.71 ^{bc} (.73 [*])	.57 ^{ab} (.24)	.44 ^{ab} (.14)	.83 ^{cc} (.77 [*])	.53 ^{ab} (.40 [*])	.56 [*]
NEUROTICISM	.03 ^a (.34 [*])	-.19 ^{ab} (.21)	-.67 ^{cc} (-.22)	-.57 ^{bc} (-.64 [*])	-.62 ^{cc} (.18)	-.10 ^a (.28)	-.26 [*]
ACHIEVEMENT	.50 ^{ab} (.54 [*])	.39 ^{aa} (.16)	.70 ^{ab} (.48 [*])	.34 ^{aa} (-.25)	.73 ^{bb} (.44)	.47 ^{ab} (.30)	.51 [*]
ORDERLINESS	.16 ^{ab} (.04)	.06 ^a (-.04)	.26 ^{ab} (.00)	.48 ^{bb} (.36)	.55 ^{bb} (-.22)	.32 ^{ab} (.16)	.26 [*]
MULTIPLE R	.66 [*]	.74 [*]	.79 [*]	.72 [*]	.86 [*]	.61 [*]	.71 [*]

*p. < .05.

NOTE. DIFFERENT ROW SUBSCRIPTS DENOTE CORRELATION DIFFERENCES
SIGNIFICANT AT THE .05 LEVEL.

criterion. Beta weights from these analyses are presented in Table 12. Multiple correlation coefficients ranged from .61 to .86 across classes, with an average of .71, indicating that on average 50% of the within-class variability of student instructional ratings was accounted for by student perceptions of teacher personality.

Insert Table 12 About Here

Between-class differences in mean ratings of composite teacher personality measures.

Single-factor analyses of variance followed by post-hoc tests (Tukey's Honestly Significant Difference) were performed to determine whether classes differed in terms of perceptions of teacher personality. As seen in Table 13, each composite measure of teacher personality differed significantly across classes. The instructors of Classes 1 and 2 received significantly lower ratings for Nurturance than those of Classes 3, 4, and 6 (5.11 and 4.87 vs. 5.85, 5.67, and 5.60), and the instructor of Class 5 had a significantly lower Nurturance rating than that of Class 3 (5.22 vs. 5.85). Instructors 4 and 5 had significantly higher scores for Neuroticism than instructors 1, 2, 3, and 6 (3.42 and 3.47 vs. 2.40, 2.39, 2.52, and 2.45). Similarly, instructors 4 and 5 had significantly lower scores for Achievement than instructors 1, 2, and 6 (4.38 and 4.78 vs. 5.51, 5.56, and 5.46. Instructor 4 also had a

TABLE 13: MEAN FREQUENCY ESTIMATES FOR FOUR TEACHER PERSONALITY DIMENSIONS, IN SIX CLASSES.

PERSONALITY DIMENSION	CLASS					
	1 N=79	2 N=48	3 N=33	4 N=29	5 N=25	6 N=64
NURTURANCE	5.11a	4.87a	5.85c	5.67bc	5.22ab	5.60bc
NEUROTICISM	2.40a	2.39a	2.52a	3.42b	3.47b	2.45a
ACHIEVEMENT	5.51a	5.56a	5.16ab	4.38c	4.78bc	5.46a
ORDERLINESS	6.03a	4.93cd	5.36bc	5.88ab	4.56d	5.92ab

NOTE. DIFFERENT ROW SUBSCRIPTS DENOTE MEAN DIFFERENCES SIGNIFICANT AT THE .05 LEVEL.

SCALES FOR ALL DIMENSIONS RANGED FROM 1 TO 7.

significantly lower Achievement score than instructor 3 (4.38 vs. 5.16). Finally, instructor 1 had a higher Orderliness score than instructors 2, 3, and 5 (6.03 vs. 4.93, 5.36, and 4.56); instructors 4 and 6 had significantly higher Orderliness scores than instructor 2 (5.88 and 5.92 vs. 4.93); and instructor 3 had a higher Orderliness score than instructor 5 (5.36 vs. 4.56). In sum, instructors 4 and 5 received the lowest ratings for Achievement, whereas instructors 1, 2, 3, and 6 received the highest ratings. The reverse pattern was found for ratings of Neuroticism. Instructors 1, 2, and 5 received the lowest ratings for Nurturance, and instructors 1, 6, and 4 received the highest Orderliness ratings.

Insert Table 13 About Here

Relation of prior interest, current interest, and change in interest to student instructional ratings.

To test whether within-class variability in ratings of instructional effectiveness was related to differences in levels of student interest, correlations between these variables were calculated separately for each class. Table 14 shows that the correlations between ratings of prior interest in the course material and ratings of Overall Effectiveness ranged from $-.16$ to $.32$, with a significant average correlation of $.18$. Individual coefficients were significant in only two of six classes, however. Correlations between ratings of interest in course

TABLE 14: CORRELATIONS OF STUDENTS' PRIOR INTEREST, CURRENT INTEREST, AND CHANGE IN INTEREST WITH OVERALL EFFECTIVENESS, FOR SIX CLASSES.

	CLASS						MEAN
	1 N=79	2 N=48	3 N=33	4 N=29	5 N=25	6 N=64	
PRIOR INTEREST	.17 ^a	.29 ^{*a}	.32 ^{*a}	.19 ^a	-.16 ^e	.14 ^a	.18 [*]
CURRENT INTEREST	.68 ^{*a}	.67 ^{*a}	.41 ^{*ab}	.48 ^{*ab}	.62 ^{*ab}	.33 ^{*b}	.55 [*]
INTEREST CHANGE	.56 ^{*b}	.38 ^{*ab}	.04 ^a	.32 ^{ab}	.69 ^{*b}	.18 ^a	.38 [*]

*p. < .05.

NOTE. DIFFERENT ROW SUBSCRIPTS DENOTE CORRELATION DIFFERENCES SIGNIFICANT AT THE .05 LEVEL.

material at the end of term (Current Interest) and Overall Effectiveness ratings ranged from .33 to .68, with a significant average correlation of .55. Although this relationship was found to be significantly weaker in Class 6 than in Classes 1 and 2 (.33 vs. .68 and .67), end-of-course interest was significantly related to Overall Effectiveness in each class. Finally, the correlations between change in interest and Overall Effectiveness ranged from .04 to .69, with a significant mean value of .38. Although coefficients were in the predicted direction in each class, significance was reached in only one half of the classes. Correlations were significantly higher in Classes 1 and 5 than in Classes 3 and 6 (.56 and .69 vs. .04 and .18).

Insert Table 14 About Here

Relation of grades, change in grades, and perceived grading leniency to student instructional ratings.

Ratings of overall instructional effectiveness were also correlated with students' grades for the course, students' overall university grade averages, the difference between these grades, and students' perceptions of grading leniency. Again, these correlations were calculated separately for each class. As seen in Table 15, students' overall university grade averages failed to correlate significantly with Overall Effectiveness in any

TABLE 15: CORRELATIONS OF STUDENTS' COURSE GRADES, OVERALL UNIVERSITY GRADE AVERAGE, DIFFERENCE BETWEEN GRADE AVERAGE AND COURSE GRADE, AND PERCEIVED GRADING LENIENCY WITH OVERALL EFFECTIVENESS, FOR SIX CLASSES.

	CLASS						MEAN
	1 N=79	2 N=48	3 N=33	4 N=29	5 N=25	6 N=64	
OVERALL AVERAGE	.06a	-.11a	-.09a	-.07a	-.08a	.11a	.00
COURSE GRADE	.14ab	-.20a	-.17a	-.05ab	.17ab	.30*b	.07
DIFFERENCE BETWEEN OVERALL AVERAGE AND COURSE GRADE	.09ab	-.16a	-.15a	.05ab	.24ab	.34*b	.09
PERCEIVED GRADING LENIENCY	.16a	.20a	-.04a	-.12a	-.19a	.20a	.10

*p < .05.

NOTE. DIFFERENT ROW SUBSCRIPTS DENOTE CORRELATION DIFFERENCES SIGNIFICANT AT THE .05 LEVEL.

of the classes. Coefficients ranged from $-.11$ to $.11$, with an average correlation of $.00$. Students' grades for the course correlated significantly with Overall Effectiveness ratings in only 1 class. Coefficients ranged from $-.20$ to $.30$, with an average of $.07$. The correlation in Class 6 was found to be significantly different from those in Classes 2 and 3 ($.30$ vs. $-.20$ and $-.17$). The difference between course grades and overall university grade averages also correlated significantly with Overall Effectiveness ratings in only 1 class. Correlations ranged from $-.16$ to $.34$, with an average of $.09$. Again, the correlation in Class 6 was found to be significantly different from those in Classes 2 and 3 ($.34$ vs. $-.16$ and $-.15$). Finally, students' perceptions of grading leniency were not found to correlate significantly with Overall Effectiveness ratings in any of the classes. Correlations ranged from $-.19$ to $.20$, with an average of $.10$.

