Teacher Personality Traits and Student Instructional Ratings in Six Types of University Courses

Harry G. Murray, J. Philippe Rushton, and Sampo V. Paunonen
University of Western Ontario
London, Ontario, Canada

Colleague ratings of 29 personality traits were studied in relation to student ratings of teaching effectiveness in a sample of 46 psychology teachers. Instructors were evaluated in six different types of university courses, ranging from freshman lecture classes to graduate research seminars. Major findings were as follows: (1) Rated teaching effectiveness varied substantially across different types of courses for a given instructor; (2) teaching effectiveness in each type of course could be predicted with considerable accuracy from colleague ratings of personality; and (3) the specific personality traits contributing to effective teaching differed markedly for different course types. It was concluded that psychology instructors tend to be differentially suited to different types of courses and furthermore that the compatibility of instructors to courses is determined in part by personality characteristics.

Student ratings have gained widespread acceptance over the past 20 years as a measure of teaching effectiveness in North American colleges and universities. This trend has resulted in part from political factors and in part from research showing that student ratings can provide reliable and valid information on certain aspects of university teaching. Although findings are sometimes contradictory, the weight of evidence suggests that student ratings of a given instructor are reasonably stable across items, raters, and time periods; are affected to only a minor extent by extraneous factors such as class size and severity of grading; are consistent with comparable ratings made by alumni, colleagues, and trained classroom observers; and most important of all, are significantly correlated with more objective measures of teaching effectiveness, such as student performance on standardized examinations (Marsh, 1984; H. G. Murray, 1980). On the basis of these data, most writers have concluded that the use of student instructional ratings is justifiable both as a source of diagnostic feedback to instructors and as one of several measures of teaching effectiveness in administrative decisions on faculty salary, retention, tenure, and promotion.

Despite the abundance of research on the reliability, validity, and utility of student ratings, relatively little is known about characteristics of instructors that contribute to positive or negative evaluations from students. Given that teaching is in part a social or interpersonal process, it seems reasonable to expect that teacher personality traits might correlate significantly with rated teaching effectiveness. Although early investigations using self-report personality inventories failed to

support this hypothesis (e.g., Bendig, 1955; Sorey, 1968), more recent studies in which teacher personality was measured by colleague ratings, student ratings, or both, have yielded positive results (see review by Feldman, 1986). For example, H. G. Murray (1975) found that colleague ratings of instructor extraversion, leadership, objectivity, and (lack of) anxiety accounted for approximately 67% of between-teacher variance in student instructional ratings. In a study contrasting the personality profiles of "teachers" and "researchers," Rushton, Murray, and Paunonen (1983) reported correlations of .40 or higher between colleague ratings of extraversion, leadership, liberalism, supportingness, exhibitionism, objectivity, and lightheartedness and student ratings of teaching. Sherman and Blackburn (1975), using students rather than colleagues as judges of teacher personality, found that instructional effectiveness ratings were predictable from teacher potency, pragmatism, amicability, and intellectual competence. Also using student judges, Tomasco (1980) reported that teacher affiliation, achievement, endurance, nurturance, definitiveness, changeability, and exhibitionism were significant correlates of global effectiveness ratings. In summary, previous research suggests a reasonably consistent pattern of personality characteristics contributing to effective college teaching, in which successful teachers are viewed both by colleagues and by students as showing leadership, objectivity, and high intellect on the one hand, and extraversion, liberalism, and nurturance on the other.

The present study investigated relations between peer ratings of teacher personality traits and student ratings of teaching effectiveness in six types of university psychology courses. Previous research on student instructional ratings suggests that although evaluations of a given instructor are reasonably stable across different years for the same course, they are much less consistent across different courses (or course types) taught in the same year. H. G. Murray (1980) reported reliability coefficients ranging from .62 to .89 ($M = .74$) for the same course taught by the same instructor in successive years, as compared with reliability coefficients ranging from

.33 to .55 (*M* = .42) for different courses taught by the same instructor in the same year. Similarly, Marsh (1981) found an average correlation of .71 in same course/different year comparisons, as compared with an average correlation of .52 in different course/same year comparisons. The relatively low correlation of student ratings across different courses taught in the same year suggests that college teaching effectiveness may be to some extent context-dependent. In other words, instructors may be differentially suited to different types of courses rather than uniformly effective or ineffective in all types of courses. Another possibility, given previous research on teacher personality, is that differences in teaching effectiveness in different types of courses are predictable from teacher personality traits. For example, it may be that instructors with personality traits A, B, and C perform well in large lecture classes but poorly in small discussion groups, whereas instructors with personality traits X, Y, and Z tend to be good discussion leaders but poor lecturers. Consistent with this view, Sherman and Blackburn (1975) found that instructor pragmatism was related to student ratings in natural science courses but not in humanities or social science courses, whereas instructor amicability was related to teaching effectiveness in humanities courses but not in natural or social science courses.

The types of courses investigated in the present study ranged from freshman lecture classes to required methodology courses to graduate research seminars. Teacher personality traits were measured by peer ratings, whereas teaching effectiveness was assessed independently through archival student ratings. These procedures minimize the possibility of spurious correlations between personality and teaching because of "halo effect" or "implicit personality theory," a situation that can arise when all variables are rated by the same judges. On the basis of previous research, it was expected that instructors would differ in their relative standing in different types of courses, and, furthermore, that these differences would be related to instructor personality traits.

Method

Teachers

The sample of teachers consisted of 46 faculty members in the Department of Psychology at the University of Western Ontario. Each instructor had held a full-time appointment in the Department for at least 5 years. The breakdown of the sample in terms of gender and academic rank was as follows: 40 men and 6 women; 19 assistant professors, 17 associate professors, and 10 full professors. Each of the participating instructors had taught and received student ratings in at least 10 previous courses, including at least two courses from each of three or more of the categories defined below. Each instructor signed a consent form agreeing to participate in the study.

Courses

The Department of Psychology at the University of Western Ontario offers approximately 110 different courses (140 class sections) in a given year. Total student enrollment in the department is approximately 7,000. Undergraduate majors may complete either a 3-year "area of concentration" or a 4-year "honors" degree. Both MA and PhD programs are offered at the graduate level.

