DOCUMENT RESUME

ED 129 903 TM 005 746

AUTHOR McCook, William M.
TITLE A Multivariate Study of Variables Effecting Student Ratings of Teaching and Course Outcomes Within a Multiple Instructor Mode.
SPONS AGENCY National Institutes of Health (DHEW), Bethesda, Md. Bureau of Health Manpower Education.
PUB DATE Apr 76
NOTE 23p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (San Francisco, California, April 19-23, 1976); Not available in hard copy due to marginal legibility of original document

EDRS PRICE MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS College Students; College Teachers; *Course Evaluation; Effective Teaching; Factor Analysis; *Higher Education; Multiple Regression Analysis; *Predictive Validity; Predictor Variables; *Rating Scales; Self Evaluation; Statistical Analysis; *Student Evaluation of Teacher Performance; Teacher Behavior; *Team Teaching; Test Validity
IDENTIFIERS Halo Effect

ABSTRACT

The study investigated the interrelationships of student ratings of teaching, course outcome and self ratings within a multiple instructor course mode, and the appropriateness of these variables for predicting overall instructor and course evaluations. Student rating of teaching and course related items and student self ratings were gleaned from a twenty item questionnaire. One hundred fifty-nine students in two courses responded to the questionnaire at the end of the spring 1975 semester. Students rated themselves, the four instructors in each course, and the course itself. Regression analyses indicated a very strong halo effect of instructor ratings over course ratings; student self ratings appeared not to bear any relationship to course ratings; and it was found that many of the variables that are often implicitly and explicitly criticized by students as having an effect on their educational experience (e.g. the quality of exam questions, course level, utility of textbooks, reasonableness of assignments) did not predict their rating of that experience very well. On the other hand, nearly every item that was concerned with teacher behavior accounted for a significant amount of variation in the instructors' overall evaluation. (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED129903

A Multivariate Study of Variables Effecting Student Ratings of Teaching and Course Outcomes Within A Multiple Instructor Mode

by

William M. McCook
University of Connecticut

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

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

TM005746

Paper presented as part of a symposium entitled "Progress and Problems in Validity Issues of Student Ratings," at the annual meeting of the National Council on Measurement in Education, San Francisco, April, 1976.

The research reported herein was supported in part by a Special Projects Grant from the Bureau of Health Manpower Education of the National Institutes of Health, U. S. Public Health Service.

A MULTIVARIATE STUDY OF VARIABLES EFFECTING STUDENT RATINGS OF TEACHING AND COURSE OUTCOMES WITHIN A MULTIPLE INSTRUCTOR MODE

William M. McCook
University of Connecticut

This study investigated the interrelationships of student ratings of teaching course outcome and self ratings within a multiple instructor course mode, and the appropriateness of these variables for predicting overall instructor and course evaluations. Student ratings of teaching and course related items and student self ratings were gleaned from a twenty item questionnaire. One hundred fifty-nine students in two courses responded to the questionnaire at the end of the spring 1975 semester. Students rated themselves, the four instructors in each course, and the course itself.

Regression analyses indicated a very strong halo effect of instructor ratings over course ratings; student self ratings appeared not to bear any relationship to course ratings; and it was found that many of the variables that are often implicitly and explicitly criticized by students as having an effect on their educational experience (e.g. the quality of exam questions, course level, utility of textbooks, reasonableness of assignments) did not predict their rating of that experience very well. On the other hand, nearly every item that was concerned with teacher behavior accounted for a significant amount of variation in the instructors' overall evaluation.

A MULTIVARIATE STUDY OF VARIABLES EFFECTING STUDENT RATINGS OF TEACHING AND COURSE OUTCOMES WITHIN A MULTIPLE INSTRUCTOR MODE

William M. McCook
University of Connecticut

There has been considerable effort in recent years to better understand, evaluate and improve the teaching-learning process at institutions of higher education. One aspect of this effort has been concerned with providing evaluative data on college teaching. Pressure to provide this data have come in part, from students who want information which will guide them in course and instructor selection, from teaching faculty who want information which will guide them in improving instruction, and from administrators who want information to guide them in promotion, tenure, and reappointment decisions.