 Insert Table 15 About Here

Relation of student gender, university level, major, and attitude toward university instruction to student instructional ratings.

Several background characteristics of students were correlated with ratings of Overall Effectiveness for each class. Table 16 shows that student gender failed to correlate significantly with Overall Effectiveness in any of the classes. Coefficients ranged from $-.10$ to $.31$, with

TABLE 16: CORRELATIONS OF STUDENT GENDER, UNIVERSITY LEVEL, MAJOR,
AND ATTITUDE TOWARD UNIVERSITY INSTRUCTION WITH OVERALL
EFFECTIVENESS RATINGS, FOR SIX CLASSES.

	CLASS						MEAN
	1 N=79	2 N=48	3 N=33	4 N=29	5 N=25	6 N=64	
GENDER (1=MALE, 2=FEMALE)	.06a	-.06a	-.10a	.20a	.31a	.06a	.06
UNIVERSITY LEVEL (2nd, 3rd, 4th yr.)	-.08a	-.50*b	-.12ab	-.22ab	-.35*ab	--	-.24*
AREA OF MAJOR (1=YES, 2=NO)	-.09a	.20a	.10a	-.08a	-.70*b	.10a	-.04
ATTITUDE TOWARD UNIVERSITY INSTRUCTION	.27*ab	.37*b	.28ab	-.10a	.14ab	.06ab	.20*

*p. < .05.

NOTE. DIFFERENT ROW SUBSCRIPTS DENOTE CORRELATION DIFFERENCES
SIGNIFICANT AT THE .05 LEVEL.

an average of .06. The university level of students, however, correlated significantly with Overall Effectiveness in 2 of 5 classes. Coefficients ranged from -.08 to -.50, with a significant average correlation of -.24. Thus, senior-level students tended to give lower ratings of instructional effectiveness than did junior-level students.

 Insert Table 16 About Here

Table 16 also shows that area of major was related to Overall Effectiveness in only 1 of 6 classes. Coefficients ranged from .20 to -.70, with an average of -.04. Finally, it was found that in 2 of 6 classes students with favorable attitudes toward university instruction rated their instructor as significantly more effective than did students with less favorable attitudes. Coefficients ranged from -.10 to .37, with a significant average correlation of .20.

Relation between composite measures of student personality and student instructional ratings.

To test whether student personality was differentially related to instructional effectiveness, correlations between composite student personality dimensions and Overall Effectiveness ratings were computed separately in each of the six classes. Table 17 shows that correlations differed significantly between classes for 3

TABLE 17: CORRELATIONS BETWEEN FIVE STUDENT PERSONALITY DIMENSIONS
AND OVERALL EFFECTIVENESS RATINGS, FOR SIX CLASSES.

PERSONALITY DIMENSION	CLASS						MEAN
	1 N=79	2 N=48	3 N=33	4 N=29	5 N=25	6 N=64	
ACHIEVEMENT	.09 ^{ab}	.35 ^{ab}	.24 ^b	-.45 ^{ab}	-.35 ^a	.10 ^{ab}	.06
NURTURANCE	-.05 ^a	.09 ^a	.12 ^a	.31 ^a	.00 ^a	.11 ^a	.07
NEUROTICISM	.27 ^{ab}	-.26 ^b	-.13 ^{ab}	-.50 ^{ab}	-.17 ^{ab}	-.02 ^{ab}	-.06
EXTRAVERSION	-.08 ^a	.11 ^a	.17 ^a	-.32 ^a	-.29 ^a	-.15 ^a	-.08
ORDERLINESS	-.12 ^a	.28 ^b	.14 ^{ab}	.40 ^{ab}	-.30 ^a	.14 ^{ab}	.06

*p. < .05.

NOTE. DIFFERENT ROW SUBSCRIPTS DENOTE CORRELATION DIFFERENCES
SIGNIFICANT AT THE .05 LEVEL.

of 5 personality dimensions. Moreover, correlations differed between classes in direction rather than just magnitude (as was the case for teacher personality measures).

Insert Table 17 About Here

Student Achievement correlated significantly with Overall Effectiveness in 2 of 6 classes. Coefficients ranged from $-.45$ to $.35$, with an average of $.06$. The relation between Achievement and Overall Effectiveness was found to be significantly different in Classes 2 and 3 than in Classes 4 and 5 ($.35$ and $.24$ vs. $-.45$ and $-.35$). It is interesting to note that the instructor of Class 2 received significantly higher mean ratings of Achievement than did those of Classes 4 and 5, and the instructor of Class 3 received significantly higher Achievement ratings than did that of Class 4 (see Table 13). Thus, positive correlations were found between student evaluations and student achievement in those classes taught by instructors perceived to be high in achievement; whereas negative correlations were found in classes taught by instructors perceived to be lower in achievement.

Student Nurturance failed to correlate significantly with Overall Effectiveness ratings in any of the classes. There were also no significant differences between classes in the relationship between Nurturance and Overall Effectiveness. Coefficients ranged from $-.05$ to $.31$, with

an average correlation of .07. It is nevertheless interesting to observe that the classes with the highest coefficients (Classes 3, 4, and 6) were taught by the instructors with the highest mean ratings for teacher Nurturance (see Table 13).

Student Neuroticism correlated significantly with Overall Effectiveness in 2 of 6 classes. Coefficients ranged from .27 to $-.50$, with an average of $-.06$. The relationship was found to be significantly different in Class 1 than in Classes 2 and 4 (.27 vs. $-.26$ and $-.50$). Again, it is interesting to note that the instructor of Class 1 had a significantly lower mean rating for instructor Neuroticism than did the instructor of Class 4, although not lower than that of the instructor of Class 2.

Overall Effectiveness and student Extraversion failed to correlate significantly in any of the classes. Coefficients ranged from .17 to $-.32$, with an average of $-.08$. Finally, student Orderliness correlated significantly with Overall Effectiveness in only 1 class. Coefficients ranged from .40 to $-.30$, with an average of .08. The relationship was found to be significantly different in Classes 2 and 4 than in Classes 1 and 5 (.28 and .40 vs. $-.12$ and $-.30$). It may be recalled that the instructor of the class with the highest negative relation (Class 5) received a significantly lower mean rating for teacher Orderliness than did the instructor of the class

with the highest positive relation (Class 4). The reverse pattern was found, however, for the instructors of Classes 1 and 2.

Between-class differences in mean ratings of composite student personality measures.

Single-factor analyses of variance followed by post-hoc tests (Tukey's Honestly Significant Difference) were performed to determine whether differences existed between classes in mean levels for the 5 student personality measures. As seen in Table 18, mean ratings for student Achievement, Nurturance, Extraversion, and Orderliness failed to differ significantly across classes. Mean ratings for student Neuroticism differed significantly between only 2 classes. Class 6 was found to have a significantly higher mean rating than did Class 3 (3.55 vs. 3.08).

 Insert Table 18 About Here

Relation between perceived student-teacher personality dissimilarity and ratings of instructional effectiveness.

Perceived dissimilarity scores for five composite personality measures were correlated with instructional effectiveness ratings in each of 6 classes. These coefficients are presented in Table 19. Five of 6 correlations between perceived dissimilarity in Nurturance and Overall Effectiveness were found to be significant, ranging from .22 to .73, with a significant average

TABLE 18: MEAN FREQUENCY ESTIMATES FOR FIVE STUDENT PERSONALITY DIMENSIONS, IN SIX CLASSES.

PERSONALITY DIMENSION	CLASS					
	1 N=79	2 N=48	3 N=33	4 N=29	5 N=25	6 N=64
ACHIEVEMENT	5.26a	5.29a	5.30a	5.08a	5.19a	5.26a
NURTURANCE	5.31a	5.15a	5.51a	5.45a	5.24a	5.25a
NEUROTICISM	3.33ab	3.33ab	3.08a	3.38ab	3.52ab	3.55b
EXTRAVERSION	4.26a	4.43a	4.37a	4.41a	4.42a	4.38a
ORDERLINESS	5.06a	4.50a	5.41a	5.26a	5.06a	4.95a

NOTE. DIFFERENT ROW SUBSCRIPTS DENOTE MEAN DIFFERENCES SIGNIFICANT AT THE .05 LEVEL.