For purposes of the present research, psychology courses were divided into the six categories defined in Table 1. It may be noted that the six course types differed substantially in class size, student composition, and method of instruction. For example, class size ranged from 3 to 450, and method of teaching ranged from straight lecture to lecture-discussion to seminar.

Measures of Personality

Peer ratings of the 29 personality traits defined in Table 2 were obtained for each participating instructor. A full range of personality variables was included to represent the diversity of characteristics that might contribute to teaching effectiveness in six different types of courses. The first 20 personality traits in the table were derived from Jackson's (1984) Personality Research Form, an omnibus personality inventory based on H. A. Murray's (1938) need definitions. The last 2 traits, namely extraversion and neuroticism, were adapted from the Eysenck Personality Questionnaire (Eysenck & Eysenck, 1975), whereas the remaining 7 traits were selected because of their predictive validities in H. G. Murray's (1975) study of personality and college teaching. Peer ratings of personality were solicited from a rater group consisting of 48 full-time faculty members in the Department of Psychology. Included in this group were 36 of the 46 instructors serving as subjects in the present study. Each faculty rater was provided with a list of trait names and trait definitions, as in Table 2, and was asked to rate a random sample of 12 to 18 colleagues on each trait, using a 9-point rating scale. Instructions emphasized that ratings were to be based solely on personal observation and were to be made relative to other university professors rather than the population at large. The number of peer raters per instructor ranged from 9 to 15; the mean was 12.2. This corresponds to an overall return rate of 77.9%. Ratings were averaged across raters to obtain mean ratings of 29 personality traits for each of 46 instructors.

Measure of Teaching Effectiveness

Teaching effectiveness was measured from archival student rating data collected since 1969. The University of Western Ontario requires that student instructional ratings be solicited annually in all courses, with results to be considered in salary, promotion, and tenure decisions. The Department of Psychology uses two standardized student evaluation forms for this purpose, one for undergraduate courses and one for graduate courses. The undergraduate form consists of 9 items focusing on expositional skills of the instructor (e.g., clarity, use of examples), plus a final item assessing instructor "overall effectiveness." All 10 items are rated on 5-point scales. The graduate evaluation form consists of only two items, one dealing with course quality and the other with overall effectiveness of the instructor. Both items are rated on 7-point scales. In both undergraduate and graduate courses, student instructional ratings are obtained in a regular class period during the last 2 weeks of the academic term. Students respond anonymously to the questionnaire, the instructor is absent during the evaluation period, and results are released to the instructor only after final grades have been submitted. The criterion of teaching effectiveness used in the present study was the item assessing overall effectiveness of the instructor, which was common to both undergraduate and graduate evaluation forms. Thus, students in all types of courses rated the same global characteristic, namely overall teaching effectiveness, although sometimes on a 5-point and sometimes on a 7-point rating scale. The latter anomaly caused no difficulties in statistical analyses

Table 1
Characteristics of Six Types of Courses

Course type	Definition	Class size	Student composition	Method of instruction
Introductory	Multiple-section, freshman survey course	200–250	Nonmajors, Year 1	Lecture-laboratory
General	Elective survey courses for non-psychology students	150–450	Nonmajors, Years 2 & 3	Lecture
Required honors	Mandatory core courses in research methodology and experimental psychology	30–60	Psychology majors, Years 2, 3, & 4	Lecture-laboratory
Optional junior honors	Elective survey courses in basic content areas (e.g., social, developmental)	20–60	Psychology majors, Year 2	Lecture-discussion
Optional senior honors	Senior elective courses in special-ized topics	5–25	Psychology honors students only, Years 3 & 4	Seminar
Graduate	Advanced seminars for MA and PhD candidates	3–15	Graduate students	Seminar

because ratings were standardized separately for each course type in the computation or correlation coefficients.

Student instructional ratings were averaged across courses for each instructor (unweighted by class size) to obtain a cumulative mean effectiveness rating for each eligible course type. No data were computed in cases where an instructor had taught fewer than two courses in a given category. As noted previously, all 46 instructors had taught the requisite number of courses in at least three different course categories. However, only 6 instructors had taught at least two courses in all six categories, and only 21 instructors had taught at least two courses in five or more categories.

Results

Reliability of Personality Ratings

Table 2 shows the grand mean, standard deviation, and reliability of mean peer ratings of 29 personality traits for 46 instructors. It may be noted that the grand mean fell within 1 point of the midpoint of the 9-point rating scale (i.e., between 4.00 and 6.00) for 28 of 29 traits. Also, the standard deviation of mean peer ratings equalled or exceeded 1.00 for all 29 traits. These results suggest that peer ratings of instructor personality were distributed throughout the full range of the 9-point rating scale.

The reliability of peer ratings of personality was estimated using intraclass correlation procedures advocated by Shrout and Fleiss (1979). Each personality trait was subjected to a one-way analysis of variance (ANOVA) in which teachers served as "treatments" and raters as "subjects." Within- and between-teacher mean squares from the ANOVA were then used to estimate the reliability of the mean rating of k raters (mean k = 12.2). Reliability coefficients computed in this way ranged from .69 to .94 and averaged .83, indicating that peer assess-

ments of personality showed substantial interrater reliability, or generalizability across potential sets of raters.

Reliability of Teacher Effectiveness Ratings

Table 3 indicates the number of instructors teaching at least two courses within each of the six course categories, as well as the grand mean, standard deviation, and retest reliability of mean student instructional ratings for the sample of instructors teaching each course type. For example, it may be noted that for the 29 of 46 instructors who taught introductory psychology at least twice, the grand mean and standard deviation of cumulative mean student ratings for this type of course were 3.27 and .67, respectively. Student ratings were higher for graduate than for undergraduate courses, reflecting the use of a 7-point rating scale in the former case. Also, consistent with previous research, ratings were higher in senior than in junior courses and higher in optional than in required courses (Feldman, 1978).

The reliability of teacher effectiveness ratings was estimated by intraclass correlation procedures as described above, with the exception that within-teacher mean squares were computed across k years per instructor (minimum k = 2, M = 4.66), rather than across raters. Retest reliabilities ranged from .78 to .93 and averaged .86, indicating that student instructional ratings were highly stable across years for a given teacher and a given type of course.