Because of the desire of students to participate in the evaluation of courses, because of the view that students, as consumers of the educational service, are in the best position to evaluate its worth (Gessner, 1973, Rodin and Rodin, 1972, Rous et al, 1972,) and to a certain extent because it is the most convenient and accessable method of gathering information on instructional effectiveness (Rosenshine, 1970), student ratings of instruction are becoming more popular. The research on this topic (Costin, Greenough and Menges, 1971; Rosenshine, 1970), has suggested that while there are problems with the use of rating forms such as halo effects, homogeneity of items, and vaguely written items, there are also benefits of using rating scales to evaluate teachers. They are inexpensive means of collecting evaluative data because the students already enrolled in a course can serve as raters. Rating systems, unlike category observational systems require no specific training for raters. Rating scales have also been reported to be reliable measures of student impressions of teachers. Reliability estimates of rating scales of .80 to .90 have been reported (Villano and Rosenstock, 1973, Widlak et al., 1973). There is also evidence

that ege students can be objective consumers of the teaching process and their judgements should be solicited to identify variables important for teacher effectiveness (Grush and Costin, 1975).

While, indeed, there has been considerable research on the assessment of instructional effectiveness with college student ratings, nearly all of this effort has been directed at the evaluation of teaching in single instructor courses and toward issues related to the evaluation of teaching within the framework of the single instructor course. The generalizability of research findings gleaned from these studies to other teaching situations must be empirically verified. A general objective of this research was to begin to examine issues of validity related to student ratings of instruction within a multiple instructor mode with one rating scale. More specifically, the objectives of this research are threefold:

> 1) To examine the relationships of student ratings of specific rating items of teaching and course outcomes to the global items of overall instructor and course ratings;
> 2) To examine the relationships of student self-ratings to course rating items and the incremental validity of the self rating items to the prediction of overall course rating;
> 3) To assess the ability of student raters to evaluate teaching effectiveness and course outcomes on separate dimensions.

The first objective is concerned with the criterion-related validity of specific teacher and course ratings to predict global ratings. The second objective is concerned with the ability of students to rate courses objectively. The third objective is concerned with the presence of halo effect, or the ability of the raters to discriminate between traits (Hoyt, 1969; Widlak et al., 1973).

Method

Sample:

Data consisted of student responses on an instrument designed to evaluate teaching effectiveness and course characteristics. The students were completing their first year (third college year) in the school of

pharmacy at the University of Connecticut. The students (N = 159) rated the instructors and course characteristics in two lecture courses. The courses were required to be taken as part of the professional curriculum. Each course was worth five semester hours of credit and had four instructors. The amount of lecture time contributed by each instructor varied from 3 to 31 lecture hours with a mean of 16.92 hours and a standard deviation of 10.58 hours.

Data was gathered near the end of the Spring 1975 semester before final grades were assigned.

Instrument

The rating instrument is a variation of the University of Connecticut Rating Scale for Instruction (UCRSI, Bureau of Institutional Research, University of Connecticut, 1971). The difference being that the UCRSI was developed for single instructor lecture courses. The first eight items of the rating instrument used in this research is in fact the UCRSI. The validity of the UCRSI has been discussed by Brown (1974). In addition to the eight teacher evaluation items, the rating form contained ten items on course evaluation, two items on self evaluation, and items that could be used to identify sample characteristics. Three of the course evaluation items are relevant for laboratory courses, and were not used in this study.

Each teacher and course item on the rating scale requests the student to indicate his or her evaluation of each item on a 10 point scale (0-9), where zero would indicate the most negative response and nine would indicate the most positive response. For item 14 (Course level) a zero response would indicate that the student perceived the course as being very easy, a response of nine would indicate that the student perceived the course as being extremely difficult. Because the positive response for this item was the middle response, it became necessary to rescale scores on this item to reflect the positive-negative directionality of the other items. The eight teacher evaluation items are summarized as follows: (1) knowledge of the subject matter, (2) presentation of material, (3) balance of breadth and detail, (4) enthusiasm for subject, (5) fairness in marking, (6) attitude toward student, (7) personal mannerisms, (8) overall summary as a teacher.

The course evaluation items included the following: (1) overall educational experience in the course, (2) ratings on the value of the subject material of lectures, (3) ratings on usefulness of assignments, (4) ratings on the usefulness of textbooks, (5) ratings on the relevance of exam questions, (6) ratings on the difficulty level of the course, and (7) ratings on the effectiveness of team teaching. The self evaluation items were concerned with: (1) how much effort the student put into the course, and (2) how well the student felt he or she had done.