SCALES FOR ALL DIMENSIONS RANGED FROM 1 TO 7.

TABLE 19: CORRELATIONS BETWEEN PERCEIVED STUDENT-TEACHER DIFFERENCE SCORES FOR FIVE PERSONALITY DIMENSIONS AND OVERALL EFFECTIVENESS RATINGS, FOR SIX CLASSES.

PERSONALITY DIMENSION	CLASS						MEAN
	1 N=79	2 N=48	3 N=33	4 N=29	5 N=25	6 N=64	
NURTURANCE	.40*ab	.59*bc	.39*abc	.22a	.73*c	.38*ab	.41*
NEUROTICISM	-.13ab	.08a	-.28*ab	-.38*bc	-.74*c	-.08ab	-.19*
ACHIEVEMENT	.24*ab	.06a	.36*abc	.42*abc	.71*c	.42*bc	.32*
ORDERLINESS	.25*bc	-.24a	.10abc	-.10ab	.55*c	.19bc	.13*
EXTRAVERSION	.27*a	.25*a	.27ab	.55*ab	.67*b	.33*ab	.35*

*p. < .05.

NOTE. DIFFERENT ROW SUBSCRIPTS DENOTE CORRELATION DIFFERENCES SIGNIFICANT AT THE .05 LEVEL.

correlation of .41. This relationship was significantly stronger in Class 5 than in Classes 1, 4, and 6 (.73 vs. .40, .22, and .38), and significantly stronger in Class 2 (.59) than in Class 4 (.22). In general, therefore, students who perceived their instructor to be higher in Nurturance than themselves rated the instructor more highly than did students who perceived their instructor to be lower in Nurturance than themselves. It will be recalled that perceived teacher Nurturance was found to be positively related to Overall Effectiveness (see Table 12).

Insert Table 19 About Here

Three of 6 correlations between perceived dissimilarity for Neuroticism and Overall Effectiveness were significant. Coefficients ranged from .08 to $-.74$, with a significant average correlation of $-.19$. The relationship was found to differ significantly between Class 5 and Classes 1, 2, 3, and 6 ($-.74$ vs. $-.13$, $.08$, $-.28$, and $-.08$), and between Class 4 ($-.38$) and Class 2 ($.08$). In general, however, students who perceived their instructor to be more Neurotic than themselves rated the instructor as less effective than did those who perceived the instructor as less Neurotic than themselves. Again, it will be recalled that teacher Neuroticism was found to be negatively related to Overall Effectiveness ratings.

Five of 6 correlations between perceived dissimilarity in Achievement and Overall Effectiveness were found to be significant. Coefficients ranged from .06 to .71, with a significant average of .32. Again, this relationship was significantly stronger in Class 5 than in Classes 1 and 2 (.71 vs. .24 and .06), and Class 6 differed significantly from Class 2 (.42 vs. .06). However, in general, students who perceived their instructor to be higher in Achievement than themselves rated the instructor more highly than did students rating their instructor as less achieving than themselves. Again, perceived teacher Achievement was found to be positively related to Overall Effectiveness.

Perceived dissimilarity in Orderliness was found to be significantly related to Overall Effectiveness in 2 of 6 classes. Coefficients ranged from -.24 to .55, with a significant average correlation of .13. The relationship was found to differ significantly between Class 5 and Classes 2 and 4 (.55 vs. -.24 and -.10), and between Class 2 and Classes 1 and 6 (-.24 vs. .25 and .19). Thus, across classes, students perceiving their instructor to be more Orderly than themselves rated the instructor as more effective than did those students who perceived the instructor to be less Orderly than themselves. Perceived teacher Orderliness was found to be positively related to Overall Effectiveness ratings.

Finally, perceived dissimilarity in Extraversion was

found to be significantly related to effectiveness ratings in 5 of 6 classes. Coefficients ranged from .25 to .67, with a significant average correlation of .35. Again, this relationship was stronger in Class 5 (.67) than in Classes 1 (.27) and 2 (.25). In general, therefore, students who perceived their instructor to be more Extraverted than themselves rated the instructor as more effective than did students perceiving the instructor to be less Extraverted than themselves.

In sum, 20 of 30 correlations between perceived personality dissimilarity and Overall Effectiveness were found to be significant. Generally, students perceiving their instructor to be more Nurturant, Achieving, Orderly, and Extraverted but less Neurotic than themselves rated the instructor as more effective than did students who perceived their instructor to be less Nurturant, Achieving, Orderly, and Extraverted but more Neurotic than themselves. It is interesting to note that teacher Nurturance, Achievement, and Orderliness were found to be positively related to ratings of instructional effectiveness; whereas teacher Neuroticism was found to be negatively related (see Table 12).

Relation of perceived student-teacher personality similarity to student instructional ratings.

Perceived similarity scores for 5 composite measures of personality (i.e., the absolute value of difference

scores) were correlated with Overall Effectiveness ratings, in each of 6 classes. As can be seen in Table 20, however, only 7 of 30 coefficients were found to be significant. Correlations between perceived similarity in Nurturance and Overall Effectiveness ranged from $-.19$ to $.51$, with a significant average correlation of $.17$. This relationship was found to differ significantly between Class 2 and Classes 3, 4, and 6 ($.51$ vs. $-.19$, $-.13$, and $-.01$), and between Class 1 ($.30$) and Class 3 ($-.19$). Correlations between perceived similarity in Neuroticism and Overall Effectiveness ranged from $-.13$ and $.55$, with an average of $.02$. This relationship differed significantly between Class 5 ($.55$) and Classes 1, 2, 3, 4, and 6 ($-.09$, $-.01$, $-.12$, $-.13$, and $.10$). Correlations between perceived similarity in Achievement and Overall Effectiveness ranged from $-.27$ to $.53$, with an average of $.03$. Again, this relationship differed significantly between Class 5 ($.53$) and Classes 1, 2, 3, and 4 ($-.21$, $-.03$, $.07$ and $-.27$). Correlations between perceived similarity in Orderliness and Overall Effectiveness ranged from $-.09$ to $.22$, with an average of $.03$. This relationship did not differ significantly between classes. Finally, correlations between perceived similarity in Extraversion and Overall Effectiveness ranged from $-.25$ to $.54$, with an average of $.04$. This relationship was significantly different in Classes 5 ($.54$) and 4 ($.53$) than in Classes 1, 2, and 6 ($-.07$, $-.17$, and $-.25$), and

TABLE 20: CORRELATIONS BETWEEN PERCEIVED STUDENT-TEACHER SIMILARITY SCORES FOR FIVE PERSONALITY DIMENSIONS AND OVERALL EFFECTIVENESS RATINGS, FOR SIX CLASSES.

PERSONALITY DIMENSION	CLASS						MEAN
	1 N=79	2 N=48	3 N=33	4 N=29	5 N=25	6 N=64	
NURTURANCE	.30*ab	.51*a	-.19c	-.13bc	.28abc	-.01bc	.17*
NEUROTICISM	-.09a	-.01a	-.12a	-.13a	.55*b	.10a	.02
ACHIEVEMENT	-.21a	-.03ab	.07ab	-.27a	.53*c	.26bc	.03
ORDERLINESS	-.09a	.21a	.08a	.12a	.22a	-.09a	.03
EXTRAVERSION	-.07ab	-.17a	.32*bc	.53*c	.54*c	-.25a	.04

*p. < .05.

NOTE. DIFFERENT ROW SUBSCRIPTS DENOTE CORRELATION DIFFERENCES SIGNIFICANT AT THE .05 LEVEL.

significantly different in Class 3 (.32) than in Classes 2 (-.17) and 6 (-.25).