Consistency of Teacher Ratings Across Course Types

Table 4 shows correlations between instructor effectiveness ratings in different types of courses (entries above the diagonal), as well as the number of instructors teaching each

Table 2

Grand Mean, Standard Deviation, and Reliability of Mean Peer Ratings of Instructor Personality (N = 46)

Personality trait	Abbreviated trait definition	M	SD	Rater reliability
Meek	Mild-mannered, accepts blame or criticism, subservient	4.07	1.24	.86
Ambitious	Aspires to accomplish difficult tasks, maintains high standards	5.87	1.60	.94
Sociable	Friendly, outgoing, enjoys being with people	5.42	1.26	.87
Aggressive	Argumentative, quarrelsome, gets angry easily	4.79	1.46	.89
Independent	Avoids restraints and confinement, enjoys being free	5.43	1.04	.77
Changeable	Flexible, restless, likes new and different experiences	4.68	1.11	.79
Seeks definiteness	Does not like ambiguity or uncertainty, seeks structure	5.78	1.12	.87
Defensive	Suspicious, guarded, takes offense easily	4.70	1.23	.85
Dominant	Forceful, decisive, attempts to control environment	5.20	1.61	.90
Enduring	Perservering, steadfast, does not give up quickly	6.03	1.60	.87
Attention-seeking	Dramatic, colorful, wants to be center of attention	4.76	1.51	.89
Harm-avoiding	Careful, cautious, avoids excitement or danger	4.82	1.19	.78
Impulsive	Spontaneous, impetuous, acts on spur of moment	4.76	1.39	.85
Supporting	Gives sympathy and comfort, helpful, indulgent	5.15	1.35	.85
Orderly	Neat and organized, dislikes clutter and confusion	5.43	1.33	.88
Fun-loving	Easygoing, playful, does things just for fun	5.02	1.38	.89
Aesthetically sensitive	Sensitive to sights, sounds, tastes, and other experiences	5.21	1.14	.89
Approval-seeking	Works for approval and recognition of others, agreeable	5.28	1.07	.75
Seeks help and advice	Desires and needs support, sympathy, and advice from others	4.29	1.33	.82
Intellectually curious	Reflective, seeks understanding and synthesis of ideas	5.85	1.26	.85
Anxious	Tense, nervous, uneasy	4.82	1.10	.78
Intelligent	Bright, quick, clever, excels in general cognitive ability	5.93	1.38	.88
Liberal	Progressive, modern, adaptable, seeks change	5.24	1.14	.81
Shows leadership	Takes initiative and responsibility for getting things done	5.20	1.53	.91
Objective	Just, fair, free of bias	5.42	1.21	.76
Compulsive	Meticulous, perfectionistic, concerned with details	5.65	1.20	.76
Authoritarian	Rigid, inflexible, obedient to authority, opinionated	4.77	1.11	.80
Extraverted	Has many friends, likes parties, craves excitement, optimistic	4.94	1.48	.89
Neurotic	Emotional, moody, constantly worried things will go wrong	4.76	1.00	.69

Table 3

Grand Mean, Standard Deviation, and Reliability of Mean Student Ratings of Instructor Effectiveness in Six Types of Courses

Course type	N	M	SD	Retest reliability
Introductory	29	3.27	0.67	.91
General	30	3.27	0.57	.80
Required honors	29	3.20	0.63	.87
Optional junior honors	35	3.54	0.59	.93
Optional senior honors	38	3.66	0.62	.88
Graduate	40	5.66	0.66	.78

combination of course types (entries below the diagonal). For example, 23 instructors taught two or more courses in both introductory and optional junior honors categories, and the correlation between instructor effectiveness ratings in these two types of courses was .52. Although all instructors did not teach all types of courses or all combinations of course types, a chi-square analysis showed that there was no significant tendency for instructors teaching one type of course to be either more or less likely to teach another type of course than would be expected on the basis of marginal frequencies, $\chi^2(14, N = 46) = 4.32, p > .05$. It would appear, then, that correlations between course types were not appreciably distorted by overrepresentation or underrepresentation of certain groups of instructors in certain combinations of course types.

It may be noted in Table 4 that correlations between instructor mean ratings in different types of courses were positive in all cases. Correlation coefficients ranged from .06 to .78, and averaged .49. This result suggests that, although there appears to be a general or g factor in college teaching performance, instructor effectiveness can vary substantially across different types of courses. In other words, receiving high or low instructional ratings in one type of course is no guarantee that ratings will be similarly high or low in another type of course. Teacher ratings showed much higher consistency across undergraduate course types (mean $r = .66$) than for undergraduate versus graduate course types (mean $r = .15$). Furthermore, for each type of course, the retest reliability coefficient shown in Table 3 was higher than any of the between-category correlations shown in Table 4. This result is consistent with previous evidence that instructor ratings correlate higher across years for the same course than across

different courses taught in the same year (e.g., Marsh, 1981). It should be noted that for undergraduate courses, both between-year and between-course correlations (mean values = .88 and .66, respectively) were higher in absolute terms in the present study than in previous studies, presumably because of aggregation of teacher ratings across several courses in this study.

Further information on the cross-situational consistency of teaching performance was obtained by transforming each instructor's mean rating for each applicable course type to a standard (z) score computed in relation to the overall mean and standard deviation of ratings for that type of course. The average range of z scores *within* instructors was 1.48. This result implies that an instructor who scores at the 32nd percentile of the ratings distribution ($z = -0.48$) in his or her worst course type would, on average, score at the 84th percentile ($z = 1.00$) in his or her best course type. Although there is no denying the between-course consistency that exists in these data, it would appear that teaching effectiveness can nonetheless vary markedly for a given instructor as a function of course type. An even more dramatic demonstration of this point is the fact that only 3 of the 46 instructors participating in this study scored in the top third of the department ($z > 0.45$) in all course types taught, and only 3 scored in the bottom third of the department ($z < -0.45$) in all course types taught. These results show that it is rare indeed for an instructor to be either uniformly "good" or uniformly "poor" in all types of teaching.