• -----
• Insert Table 20 About Here
• -----

Discussion

In general, the present findings both support and extend previous research concerning student evaluations of instructional effectiveness. The factor structure of the teaching effectiveness inventory found in the present study corresponded to those reported in previous investigations (Kulik & Kulik, 1974; Linn, Centra, & Tucker, 1975; Frey, 1978; Marsh, 1984). Teaching dimensions reflecting Organization, Rapport, Availability, and Difficulty have consistently been found in factor analyses using within-class, pooled, and class mean ratings. Moreover, the pattern of correlations between dimensions in the present study was quite similar to those reported in previous investigations. For example, Marsh (1984) found correlations of .33 and .43 between measures of Organization/Clarity and Individual Rapport in two samples. Workload/Difficulty, however, correlated $-.03$ and $-.05$ with Individual Rapport, and $-.15$ and $.02$ with Organization/Clarity, in the two samples. Similarly, in the present study, moderate correlations were found between Organization, Rapport, and Availability dimensions, but Difficulty failed to correlate significantly with any of these dimensions. Averaged across classes, Organization correlated $.66$ and $.52$ with Rapport and Availability, respectively, and these latter two dimensions correlated $.42$. Difficulty, however, correlated $.03$, $.06$, and $.08$ with Organization, Rapport,

and Availability, respectively. It seems from both present and previous results that course difficulty or workload is unrelated to other components of teaching effectiveness, and therefore might be regarded as an attribute of courses which is independent of the instructors' teaching effectiveness. Items concerning course difficulty or workload, therefore, should be omitted from formal student evaluations of instructors, since it would appear that they will only contribute to error variance in scores.

The finding that within-class variability in student ratings of teaching effectiveness was positively related to variability in other indices of teaching effectiveness in the present study supports results of previous investigations using class mean ratings (Murray, 1983a) and pooled individual ratings (Feldman, 1977; Whitely & Doyle, 1979; Brunton & Crull, 1982). Averaged across classes, overall effectiveness ratings correlated .58 and .50 with ratings of amount learned in the course and likelihood of future course enrollment, respectively. These latter two measures correlated .47. Murray (1983a) reported coefficients of .43 and .55 between class mean student evaluations and ratings of amount learned and senior course registration, respectively, with a correlation of .44 between class means for amount learned and course registration. Similarly, using pooled individual ratings, Whitely and Doyle (1979) reported a

correlation of .63 between teaching evaluations and ratings of amount learned. At the level of class means, such correlations are typically considered to be evidence for the convergent validity of student ratings. Instructors who are rated as effective by students are also those whose students report high levels of learning and register in senior level courses. At the level of individual ratings, such correlations indicate that within-class variability in student ratings cannot be considered simply error variance. Ratings of the effectiveness of instructors correlate substantially with both ratings of amount learned and ratings of the likelihood of future course enrollment. Moreover, as with class mean ratings, such correlations might be interpreted as evidence for the convergent validity of student ratings. That is, students for whom the instructor is most effective are also those who learn the most and who will enroll in similar courses. Alternatively, it may be that the positive correlations found between evaluations of teachers and ratings of amount learned and likelihood of future enrollment simply reflect a need for student raters to maintain consistency in their assessments of teaching.

The finding that student perceptions of teacher personality were related to student evaluations of teaching also supports results of previous investigations using class mean ratings or pooled individual ratings (Rushton et al, 1983; Abrami et al, 1982). Averaged across

classes, student ratings of effectiveness correlated .56, -.26, .51, and .26 with student ratings of teacher nurturance, neuroticism, achievement, and orderliness, respectively. The pattern of correlations was found to be reasonably consistent across classes, with differences being primarily ones of magnitude. Calculated within classes, multiple correlations between student evaluations and perceptions of teacher personality ranged from .61 to .86, with an average of .71, indicating that on average approximately 50 percent of the within-class variability in student evaluations of instructors can be explained by variability in student perceptions of teacher personality. These findings again demonstrate that within-class variability in student evaluations is not entirely a result of random error.

At the level of class means, the relation between teacher personality and student evaluations is viewed as evidence for the convergent validity of student ratings. For example, Erdle et al (1985) found that instructors with certain personality traits (Achievement and Interpersonal Orientation) were rated as more effective partly because they exhibited various organizational and charismatic classroom behaviors which were related to teaching effectiveness. At the level of individual ratings within classes, correlations between evaluations of teaching effectiveness and perceptions of teacher

personality might also be viewed as evidence for the convergent validity of student ratings. A teacher may actually be more effective for students who perceive him or her to exhibit high levels of achievement, nurturance, and orderliness than for students who perceive him or her to have low levels of these traits. It is possible that students differ in their impressions of teacher personality, in part, because they differ in their interpretation of teachers' classroom behavior. This would suggest that within-class relationships between student perceptions of teacher personality and their evaluations of teaching effectiveness are mediated by impressions of relevant classroom teaching behaviors. This interpretation is consistent with results of studies employing class means (Erdle et al, 1985), but remains to be empirically demonstrated. An alternative interpretation of the relation between perceived teacher personality and evaluations of teaching, however, is that students attribute traits to teachers on the basis of their evaluations of the teachers' instruction rather than vice-versa.

As found in previous research using class means or pooled data (reviewed by Feldman, 1977; Murray, 1980; Marsh, 1984), the present within-class investigation found that students' level of prior interest in course material was minimally related to ratings of instructional effectiveness. Averaged across classes, prior interest

correlated significantly ($r=.18$) with individual ratings. Thus, although students with high initial interest in course material tended to rate teachers as more effective than did students with lower initial interest levels, the relationship was quite weak. The association between prior interest and class average evaluations has often been viewed as evidence of a component of bias in student ratings of effectiveness. Some authors, however, have offered a second interpretation of the relation between mean student evaluations and prior subject interest, namely that teachers are actually more effective in classes where students have high levels of interest, perhaps because of increased levels of interaction (Marsh, 1984). Similarly, within classes, it might be that teachers are actually more effective for those students who have high initial interest levels in the course, perhaps because it is these students who benefit most from classroom interactions.

It is important to note that the present findings indicated a much stronger relation between end-of-course interest levels and student ratings than between prior interest levels and student ratings. Coefficients ranged across classes from .33 to .55, with a significant average correlation of .55. Similarly, interest change scores were found to be related to student ratings within classes. Coefficients ranged from .04 to .69, with a significant

average of .38. Thus, students who reported high end-of-course interest in material and a positive change in interest were those who gave high ratings of teaching effectiveness. As was the case for ratings of amount learned and likelihood of future enrollment, these findings can be interpreted as evidence for the convergent validity of individual student ratings, since end-of-course interest and change in interest might be viewed as indicants of instructional effectiveness. Alternatively, however, these results may simply reflect an effort by students to maintain consistency in their assessments of teaching.

Unlike some previous research (reviewed by Feldman, 1976), little evidence was found in the present study for a relation between course grades and ratings of teaching effectiveness at the level of individual ratings. It may be that previous moderate correlations reported in some investigations using pooled data resulted from confounding between-class and within class sources of variance. That is, effects may have resulted from differences between classes in the effectiveness of instructors. Yonge and Sassenrath (1968), using a within-class design, reported that grades correlated significantly with ratings in only 1 of 3 classes. In the present within-class study, course grades were found to correlate with effectiveness ratings in only 1 of 6 classes, with an average correlation across classes of .07. Similarly, the correlation between

students' overall university grade levels and their ratings of instructional effectiveness ranged from $-.11$ to $.11$, with an average correlation of $.00$. The within-class correlations of student ratings with the discrepancy between course grades and university grade average (a measure of grading leniency) ranged from $-.16$ to $.34$, with an average of $.09$. Finally, direct perceptions of grading leniency were also found to be unrelated to effectiveness ratings, ranging from $-.19$ to $.20$, with an average of $.10$. In general, these results indicate that individual ratings within classes were unrelated to measures of student ability, and grading leniency.

As in previous research using class average or pooled individual ratings (Feldman, 1977; Marsh, 1984), the present within-class analysis found no significant relation between student gender and ratings of instructional effectiveness. Correlation coefficients ranged across classes from $-.10$ to $.31$, with an average correlation of $.06$. Similarly, averaged across classes, whether or not the course was in the area of students' major had no relation to effectiveness ratings (average $r = -.04$). In one class, however, a significant negative relationship was found ($r = -.70$). Students for whom the course was in the area of their major rated the instructor as more effective than did students for whom the course was not in the area of their major.

Using class mean evaluations, instructors of upper level classes have typically been found to receive higher effectiveness ratings than have instructors of lower level classes (Murray, 1980; Marsh, 1984), a result not always replicated when using individual ratings (Feldman, 1977). The present study found that higher level students rated instructors as less effective than did lower level students, with coefficients ranged from $-.08$ to $-.50$ across five classes, with a significant average correlation of $-.24$. This is the reverse of the pattern found using class means as the units of analysis. It is important to note, however, that these classes were all lower level classes. It is possible that teachers are perceived as most effective by students for whom the course is at the appropriate level, and that a reverse pattern, therefore, would be found in higher level classes.