Zero-Order Correlations Between Personality and Teaching

Given that instructor ratings show less than perfect consistency across different types of courses, the next question is whether between-course differences in teaching effectiveness are related to instructor personality characteristics. In other words, do teachers who perform well in certain types of courses share certain personality traits that differ from those of teachers who excel in other types of courses? Table 5 shows zero-order correlations between peer ratings of personality and student ratings of teaching effectiveness in different types of courses. The first noteworthy finding is that, consistent with previous research (e.g., Rushton et al., 1983), there was a substantial overall or main effect relationship between personality and teaching. Seventy-two of 174 correlations between personality traits and teacher ratings were statistically

Table 4

Intercorrelation of Instructor Mean Teaching Effectiveness Ratings in Six Types of Courses

Course type	1	2	3	4	5	6
1. Introductory	—	.78	.78	.52	.66	.12
2. General	20	—	.60	.53	.62	.06
3. Required honors	21	15	—	.76	.68	.11
4. Optional junior honors	23	24	17	—	.71	.13
5. Optional senior honors	24	27	20	31	—	.33
6. Graduate	26	27	22	31	36	—

Note. Below-diagonal entries are cell *ns*.

Table 5

Correlations Between Peer Ratings of Personality and Student Ratings of Teaching in Six Types of Courses (decimals omitted)

Personality trait	Introductory	General	Required honors	Optional junior honors	Optional senior honors	Graduate
				Course type		
Meek	02	−30	−11	06	21	02
Ambitious	24	19	31	14[a]	03	43*
Sociable	64[a]*	33	37*	55*	57*	−04
Aggressive	07	20	12	−03	−21[a]	−08
Independent	−09	13	22	09	−03	12
Changeable	46*	37[a]*	49*	50*	44*	−08
Seeks definiteness	06	12	34	02	−16	32*
Defensive	−35	−16	−19	−34*	−56*	−24
Dominant	30	36*	40*	30	01	08
Enduring	29	26	38[a]*	25	18	52[a]*
Attention-seeking	56*	50[a]*	42*	51*	29	−18
Harm-avoiding	−16	−28	−26	−45[a]*	−48[a]*	−06
Impulsive	36*	39*	01	37*	12	−31*
Supporting	53*	11[a]	43*	58[a]*	64*	11
Orderly	34[a]	25[a]	37[a]*	07	25[a]	29*
Fun-loving	56[a]*	32	35	49*	48*	−22
Aesthetically sensitive	27	40*	18	40*	52[a]*	−22
Approval-seeking	48*	29	16	24	12	11[a]
Seeks help and advice	46*	15	14	32	34*	−34[a]*
Intellectually curious	29	38*	42*	23	24	46*
Anxious	−22	−16	−35[a]	−24	−41*	−19
Intelligent	28	37*	42*	19	11	35*
Liberal	60*	52[a]*	35[a]	55*	65[a]*	−06
Shows leadership	51*	40*	56[a]*	54*	43*	44[a]*
Objective	40*	20	31	34[a]*	59*	40[a]*
Compulsive	32	20	39*	12[a]	11	46*
Authoritarian	−34[a]	−24	−12	−30	−42*	05
Extraverted	64[a]*	41*	50*	54*	45*	−12
Neurotic	01	03	−15	−08	−25	−27
Mean *r*	34	28	30	30	32	23
Stepwise *R* (*k* = 5)	.801*	.786*	.737*	.771*	.877*	.695*
Variance explained (adjusted *R²*)	56.3%	53.7%	43.0%	52.6%	73.2%	41.5%

[a] One of five variables used to obtain stepwise *R* for each course type.
* $p < .05$.

significant, whereas only 8.7 (0.5 × 174) significant correlations would be expected under the null hypothesis. The mean absolute value of the 174 correlations was .295. Secondly, it is apparent that teaching effectiveness in each of the six types of courses was separately predictable from ratings of instructor personality. Each course type showed at least 10 significant correlations with individual personality traits, and the mean absolute correlation with personality variables ranged from .23 (for graduate courses) to .34 (for introductory courses). Stepwise multiple regression analysis yielded significant multiple *R*s ranging from .695 to .877 between the best five personality predictor variables for a given type of course and cumulative mean teacher ratings. With correction for shrinkage, corresponding estimates of variance accounted for (adjusted *R²*) ranged from 41.5% to 73.2%, indicating that teaching effectiveness in each type of course could be predicted with considerable accuracy from peer ratings of as few as five personality traits. A similar analysis of mean standard scores for all course types taught by a given instructor yielded a multiple *R* of .754 and an adjusted *R²* of .521, indicating that

composite or global teaching effectiveness is also highly predictable from personality variables. The five best predictors of the composite effectiveness measure were leadership, liberalism, seeks definiteness, supportingness, and extraversion.

The third and most important finding in Table 5 is that the pattern of relations between personality and teaching differed markedly for different types of courses. In other words, personality traits that contributed to effective teaching in one type of course did not necessarily contribute similarly in other types of courses. For example, the traits of sociability, changeableness, attention-seeking, liberalism, and extraversion correlated positively and significantly with teacher ratings in undergraduate courses, whereas the same traits tended to show nonsignificant negative correlations with teacher ratings in graduate courses. On the other hand, ambitiousness, endurance, orderliness, and compulsiveness were positively related to performance in graduate courses and in required honors courses, but did not correlate significantly with performance in other types of courses. Two other traits, impulsiveness and seeking help and advice, correlated significantly but in oppo-

site directions with instructional ratings in different types of courses. In general, it appears that different and sometimes incompatible combinations of personality traits are necessary for effective performance in different types of courses, which of course would explain why very few teachers are outstanding in all types of courses taught. It is interesting to note that only 1 of the 29 personality traits, namely leadership, correlated significantly and in the same direction with instructor ratings in all six types of courses. It may be that teachers with leadership ability are sufficiently flexible to adapt or modify their teaching style to the demands of different types of courses.

Factor Analysis

To derive a smaller set of uncorrelated personality variables for use in further analyses, instructor mean ratings on the 29 personality scales were subjected to a principal-components, varimax-rotation factor analysis. As indicated in Table 6, the analysis yielded five factors with eigenvalues greater than 1.0 that collectively accounted for 86.6% of the total variance in mean personality ratings. The orthogonal, 5-factor solution

provided good approximation to simple structure. Most traits loaded highly on only one factor, and each factor was defined by at least three high-loading traits. Factor 1, interpreted as Extraversion, accounted for 29.0% of ratings variance. High scorers on this factor were perceived by peers as friendly, lighthearted, colorful, and charismatic. Factor 2, Achievement, accounted for 26.3% of variance. High scorers were rated as dominant, intelligent and hardworking. Factor 3, Negative Affect, which accounted for 14.8% of variance, was a bipolar factor on which high scorers were perceived as aggressive, defensive, and impulsive, and low scorers as objective, mild-mannered, and supporting. Factor 4, Liberalism, accounted for 11.3% of variance and was defined by traits such as aesthetic sensitivity, flexibility, and nonauthoritarianism. Factor 5, interpreted as Neuroticism, accounted for 5.2% of variance and included traits such as fearfulness and cautiousness.

Instructors were assigned factor scores on each of the above dimensions according to the SPSS factor score procedure (Nie, Hull, Jenkins, Steinbrenner, & Bent, 1975). Factor scores computed in this way had means of 0, standard deviations of 1.0, and intercorrelations of 0. Alpha reliabilities of factor

Table 6
Factor Loadings From Principal-Components Factor Analysis of Peer-Rated Personality Traits (decimals omitted)

Personality trait	Factor				
	I Extraversion	II Achievement	III Negative affect	IV Liberalism	V Neuroticism
Extraverted	93	−04	15	23	−16
Sociable	91	00	−23	20	−07
Attention-seeking	82	12	45	20	−05
Fun-loving	74	−21	−02	48	−24
Approval seeking	74	23	−18	−10	38
Shows leadership	71	58	−01	−22	−18
Seeks help and advice	65	−36	−32	28	48
Enduring	06	92	−01	−13	−09
Intelligent	04	92	14	18	−09
Intellectually curious	03	91	00	25	−14
Ambitious	11	91	16	−06	−09
Seeks definiteness	−13	81	23	−21	09
Compulsive	00	70	−10	−49	33
Dominant	44	60	58	−15	−06
Aggressive	18	36	86	03	06
Objective	25	24	−84	11	−20
Meek	−14	−45	−79	04	16
Defensive	−15	−07	75	−29	44
Supporting	53	−04	−74	−23	−07
Independent	−15	47	65	35	−30
Impulsive	56	−04	64	41	04
Aesthetically sensitive	35	−07	−15	76	08
Liberal	59	17	−25	75	−13
Authoritarian	−22	06	52	−70	17
Changeable	55	22	09	70	−17
Orderly	07	44	−15	−63	19
Neurotic	05	−10	15	−06	93
Anxious	−19	−03	19	−12	89
Harm-avoiding	−37	−09	−40	−38	54
Eigenvalue	8.41	7.63	4.29	3.27	1.52
Cumulative variance	29.0	55.3	70.1	81.4	86.6
Alpha reliability	.90	.93	.88	.84	.79

Note. Boxes indicate traits that loaded highest on a given factor.

scores (based solely on the marker variables designated in Table 6) ranged from .79 to .93.

Analysis of Factor Scores

A preliminary analysis of personality factor scores was undertaken to determine whether instructors who had taught a particular type of course at least twice differed significantly in mean scores on any of the five personality factors from instructors who had not taught that type of course. Of 30 such comparisons (6 course types × 5 personality factors), only 1 was statistically significant. Given that approximately 1.5 significant contrasts would be expected by chance alone under these circumstances, it seems reasonable to attribute the obtained result to Type I error and to conclude that assignment of instructors to course types was not systematically related to instructor personality. A direct implication of this conclusion is that the occurrence of missing data for particular instructor/course type combinations was unrelated to instructor personality. Results presented earlier indicated that the presence or absence of data for a given instructor/course type combination was unrelated to the presence or absence of data for other types of courses. Thus, it would

appear that the occurrence of missing data in the present study was random with respect to both of the independent variables under investigation (i.e., personality factors and course types) and that, for this reason, the results obtained are representative of what would have been found with a complete data set (Cohen & Cohen, 1975).

Factor scores were next analyzed to determine personality profiles of highly successful instructors in different types of courses. Figure 1 shows mean factor score profiles for instructors whose cumulative mean ratings ranked in the top 16% ($z > 1.0$) among those teaching each type of course. The number of "top teachers" varied slightly for different course types as a function of the total number of instructors teaching each type of course (see Table 3). A two-way ANOVA, in which course type was treated as a between-subjects variable (despite some overlap in membership among high-ranking groups) and personality factor was a within-subjects variable, yielded a significant interaction effect, $F(20, 112) = 3.03$, $MS_e = .294$, $p < .001$, reflecting the fact that personality profiles differed substantially across course types. Inspection of Figure 1 suggests that the six course types can be divided into three subcategories with respect to profile shape, or in other words, with respect to the relative contribution of different person-
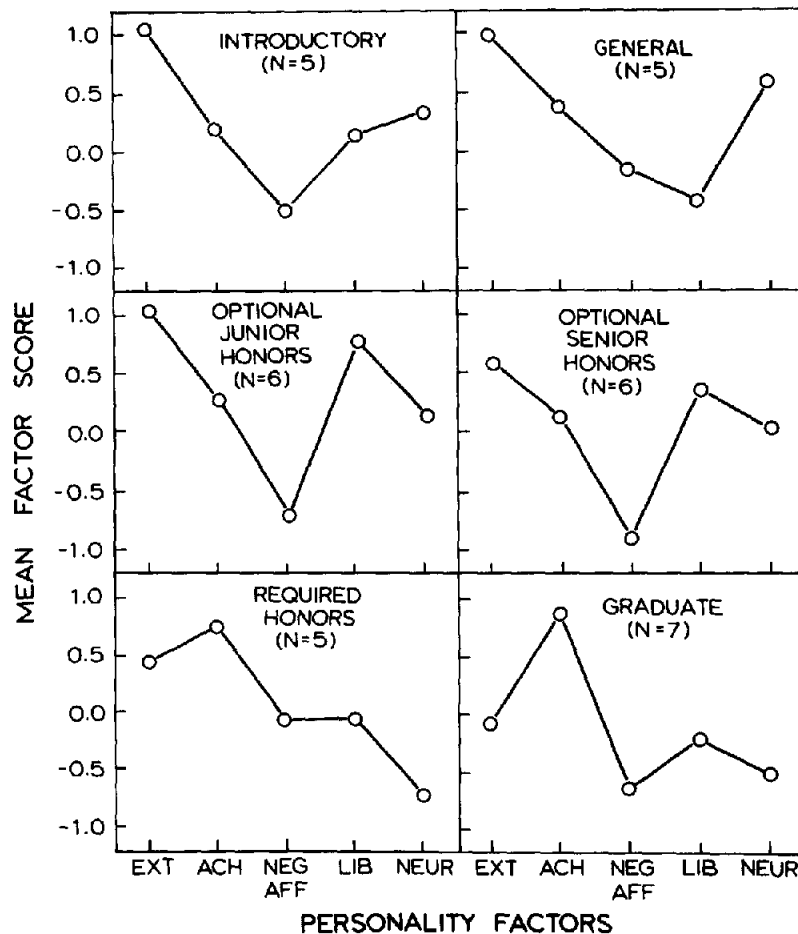