The negative within-class relationship between students' university level and evaluations of teaching also may have implications for the interpretation of the positive relation found between student level and class average effectiveness ratings. It is possible that relation between these variables at the level of class means is an artifact of class size. Instructors of smaller classes typically receive higher mean ratings than do those of larger classes (Murray, 1980; Marsh, 1984), and upper level courses typically are smaller in size than

lower level courses. The present study suggests that the relation between students' university level and individual evaluations may depend upon the level of the course.

The present study also replicated previous studies using pooled individual ratings (Hofman & Kremer, 1980) in finding a modest correlation between students' attitudes toward university instruction and their ratings of instructional effectiveness. Coefficients were found to range across classes from $-.10$ to $.37$, with a significant average correlation of $.20$. Although significant differences were found between classes, in general, students with a more favorable attitude toward university instruction rated their instructor as more effective than did students with less favorable attitudes. This relation may be interpreted in two ways. First, it may be that teachers are actually most effective for students with favorable attitudes toward instruction perhaps because these students participate in classroom interactions. Alternatively, these students may give high evaluations to instructors regardless of instructional ability.

The finding that student personality was differentially related to effectiveness ratings across classes in this study supports results of previous within-class investigations (Yonge & Sassenrath, 1968), and demonstrates the utility of examining associations within separate classes. Correlations between 3 of 5 personality

dimensions were found to differ significantly across classes in this study. Thus, although the average correlation between student achievement and effective ratings was found to be only .06, coefficients varied from -.45 to .35 across classes. Correlations in Classes 2 and 3 (.35 and .24) differed significantly from those in Classes 4 and 5 (-.45 and -.35). It is interesting to observe that the instructors of Classes 2 and 3 received higher mean ratings of teacher achievement than did instructors of Classes 4 and 5. Thus, for instructors perceived as low in achievement, students high in achievement rated them as less effective than did students lower in achievement; whereas the reverse was found for instructors perceived as high in achievement.

Similarly, although the average correlation between student neuroticism and effectiveness ratings was found to be only -.06, coefficients ranged across classes from .27 to -.50. The correlation in Class 1 (.27) was found to differ significantly from that in Class 4 (-.50). Again, it is interesting to note that the instructor of Class 1 received a significantly lower mean rating for teacher neuroticism than did the instructor of Class 4. Thus, for the instructor perceived as high in neuroticism, students high in neuroticism gave lower effectiveness ratings than did students low in neuroticism; whereas the reverse was true for the instructor low in neuroticism.

Although student orderliness was found to correlate

only .08 with evaluations across classes, coefficients ranged from $-.30$ to $.40$. Correlations in Classes 1 and 5 ($-.12$ and $-.30$) were found to differ significantly from those in Classes 2 and 4 ($.28$ and $.40$). Again, it is interesting to observe that for the classes with the largest difference in correlation (Classes 4 and 5), a significant difference was found for mean ratings of teacher orderliness. The instructor of Class 4 received a significantly higher mean rating of orderliness than did the instructor of Class 5. Thus, the instructor perceived as low in orderliness received lower effectiveness ratings from students high in orderliness than from students low in orderliness, whereas the reverse was found for the instructor perceived as high in orderliness. This corresponds to the pattern of results found for student achievement. It is important to note, however, that although an identical pattern of correlation differences between student orderliness and effectiveness ratings was found between Classes 1 and 2, the difference between instructors in mean ratings of teacher orderliness was reversed.

Finally, the relationships between student evaluations and student nurturance and extraversion were not found to be significant in any of the classes tested. It is interesting, however, that for student nurturance the higher positive correlations with effectiveness were

in those classes taught by instructors receiving the higher mean ratings of teacher nurturance.

In sum, differences were found between classes in correlations between 3 of 5 student personality dimensions and evaluations of teaching. Moreover, these differences were found to be related to differences in relevant teacher personality traits. These results might be interpreted as evidence for the validity of individual ratings. If student personality was simply a source of bias in ratings, the relationship between these variables would not be expected to differ across classes. The present results, however, suggested that certain types of students find certain types of instructors to be effective; whereas other types of students find other types of instructors to be effective. In general, for traits perceived to be positively related to teaching effectiveness (achievement, nurturance, and orderliness), students with high scores tended to rate teachers as more effective than did students with low scores when teachers were perceived to have high levels of the traits. When teachers were perceived to have low levels of the traits, students with low scores tended to rate teachers as more effective than did students with high scores. For the trait perceived to be negatively related to teaching effectiveness (neuroticism), students with high scores tended to rate teachers as more effective than did students with low scores when teachers were perceived to

have low levels of the trait; whereas the reverse was found when teachers were perceived to have high levels of the trait. These results suggest, therefore, that for personality traits perceived to be positively related to effective instruction, student-teacher similarity predicts effectiveness ratings; whereas dissimilarity predicts ratings for traits perceived to be negatively related to instructional effectiveness. It should be noted that this conclusion applies to actual similarity and dissimilarity across classes.

The relationships between perceived student-teacher similarity and dissimilarity in personality within classes were also examined in the present study. Averaged across classes, perceived dissimilarity in nurturance, neuroticism, achievement, orderliness, and extraversion were found to predict student evaluations (r 's = .41, -.19, .32, .13, and .35, respectively). Thus, in general, students perceiving their teacher to be more nurturant, achieving, orderly, and extraverted but less neurotic than themselves rated the instructor as more effective than did students perceiving the teacher to be less nurturant, achieving, orderly, extraverted, but more neurotic than themselves. It should be recalled that perceptions of teacher nurturance, achievement, orderliness, and extraversion were found to be positively related to evaluations; whereas teacher neuroticism was found to be

negatively related to evaluations. These results are consistent with and extend those reported by Grush et al (1975) concerning positive student-teacher personality dissimilarity on relevant traits. As would be expected, given the present findings concerning perceived positive personality dissimilarity, little evidence was found for a relation between perceived student-teacher personality similarity and evaluations of teaching effectiveness.

In general, the present results indicated that within-class variability in student evaluations of teachers was systematically related to student characteristics and perceptions. On average, approximately 50 percent of the variability in student evaluations within classes was found to be accounted for by perceptions of teacher personality. Similarly, across classes, evaluations were found to be substantially correlated with other subjective indices of teaching effectiveness: ratings of amount learned (.58), likelihood of future course enrollment (.50), and end-of-course interest (.55). Averaged across classes, student instructional ratings also correlated with prior interest (.18), student attitudes toward college instruction (.20), student university level (-.24), and perceived student-teacher personality dissimilarity. Student gender, area of major, and grades, however, were found to be unrelated to evaluations. Finally, the relationship between evaluations and student personality was found to differ across classes for 3 of 5

dimensions. Moreover, these differences seemed to be meaningfully related to differences between classes in mean perceptions of relevant teacher personality traits.

Several limitations of this research, however, should be kept in mind. First, although the present results are consistent with the view that within-class variability in student evaluations of teaching reflects actual differences in the effectiveness of the instructor for different students, it is possible that within-class variability simply reflects rating biases. Positive correlations between student evaluations and other subjective indices of teaching effectiveness could have resulted from an effort by students to be consistent in their assessments, and correlations between student characteristics and evaluations of teaching could have resulted if some types of students rated instructors higher than other types for reasons other than the effectiveness of the instructor. Second, the present findings may not generalize to methods other than the lecture method, to teaching content areas other than psychology or to instructors who are female (given that the present sample was composed entirely of male instructors). For example, it seems likely that personality traits might be perceived by students as differentially desirable for male and female instructors. Finally, although student gender was found to be unrelated to ratings of teaching

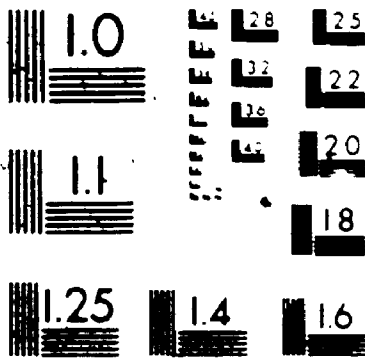
effectiveness, this does not rule out the possibility that relationships between such ratings and other variables may differ for male and female students. Future research concerning these issues is clearly necessary.