Figure 1. Mean personality factor score profiles for top-ranking instructors in six types of college courses. (EXT = extraversion; ACH = achievement; NEG AFF = negative affect; LIB = liberalism; NEUR = neuroticism.)

ality factors to teaching effectiveness. The first category, consisting of introductory and general courses, is characterized by elevated scores on the Extraversion and Neurosis factors and by relatively low scores on Negative Affect. Apparently the instructor who excels in large, lower-level, lecture classes is a "neurotic extravert" type who is friendly, warm, and approachable, has a flair for the dramatic, and is fair and reasonable in relations with students, but shows an element of neurotic worrying. Perhaps an individual with these characteristics enjoys the stimulation of speaking before a large audience, but nonetheless is compulsive enough to attend to the myriad of details involved in organizing and orchestrating a large lecture course. The second category of course types, consisting of optional junior and optional senior honors courses, is defined by high scores on the Extraversion and Liberalism factors and by very low scores on Negative Affect. It appears that successful teachers in smaller, higher-level, discussion-oriented classes tend to be friendly, gregarious, fair, and supportive, and, at the same time, flexible, adaptable, and open to change. Presumably this is the type of person who is at ease in a discussion format where students exert greater control over classroom activities. The third category of course types consisted of required honors and graduate courses and was defined by high scores on the Achievement factor, average scores on Liberalism and Negative Affect, and low scores on Neuroticism. Instructors who excel in required methodology courses and in graduate seminars appear to be ambitious, competent, and hard working, and, at the same

time, confident and worry free. One possible reason that high-achievement instructors are effective in methodology and in graduate courses is that inspiring students to work hard (either through modeling or through exhortation) is a prerequisite for success in these types of courses. In general, the picture that emerges from Figure 1, consistent with that of Table 5, is that the personality profile of the effective teacher differs substantially for different types of university courses.

Another factor score analysis was performed to identify personality characteristics of instructors who, contrary to the general rule, were uniformly effective or uniformly ineffective in all types of courses taught. Mean personality profiles were computed for 8 teachers who taught four or more course types and received above-average cumulative ratings in all cases, and for 7 teachers who taught four or more course types with below-average ratings in all cases. These data are plotted in Figure 2. The most distinguishing characteristics of uniformly "good" teachers were high scores on the Extraversion and Liberalism factors. On the other hand, uniformly "poor" teachers were characterized by very low scores on Extraversion and somewhat elevated levels of Neuroticism.

A two-way ANOVA of the data in Figure 2 yielded a significant Groups × Factors interaction effect, $F(1, 52) = 4.21$, $MS_e = .201$, $p < .05$. Follow-up Tukey Honestly Significant Difference tests showed that group differences were significant only for the Extraversion and Liberalism factors, suggesting that these two personality dimensions are critical in determining breadth of teaching effectiveness. Comparison of uni-
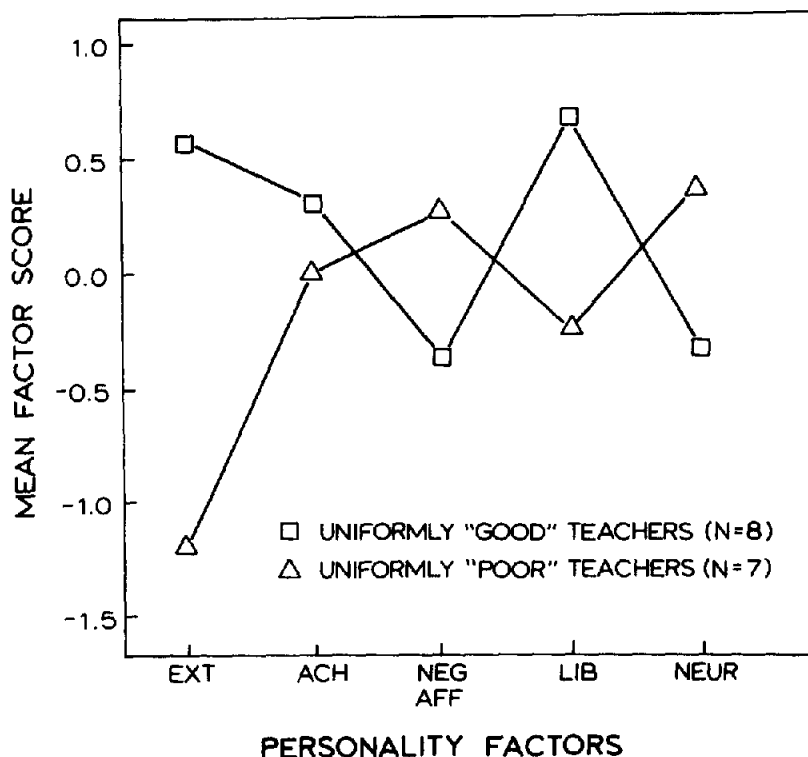

Figure 2. Mean personality factor score profiles for instructors receiving uniformly high or uniformly low student ratings in all types of courses taught. (EXT = extraversion; ACH = achievement; NEG AFF = negative affect; LIB = liberalism; NEUR = neuroticism.)

formly good and uniformly poor teachers on the 12 individual personality traits loading highest on Extraversion and Liberalism in the factor analysis reported in Table 6 revealed that the trait contributing most to the differentiation of these two groups was leadership. This result is consistent with the fact that leadership was the only personality trait to correlate significantly with instructor ratings in all six types of courses (see Table 5). As elaborated further below, leadership qualities appear to play a pivotal role in university teaching effectiveness.