In spite of possible drawbacks, this study has implications for both the understanding and improvement of university instruction. The present findings suggest that within-class variability in student evaluations may reflect actual differences in the effectiveness of instruction for different students. Thus, college instructional effectiveness might be enhanced by an understanding of the reasons for such differences. Given that student evaluations were found to be related to perceptions of certain teacher personality traits, it seems that teaching effectiveness might be improved by an effort on the part of instructors to engage in behaviors which reflect such traits. The present results also suggested that certain types of students may find certain types of instructors to be effective; whereas other types of students may find other types of instructors to be effective. Similarly, future research might find that certain types of students find certain types of courses to be effective; whereas other types of students find other types of courses to be effective. Thus, it may be possible to improve college instruction by matching types of students with types of courses and teachers. For example, instructors might include brief descriptions of their

2

OF/DE

2



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS
STANDARD REFERENCE MATERIAL 1010a
(ANSI AND ISO TEST CHART No. 2)

teaching style and personality on course syllabi. Students could then use such information to select instructors whose styles match their needs. Additional investigations along these lines will further both the understanding and improvement of college teaching.

References

- Abrami, P.C., & Mizener, D. A. (1985). Student/instructor attitude similarity, student ratings, and course performance. Journal of Educational Psychology, 77, 693-702.
- Abrami, P. C., Perry, R. P., & Leventhal, L. (1982). The relationship between student personality characteristics, teacher ratings, and student achievement. Journal of Educational Psychology, 74, 111-125.
- Bausell, R. B., & Magoon, J. (1972). Expected grade in a course, grade-point average, and student ratings of the course and the instructor. Educational and Psychological Measurement, 32, 1013-1023.
- Bausell, R. B., Schwartz, S., & Purohit, A. (1975). An examination of the conditions under which various student rating parameters replicate across time. Journal of Educational Measurement, 12, 273-280.
- Bejar, I. I. (1975). A survey of selected administrative practices supporting student evaluation of Instructional programs. Research in Higher Education, 3, 77-86.
- Brunton, B.T., & Crull, S.R. (1982). Causes and consequences of student evaluation of instruction. Research in Higher Education, 17, 195-206.
- Cattell, R. B. (1966). The scree test for the number of factors. Multivariate Behavioral Research, 1, 245-276.

- Centra, J. A. (1973). Reliability of Student Instructional Report items. SIR Report No. 3. Princeton, N.J.: Educational Testing Service.
- Centra, J. A. (1977). How universities evaluate faculty performance: A survey of department heads. GREB Research Report 75-5bR. Princeton, N.J.: Educational Testing Service.
- Clore, G. L., & Byrne, D. (1974). A reinforcement-affect model of attraction. In T. L. Huston (Ed.), Perspectives on interpersonal attraction. New York: Academic Press.
- Cohen, P.A. (1981). Student ratings of instruction and student achievement: A meta-analysis of multisection validity studies. Review of Educational Research, 51, 281-309.
- Costin, F. (1968). A graduate course in the teaching of psychology: Description and evaluation. Journal of Teacher Education, 19, 425-432.
- Costin, F., & Grush, J.E. (1973). Personality correlates of teacher-student behavior in the college classroom. Journal of Educational Psychology, 65, 35-44.
- Costin, F., Greenough, W., & Menges, R. J. (1971). Student ratings of college teaching: Reliability, validity, and usefulness. Review of Educational Research, 41, 511-535.

- Cranton, P.A., & Hillgartner, W. (1981). The relationship between student ratings and instructor behavior: Implications for improved teaching. Canadian Journal of Higher Education, 11, 73-81.
- Crittenden, K. S., & Norr, J. L. (1973). Student values and teacher evaluation: A problem in person perception. Sociometry, 36, 143-151.
- Deshpande, A.S., Webb, S.C., & Marks, E. (1970). Student perceptions of engineering instructor behaviors and their relationship to the evaluation of instructors and courses. American Educational Research Journal, 7, 289-305.
- Doyle, K. O., Jr. (1975) Student evaluation of instruction. Lexington, Mass.: Heath.
- Doyle, K. O., Jr. & Crichton, L. (1978). Student, peer, and self evaluations of college instructors. Journal of Educational Psychology, 70, 815-826.
- Dunkin, M.J., & Biddle, B.J. (1974). The study of teaching. New York: Holt, Rinehart, & Winston.
- Erdle, S., & Murray, H.G. (1986). Interfaculty differences in classroom teaching behaviors and their relationship to student instructional ratings. Research in Higher Education, 24, 115-127.
- Erdle, S., Murray, H.G., & Rushton, J.P. (1985). Personality, classroom behavior, and student ratings of college teaching effectiveness: A path analysis. Journal of Educational Psychology, 77, 394-407.

- Feldman, K. A. (1976). Grades and college students' evaluations of their courses and teachers. Research in Higher Education, 4, 69-111. 6
- Feldman, K. A. (1977). Consistency and variability among college students in rating their teachers and courses: A review and analysis. Research in Higher Education, 6, 223-274.
- Feldman, K. A. (1986). The perceived instructional effectiveness of college teachers as related to their personality and attitudinal characteristics: A review and synthesis. Research in Higher Education, 24, 139-213.
- Follman, J. (1975). Student ratings of faculty teaching effectiveness: Rater or ratee characteristics? Research in Higher Education, 3, 155-167.
- Frey, P. W. (1978). A two-dimensional analysis of student ratings of instruction. Research in Higher Education, 9, 69-91.
- Gillmore, G. M. (1972). Estimates of reliability coefficients for items and subscales of the Illinois Course Evaluation Questionnaire. Research Report No. 341, Urbana, Ill.: University of Illinois, Office of Instructional Resources.
- Glasman, N. S. & Gmelch, W. H. (1976). Purposes of evaluation of university instructors: Definitions, delineations, and dimensions. Canadian Journal of Higher Education, 6, 37-55.

- Good, K. C. & Good, L. R. (1973). Assumed attitude similarity and instructor evaluation. Journal of Social Psychology, 91, 285-290.
- Grush, J. E., Clore, G. L. & Costin, F. (1975). Dissimilarity and attraction: When difference makes a difference. Journal of Personality and Social Psychology, 32, 783-789.
- Guthrie, E. R. (1964). The evaluation of teaching: A progress report. Seattle: University of Washington.
- Hoffman, R. G. (1978). Variables affecting university student ratings of instructor behavior. American Educational Research Journal, 15, 287-299.
- Hofman, J. E., & Kremer, L. (1980). Attitudes toward higher education and course evaluation. Journal of Educational Psychology, 72, 610-617.
- Hogan, T. P. (1973). Similarity of student ratings across instructors, courses, and time. Research in Higher Education, 1, 149-154.
- Howard, G. S. & Schmeck, R. R. (1979). Relationship of changes in student motivation to student evaluations of instruction. Research in Higher Education, 10, 305-315.
- Jackson, D. N. (1974). Personality Research Form Manual. Goshen, N.Y.: Research Psychologists Press.

- Keaveny, J.J., & McGann, A.F. (1978). Behavioral dimensions associated with students' global ratings of college professors. Research in Higher Education, 9, 333-345.
- Kulik, J. A. & Kulik, C. C. (1974). Student ratings of instruction. Teaching of Psychology, 1, 51-57.
- Linn, R. L., Centra, J. A. & Tucker, L. R. (1975). Between, within, and total group factor analyses of student ratings of instruction. Multivariate Behavioral Research, 10, 277-288.
- Marsh, H.W. (1984). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases, and utility. Journal of Educational Psychology, 76, 707-754.
- McKeachie, W. J. (1966). Student affiliation motives, teacher warmth, and academic achievement. Journal of Personality and Social Psychology, 4, 457-461.
- Mintzes, J. J. (1979). Overt teaching behaviors and student ratings of instructors. Journal of Experimental Education, 48, 145-153.
- Mueller, F. J. (1951). Trends in student ratings of faculty. American Association of University Professors Bulletin, 37, 319-324.
- Murray, H. G. (1972). The validity of student ratings of faculty teaching ability. Presented at annual meetings of the Canadian Psychological Association.

- Murray, H. G. (1975). Predicting student ratings of college teaching from peer ratings of personality traits. Teaching of Psychology, 2, 66-69.
- Murray, H. G. (1978). Teacher ratings, student achievement, and teacher personality traits. Presented at annual meetings of the Canadian Psychological Association.
- Murray, H. G. (1980). Evaluating university teaching: A review of research. Toronto: Ontario Confederation of University Faculty Associations.
- Murray, H.G. (1983a). Low-inference classroom teaching behaviors in relation to six measures of college teaching effectiveness. Presented at Conference on the Evaluation and Improvement of University Teaching, Montebello, Quebec.
- Murray, H.G. (1983b). Low-inference classroom teaching behaviors and student ratings of college teaching effectiveness. Journal of Educational Psychology, 75, 138-149.
- Murray, H.G., Rushton, J.P., & Paunonen, S.V. (1986). Teacher personality and student instructional ratings in six types of college courses. Submitted manuscript.
- Rushton, J. P., Murray, H. G. & Paunonen, S. V. (1983). Personality, research creativity, and teaching effectiveness in university professors. Scientometrics, 5, 93-116.

- Seiler, L. H., Weybright, L. D. & Stang, D. J. (1977). How useful are published evaluations to students' selection of courses and instructors? Teaching of Psychology, 4, 174-177.
- Seldin, P. (1980) Successful faculty evaluation programs. Crugers, N.Y.: Coventry Press.
- Sherman, B.R., & Blackburn, R.T. (1975). Personal characteristics and teaching effectiveness of college faculty. Journal of Educational Psychology, 67, 124-131.
- Solomon, D. (1966). Teacher behavior dimensions, course characteristics, and student evaluations of teachers. American Educational Research Journal, 3, 35-47.
- Tomasco, A.T. (1980). Student perceptions of instructional and personality characteristics of faculty: A canonical analysis. Teaching of Psychology, 7, 79-82.
- Whitely, S. E. & Doyle, K. O., Jr. (1979). Validity and generalizability of student ratings from between-classes and within-class data. Journal of Educational Psychology, 71, 117-124.
- Yonge, G. D. & Sassenrath, J. M. (1968). Student personality correlates of teacher ratings. Journal of Educational Psychology, 59, 44-52.

APPENDIX 1: SURVEY OF STUDENT PERCEPTIONS OF
UNIVERSITY INSTRUCTION.

Survey of Student Perceptions of University Instruction

General Information:

In this questionnaire, you will be asked to evaluate the effectiveness of your course and instructor as well rate yourself and your instructor on several personality traits. In addition, you will be asked to provide some information regarding your interests and expectations concerning this course. Participation in this investigation is completely voluntary. Your responses to this questionnaire, however, may further the understanding and improvement of instruction at this university and, therefore, would be greatly appreciated. This survey is designed for research purposes only and should not be confused with the formal instructor evaluation for this course which you will be asked to complete at another time. Your responses, combined with those of your classmates, will be used to examine the adequacy of student ratings as a measure of university instructional effectiveness. Therefore, it is important that you answer in a completely open and frank manner. Your responses will be completely confidential and individual student ratings will never be released.

For the purposes of this investigation, it is important to obtain your final grade in this course from the instructor. In order to match your grade with your responses to this questionnaire, you will be asked to provide your Student Identification Number. After recording your final grade, however, your student number will be deleted from the questionnaire before submitting your responses for computer processing. At that time, your responses will be completely anonymous. If you do not wish to release this (or any other) information, then feel free to omit any questions. Your responses to all questions, however, would be greatly appreciated.

A brief report concerning the findings of this survey will be available in the fall of 1985. If you wish to receive a copy, then provide your name and address (as of Fall, 1985) in the space provided below. This information will be kept separate from the rest of your questionnaire to protect the anonymity of your responses.

(Note. To be completed only if feedback desired.)

Name: _____
 Address: _____
 City/Town: _____ Province: _____
 Postal Code: _____

SUBJECT CONSENT FORM

1. I agree to participate as a subject in this research project, conducted by Stephen Erdle of the Department of Psychology at The University of Western Ontario.
2. I have read the "General Information" cover sheet concerning the nature and purpose of the Survey of Student Perceptions of University Instruction.
3. I understand that my participation in this project is entirely voluntary and that I may choose not to answer any questions and to terminate my participation at any time.
4. I understand that all information I provide will be kept strictly confidential.
5. I understand that the results of the study will be made available to me if I complete the "feedback desired" sheet.
6. I understand that there are no risks to me, personally, in the research.

Signed: _____

Instructional Effectiveness Inventory

Instructions:

The questions below concern the effectiveness of your instructor and course. Please use the following rating scale in making your judgements:

- 1 - strongly disagree
- 2 - disagree
- 3 - slightly disagree
- 4 - undecided
- 5 - slightly agree
- 6 - agree
- 7 - strongly agree

Please enter your response in the space beside each question.

-
1. The student had to work hard in this course. : :
 2. Each class period was carefully planned in advance. : :
 3. Class discussion was welcome in this course. : :
 4. The student was able to get personal help in this course. : :
 5. The instructor presented the material clearly and summarized major points. : :
 6. The grading accurately reflected the student's performance. : :
 7. The course has increased my knowledge and competence in this area. : :
 8. Course objectives were clearly defined. : :
 9. The instructor was enthusiastic about course material. : :
 10. Lectures were in agreement with course objectives. : :
 11. The instructor encouraged class participation. : :
 12. The instructor used class time well. : :
 13. The instructor presented material in an interesting and relevant manner. : :
 14. The lectures were too repetitive of the textbook. : :
 15. The instructor was available for student consultation. : :

16. The instructor was well prepared for class. : :
17. The instructor knew when students didn't understand. : :
18. The scope of the course was too limited. : :
19. The instructor encouraged students to think. : :
20. The instructor summarized or emphasized major points. : :
21. The instructor was concerned with student progress. : :
22. The instructor accomplished objectives for the course. : :
23. The instructor made helpful comments on papers or exams. : :
24. The instructor raised challenging questions. : :
25. Exams reflected important aspects of the course. : :
26. The student felt free to ask questions or give opinions. : :
27. The course was difficult. : :
28. The instructor informed students as to how they would be evaluated. : :
29. The instructor explained difficult concepts in a clear, straight-forward way. : :
30. The instructor was open to other viewpoints. : :
31. The pace of the course was too fast. : :
32. Overall, the textbook was valuable. : :
33. Overall, other readings were valuable. : :
34. Overall, the exams were valuable. : :
35. Overall, the lectures were valuable. : :
36. Overall, class discussions were valuable. : :
37. Overall, the instructor was in tune (had rapport) with the students. : :
38. Overall, the instructor was effective in organizing and explaining course material. : :
39. Overall, the instructor was effective. : :
40. Overall, the course was valuable. : :

Personality Rating Form - A

Instructions.

Please rate yourself on each of the personality traits. All ratings should reflect the extent to which you feel you exhibit the trait in question. Please use the following rating scale in making your judgements:

- 1 - never
- 2 - almost never
- 3 - rarely
- 4 - sometimes
- 5 - frequently
- 6 - almost always
- 7 - always

-
1. Ambitious (competitive, industrious, achieving) : :
 2. Reserved (introverted, quiet, shy) : :
 3. Pacifistic (even-tempered, peaceful, mild-mannered) : :
 4. Defensive (suspicious, guarded, touchy) : :
 5. Clever (bright, logical, insightful) : :
 6. Dependent (relies on others, group-oriented, conforming) : :
 7. Enduring (persistent, durable, steadfast) : :
 8. Cautious (apprehensive, seeks safety, timid) : :
 9. Supportive (nurturant, helpful, comforting) : :
 10. Yielding (meek, submissive, follower) : :
 11. Orderly (neat, methodical, systematic) : :
 12. Relaxed (calm, tranquil, composed) : :
 13. Subjective (partial, biased, unfair) : :

14. Liberal (flexible, tolerant, progressive) : :
15. Arrogant (conceited, egotistic, vain) : :
16. Apathetic (unmotivated, indifferent, not interested
in accomplishments) : :
17. Sociable (extraverted, outgoing, friendly) : :
18. Aggressive (threatening, hostile, argumentative) : :
19. Open (trusting, self-disclosing, revealing) : :
20. Dull (unimaginative, slow-witted, unintelligent) : :
21. Independent (individualistic, self-reliant, autonomous) : :
22. Impatient (restless, uncommitted, fidgety) : :
23. Adventurous (takes risks, daring, enjoys
dangerous activities) : :
24. Thoughtless (unsympathetic, neglecting, uncaring) : :
25. Leadership (assertive, self-assured, dominant) : :
26. Disorganized (untidy, messy, sloppy) : :
27. Anxious (nervous, tense, uneasy) : :
28. Objective (just, unprejudiced, equitable) : :
29. Authoritarian (rigid, conservative, dogmatic) : :
30. Modest (humble, unpretentious, down-to-earth) : :

Personality Rating Form - B

Instructions.