Discussion

In summary, the results of this research suggest three major conclusions concerning teacher personality traits in relation to instructional effectiveness in different types of university-level psychology courses. First, clear evidence was found that perceived teaching effectiveness does in fact vary substantially across different types of courses for the same instructor. Correlations between cumulative instructor ratings in different types of courses ranged from .06 to .78, and averaged .49, indicating that even with aggregated and reliable measures, college teaching effectiveness shows only moderate cross-situational consistency. In other words, it is the exception rather than the rule for an instructor to perform exceptionally well or exceptionally poorly in all types of courses. The second major conclusion is that, for any given type of course or for all types combined, student instructional ratings were strongly related to peer ratings of instructor personality traits. Consistent with previous research (e.g., Rushton et al., 1983), multiple regression analyses showed that 40 to 70% of between-teacher variance in student instructional ratings was predictable from peer ratings of as few as five personality traits. Personality traits correlating highest with composite teacher effectiveness ratings included leadership, extraversion, liberalism, supportingness, intellectual curiosity, and changeableness. The third and most important finding of this study is that the specific personality traits contributing to effective teaching varied substantially for different types of courses. For example, instructor extraversion and liberalism correlated positively and significantly with student ratings in undergraduate courses, whereas these same traits showed zero or negative correlations with ratings in graduate courses. Conversely, instructor achievement and endurance contributed significantly to teaching effectiveness in graduate and required honors courses, but were unrelated to performance in other types of courses.

The results of this study provide evidence that, at least in the field of psychology, university teachers tend to be differentially suited to different types of courses rather than uniformly effective or ineffective in all types of courses. Furthermore, the compatibility of teachers to courses appears to be determined in part by personality characteristics. According to this interpretation, very few teachers will have the necessary range of personality traits to excel in all types of courses. One of the secrets of effective teaching, therefore, is to discover the conditions under which one teaches most effectively. Gage and Berliner (1984) stated this principle as follows:

> Just as plays and movies require casting, and not every actor is suitable for every role, so teaching methods require matching with the strengths and weaknesses of the teacher. If a teacher's personality is unsuited for lecturing, it will be much more worthwhile for that instructor to choose a method other than lecturing than to try to learn to use the method effectively. (p. 482)

The present findings also have important implications for the validity and utility of student instructional ratings, and more generally, for the improvement of university teaching. Some writers (e.g., Small, Hollenbeck, & Haley, 1982) have argued that a strong relationship between teacher personality and student instructional ratings implies that student ratings are invalid, in that they reduce to nothing more than a "personality context." Ware and Williams (1980) reached a similar conclusion in their analysis of the Dr. Fox effect, defined as a prepotent influence of instructor charisma or expressiveness on student ratings. An alternative interpretation, favored by the present authors, is that personality traits of the instructor (e.g., orderliness) are translated into specific classroom teaching behaviors (e.g., putting an outline on blackboard), which in turn are validly reflected in student ratings. In support of this view, Erdle, Murray, and Rushton (1985) showed by path analytic procedures that more than 50% of the relationship between teacher personality and student ratings was mediated by specific classroom behaviors. It is not unreasonable to conclude, therefore, that a correlation between teacher personality and student ratings provides positive (rather than negative) evidence with respect to the validity of student ratings, in that ratings are shown to be systematically related to pedagogically relevant instructor characteristics.

The low to moderate correlations found in the present research between student ratings in different types of courses suggest that when instructional ratings are used in administrative decisions on faculty retention, tenure, and promotion, data should be available from as many types of courses as possible. Otherwise, an instructor denied tenure or promotion on the basis of poor teaching evaluations in a limited range of course types could rightfully argue that more favorable evaluations might have been obtained in other types of courses. A further implication of the modest correlation of ratings across different course types is that the overall quality of teaching in an academic department could be significantly improved simply by assigning teachers to the types of courses in which they have received their highest ratings. Ideally this should be a posttenure arrangement, instituted only after the instructor has had the opportunity to teach several different types of courses. As a case in point, if each of the 46 instructors participating in the present study was in the future assigned only to the two types of courses in which he or she had performed best, the department mean teacher rating for undergraduate courses would be expected to increase from 3.41 to 3.75 (on a 5-point scale), whereas the department mean for graduate courses would increase from 5.66 to 6.06 (on a 7-point scale).

Assuming that peer ratings provide a valid index of stable, enduring personality characteristics, the present findings suggest that a faculty member's effectiveness as a teacher is determined, at least in part, before he or she sets foot in the classroom. In other words, preexisting personality traits (among other factors) are assumed to determine the instructor's classroom teaching behaviors, which in turn determine student ratings and student achievement. Given that current

selection and training procedures virtually guarantee that college teachers will be more heterogeneous with respect to personality characteristics than with respect to cognitive or intellectual variables, it is perhaps not surprising that personality should account for such a large proportion of variance in college teaching effectiveness. Whatever the reasons, the existence of strong correlations between personality and teaching suggests that personality measures (e.g., peer ratings) might be used in selecting new faculty members who are likely to be effective teachers. Alternatively, such measures might be used to assign instructors to the types of courses in which they are likely to be most successful. According to the present research, faculty who are extraverted, dramatic, and perhaps a bit neurotic would be expected to excel in lower-level, lecture courses; whereas those who are extraverted, liberal, and fair-minded might be predicted to do well in higher-level discussion or seminar courses; and those who are ambitious, hard-working, and organized would perhaps be most successful in graduate seminars and undergraduate methodology courses. Finally, faculty who are liberal, flexible, and high in leadership ability would be expected to excel in a wide range of different types of courses.

It is noteworthy that the personality traits of leadership and flexibility (liberalism) were found to be associated with generalized or wide-ranging teaching effectiveness in the present study. Previous writers have suggested, on the one hand, that teaching and leadership are similar in many ways (e.g., Norr & Crittenden, 1975) and, on the other hand, that a good teacher is, above all else, flexible and adaptable in approach (e.g., Berliner, 1976). Possibly the superordinate trait of "role flexibility" is part of what defines both an effective leader and a master teacher—that is, a teacher who excels in all types of courses. A master teacher, like an effective leader, presumably has the ability to modify his or her approach so as to adapt successfully to the requirements of different situations. Teachers who excel in all types of courses are capable of being either friendly and supportive or strict and demanding, either student-centered or teacher-centered, either a "therapist" or a "drillmaster," depending on what will accomplish the task most effectively. Lowman (1984) offers a similar analysis in his characterization of "complete master" teachers. These rare individuals excel both in expository and in interpersonal skills, and are able to modify their approach so as to motivate both brilliant and mediocre students and to perform well in both the lecture hall and the seminar room.

An obvious limitation of the present research is that the correlational design used does not allow direct, unambiguous interpretation of cause-effect relationships. Whereas the preceding discussion has assumed that personality traits determine teaching behaviors, which in turn determine perceived teaching effectiveness, it is conceivable that causality operates in the opposite direction, such that prior successes or failures in teaching lead to systematic changes in instructor personality characteristics. Alternatively, it is possible that actual or hearsay knowledge of instructors' prior student ratings influenced colleague assessments of personality traits. Although neither of these alternative causal models can be unambiguously eliminated with the data at hand, their plausibility is reduced both by the normal temporal precedence of personality traits

to classroom teaching behaviors and by H. G. Murray's (1975) finding that student ratings of new, previously unrated college instructors could be accurately predicted from peer ratings of personality traits obtained at least 5 months prior to student assessment of teaching.

A final comment should be added concerning possible implications of the present study for the controversial issue of whether teaching effectiveness and research productivity are positively related in university faculty members. Although many faculty are steadfastly confident that these factors are (or should be) positively associated, empirical studies have consistently reported a correlation of close to zero between measures of teaching quality and measures of research productivity (e.g., Feldman, 1987; Rushton et al., 1983). It appears, however, that no previous study has investigated research productivity in relation to rated teaching effectiveness in different types of courses. The present finding that instructor ambitiousness, endurance, compulsiveness, and intelligence correlated significantly with student ratings in graduate seminars and undergraduate methodology courses, in combination with Rushton et al's (1983) finding that these same personality traits correlated significantly with faculty publication rates and citation counts, implies that research productivity may in fact be positively related to teaching effectiveness, but only for certain very research-oriented or "work-oriented" types of courses.

References

Bendig, A. W. (1955). Ability and personality characteristics of introductory psychology instructors rated competent and empathetic by their students. *Journal of Educational Research, 48,* 705–709.

Berliner, D. C. (1976). Impediments to the study of teacher effectiveness. *Journal of Teacher Education, 27,* 5–13.

Cohen, J., & Cohen, P. (1975). *Applied multiple regression/correlation analysis of the behavioral sciences.* Hillsdale, NJ: Erlbaum.

Erdle, S., Murray, H. G., & Rushton, J. P. (1985). Personality, classroom behavior, and college teaching effectiveness: A path analysis. *Journal of Educational Psychology, 77,* 394–407.

Eysenck, H. J., & Eysenck, S. B. G. (1975). *Manual of Eysenck Personality Questionnaire.* San Diego, CA: Educational and Industrial Testing Service.

Feldman, K. A. (1978). Course characteristics and college students' ratings of their teachers: What we know and what we don't. *Research in Higher Education, 9,* 199–242.

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.

Feldman, K. A. (1987). Research productivity and scholarly accomplishment of college teachers as related to their instructional effectiveness: A review and exploration. *Research in Higher Education, 26,* 227–298.

Gage, N. L., & Berliner, D. C. (1984). *Educational psychology* (3rd ed.). Boston: Houghton-Mifflin.

Jackson, D. N. (1984). *Personality Research Form Manual.* Port Huron, MI: Research Psychologists Press.

Lowman, J. (1984). *Mastering the techniques of teaching.* San Francisco: Jossey-Bass.

Marsh, H. W. (1981). The use of path analysis to estimate teacher and course effects in student ratings of instructional effectiveness. *Applied Psychological Measurement, 6,* 47–60.

Marsh, H. W. (1984). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases, and utility. *Journal of Educational Psychology, 76,* 707–754.

Murray, H. A. (1938). *Explorations in personality.* Cambridge, MA: Harvard University Press.

Murray, H. G. (1975). Predicting student ratings of college teaching from peer ratings of personality types. *Teaching of Psychology, 2,* 66–70.

Murray, H. G. (1980). *Evaluating university teaching: A review of research.* Toronto: Ontario Confederation of University Faculty Associations.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. H. (1975). *Statistical package for the social sciences.* New York: McGraw-Hill.

Norr, J. L., & Crittenden, K. S. (1975). Evaluating college teaching as leadership. *Higher Education, 4,* 335–350.

Rushton, J. P., Murray, H. G., & Paunonen, S. V. (1983). Personality, research creativity, and teaching effectiveness in university professors. *Scientometrics, 5,* 93–116.

Sherman, B. R., & Blackburn, R. T. (1975). Personal characteristics and teaching effectiveness of college faculty. *Journal of Educational Psychology, 67,* 124–131.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86,* 420–428.

Small, A. C., Hollenbeck, A. R., & Haley, R. L. (1982). The effect of emotional state on student ratings on instructors. *Teaching of Psychology, 9,* 205–211.

Sorey, K. E. (1968). A study of the distinguishing personality characteristics of college faculty who are superior in regard to the teaching function. *Dissertation Abstracts International, 28,* 4916A. (University Microfilms No. 68-8502)

Tomasco, A. T. (1980). Student perceptions of instructional and personality characteristics of faculty: A canonical analysis. *Teaching of Psychology, 7,* 79–82.

Ware, J. E., & Williams, R. G. (1980). A reanalysis of the Dr. Fox experiments. *Instructional Evaluation, 4,* 15–18.