Please rate your instructor on each of the personality traits. All ratings should reflect the extent to which you feel your instructor exhibits the trait in question. Please use the following rating scale in making your judgements:

- 1 - never
- 2 - almost never
- 3 - rarely
- 4 - sometimes
- 5 - frequently
- 6 - almost always
- 7 - always

-
1. Ambitious (competitive, industrious, achieving) : :
 2. Reserved (introverted, quiet, shy) : :
 3. Pacifistic (even-tempered, peaceful, mild-mannered) : :
 4. Defensive (suspicious, guarded, touchy) : :
 5. Clever (bright, logical, insightful) : :
 6. Dependent (relies on others, group-oriented, conforming) : :
 7. Enduring (persistent, durable, steadfast) : :
 8. Cautious (apprehensive, seeks safety, timid) : :
 9. Supportive (nuturant, helpful, comforting) : :
 10. Yielding (meek, submissive, follower) : :
 11. Orderly (neat, methodical, systematic) : :
 12. Relaxed (calm, tranquil, composed) : :
 13. Subjective (partial, biased, unfair) : :

- 14. Liberal (flexible, tolerant, progressive) : :
- 15. Arrogant (conceited, egotistic, vain) : :
- 16. Apathetic (unmotivated, indifferent, not interested
in accomplishments) : :
- 17. Sociable (extraverted, outgoing, friendly) : :
- 18. Aggressive (threatening, hostile, argumentative) : :
- 19. Open (trusting, self-disclosing, revealing) : :
- 20. Dull (unimaginative, slow-witted, unintelligent) : :
- 21. Independent (individualistic, self-reliant, autonomous) : :
- 22. Impatient (restless, uncommitted, fidgety) : :
- 23. Adventurous (takes risks, daring, enjoys
dangerous activities) : :
- 24. Thoughtless (unsympathetic, neglecting, uncaring) : :
- 25. Leadership (assertive, self-assured, dominant) : :
- 26. Disorganized (untidy, messy, sloppy) : :
- 27. Anxious (nervous, tense, uneasy) : :
- 28. Objective (just, unprejudiced, equitable) : :
- 29. Authoritarian (rigid, conservative, dogmatic) : :
- 30. Modest (humble, unpretentious, down-to-earth) : :

Background Information

1. Student Identification Number _____
2. Gender: M F
3. University level: A - second year
 B - third year
 C - fourth year
 D - other
4. Is this course in the area of your major? Y N
5. In general, university instruction is valuable.
A - strongly agree
B - agree
C - slightly agree
D - undecided
E - slightly disagree
F - disagree
G - strongly disagree
6. What was the extent of your interest in this course prior to the start of classes?
A - extremely interested
B - moderately interested
C - slightly interested
D - neither interested nor uninterested
E - slightly uninterested
F - moderately uninterested
G - extremely uninterested
7. What is the extent of your interest in this course currently?
A - extremely interested
B - moderately interested
C - slightly interested
D - neither interested nor uninterested
E - slightly uninterested
F - moderately uninterested
G - extremely uninterested

8. Compared to other courses, how much do you feel you have learned from this course?
- A - far more
 - B - considerably more
 - C - slightly more
 - D - neither more nor less
 - E - slightly less
 - F - considerably less
 - G - far less
9. Compared to other instructors, how much do you feel you have learned from your instructor, independent of course readings and assignments?
- A - far more
 - B - considerably more
 - C - slightly more
 - D - neither more nor less
 - E - slightly less
 - F - considerably less
 - G - far less
10. Would you take another course in this area taught by your instructor, if not restricted by degree requirements?
- A - definitely
 - B - probably
 - C - possibly
 - D - probably not
 - E - definitely not
11. Would you take another course in this area taught by a different instructor, if not restricted by degree requirements?
- A - definitely
 - B - probably
 - C - possibly
 - D - probably not
 - E - definitely not
12. What is your overall university grade average (percent)?
- %

13. To what extent do you feel the grade you will receive in this course reflects the grade you deserve?

- A - far higher than deserved
- B - considerably higher than deserved
- C - slightly higher than deserved
- D - neither higher nor lower than deserved
- E - slightly lower than deserved
- F - considerably lower than deserved
- G - far lower than deserved

THANK YOU FOR YOUR PARTICIPATION

APPENDIX 2: FEEDBACK TO SUBJECTS

The Relationship between Student Characteristics and Student Ratings of University Teaching Effectiveness

This report briefly describes major findings of the study entitled "Survey of Student Perceptions of University Instruction" in which you participated at the completion of the 1984/1985 academic year. Your responses to the survey were greatly appreciated.

As you may know, student ratings of university instructors are used for such purposes as: providing feedback for instructors, providing students with information for selecting classes, and providing the university with information for administrative decisions. Although students in a class typically differ in their ratings of teaching effectiveness, the overall rating of the instructor is obtained by averaging students' individual ratings. If differences between students are simply due to chance variation, then differences will cancel out and the average rating will be an accurate measure of teaching effectiveness. If, however, such differences in ratings are due to differences in characteristics of students, then average ratings may be affected. The purpose of this study, therefore, was to determine if differences in student characteristics can account for differences in ratings of teaching effectiveness.

An average of 28 students from each of six similar undergraduate classes taught by different male professors in the Department of Psychology, University of Western Ontario were included in the study. Each student provided ratings of instructor teaching effectiveness, instructor personality, and their own personality. Students also provided background information such as gender, university level, and university grade average.

In the six classes, the amount of variability in ratings of teaching effectiveness accounted for by student personality ranged from 8% to 34%, with an average of 17%. In 5 of the 6 classes, students rating themselves as low in anxiety rated their instructor as more effective than students high in anxiety. The direction and strength of the relationships between other personality traits and ratings of teaching effectiveness, however, differed in terms of both the class and the aspect of teaching effectiveness being examined. Differences between classes may be due to differences in personality traits of instructors teaching the classes. Analyses are currently being conducted to examine this hypothesis.

The amount of variability in ratings of teaching effectiveness explained by student gender ranged from 14% to 48%, with an average of 29%. Male students rated their instructor higher than did female students in each of the classes. It should be recalled, however, that only male instructors were included in this study. Whether the same relationship exists in classes taught by female instructors is a question for future research.

The amount of variability in ratings of teaching effectiveness explained by the university level of students ranged from 1% to 13%, with an average of 4%. The amount of variability explained by students' university grade average ranged from less than 1% to 14%, with an average of 3%. Again, the nature of these relationships differed in terms of both the class and the aspect of teaching effectiveness being examined. Overall, the amount of variability in ratings of teaching effectiveness explained by student characteristics (i.e., personality, gender, university level, and university grade average) ranged from 30% to 56%, with an average of 44%.

The findings of this study demonstrate that differences in student characteristics can account, to a substantial degree, for differences in student ratings of an instructor's teaching effectiveness. Average ratings of instructors, therefore, may be affected. For example, even if two instructors were equally effective, the instructor teaching a class with less anxious students would receive a higher average rating. When using average ratings, therefore, characteristics of students making the ratings should be considered.

The results of this study can be interpreted in two ways. First, some types of students may rate an instructor higher than other types for reasons other than the effectiveness of the instructor. Second, an instructor may actually be more effective for some types of students than others. The finding that relationships between many student characteristics and student ratings of teaching effectiveness differ from class to class seems consistent with this second interpretation. Some types of instructors may be most effective for some types of students while other types of instructors may be most effective for other types of students.

If more information regarding this study is desired, please contact:

Stephen Erdle
Department of Psychology,
University of Western Ontario,
London, Ontario. N6A 5C2

- Keaveny, J.J., & McGann, A.F. (1978). Behavioral dimensions associated with students' global ratings of college professors. Research in Higher Education, 9, 333-345.
- Kulik, J. A. & Kulik, C. C. (1974). Student ratings of instruction. Teaching of Psychology, 1, 51-57.
- Linn, R. L., Centra, J. A. & Tucker, L. R. (1975). Between, within, and total group factor analyses of student ratings of instruction. Multivariate Behavioral Research, 10, 277-288.
- Marsh, H.W. (1984). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases, and utility. Journal of Educational Psychology, 76, 707-754.
- McKeachie, W. J. (1966). Student affiliation motives, teacher warmth, and academic achievement. Journal of Personality and Social Psychology, 4, 457-461.
- Mintzes, J. J. (1979). Overt teaching behaviors and student ratings of instructors. Journal of Experimental Education, 48, 145-153.
- Mueller, F. J. (1951). Trends in student ratings of faculty. American Association of University Professors Bulletin, 37, 319-324.
- Murray, H. G. (1972). The validity of student ratings of faculty teaching ability. Presented at annual meetings of the Canadian Psychological Association.