The rating scale can be characterized in terms of the evaluation plan of Smock and Crooks (1973). The two global items represent level I evaluations and give little or no specific information to guide instructional improvement. These items are primarily of use to administrators in making decisions on promotion, tenure and reappointment. The other seven teacher items, and the other six course evaluation items represent level II evaluations, that is, evaluations aimed at identifying success or failure in general areas or attitudes of instruction but not offering detailed evaluations aimed at providing diagnostic information about instructional problems. The self-evaluation items provide information about the student sample and possible relationships between their self assessment of their efforts and abilities and their evaluations of teaching and course characteristics. There are no level III items, that is, detailed, course specific evaluation aimed at providing diagnostic information about instructional problems (frequently those suggested by level II evaluation).

Data Analysis

Means, standard deviations, and product moment correlations were calculated for each of the seventeen items. Item 14 (Course level) was rescaled to reflect the same positive-negative directionality of the teacher and coures evaluation items. The rescaled item was scored on a scale of 0-4. Table 1 presents the means and standard deviations of these items. Table 2 presents correlation data for the items. The correlation of instructor lecture hours with the teacher evaluation items, and the correlation of the mean global teacher rating for each student with the course evaluation items is also presented in table 2.

Data were employed in stepwise multiple regression analyses to examine the relationship of teacher items to the overall teaching evaluation, the relationship of course items to overall course evaluation, the relationship of teacher evaluation to course evaluation. Results of these analysis are presented in tables 3 and 6.

Results of the analysis of teacher items to predict overall teacher evaluations are presented in table 3. The analysis also included number of lecture hours for each instructor as a predictor. The latter was included as a predictor because the variation in the number of lecture hours for each instructor could have been a factor in student ratings of instructors.

Regression analysis of course items on overall course evaluation are presented in table 4. Table 5 presents the same analysis with the self-evaluation items added as predictors. Table 6 presents the regression analysis of course items on the global course ratings with mean global teacher rating added as a predictor. This variable is the mean global teacher rating of the four instructors rated by each student and was included in the analysis to assess the presense of halo effect of teacher impression on course rating.

Finally, two principal component analyses of the seventeen item instrument were carried out. One analysis employed the individual instructor as the basic unit of analysis. The second analysis used the rater as the basic unit of analysis. Two analyses were necessary because while each student in each course evaluated four instructors, students rated themselves and course characteristics once in each course. The first analysis assumed that course and self ratings for each of the four instructors rated by each strudent would be equal. The second anlysis was done with the mean teacher item ratings of the four instructors rated by each student, the student's course ratings, and the student's self-evaluations. This analysis did not assume equal course and self-evaluations for each instructor, but does minimize between instructor variability. The results of the two analysis are presented in tables 7 and 8.

Results

Table 1 displays the item score means and standard deviations. Raw

score means range from 3.34 (exam questions) to 7.19 (knowledge of subject matter). A comparison of item means for the global items of instructor and course evaluation indicate that the student raters tended to rate instructors in a more positive direction than courses. The difference between these item means is significant ($t(312)=6.01, p < 0.001$). This finding is consistent with that of Widlak (1973) which reported that instructors tend to be rated higher than courses for this kind of scale.

Inspection of the intercorrelation matrix presented in table 2 reveals high correlations among the teacher evaluation items except for item 1 (knowledge of subject matter) with item 2 (presentation of material) and item 3 (balance of breadth and detail). The criterion of overall evaluation of teaching (item 8) is highly correlated with each of the specific teacher evaluation items.

An examination of course evaluation and self-evaluation intercorrelations also present a considerable number of highly correlated items. The self-evaluation items appear not to be highly correlated with the course evaluation items with the exception that item 17 (how well the student felt he had done) and item 14 (course level) appear to be moderately correlated. The positive correlation between these two items may reflect the students perceived ability to cope with the course.

Course items appear to be correlated with mean teacher ratings suggesting a possible halo effect of teacher rating over course rating. The high correlation between the global course evaluation and mean global teacher rating could indicate a high degree of perceived relationship between these concepts. The low relationship between the self-evaluations and the course and the mean teacher evaluation indicates that the students appear to make these self-evaluations independently of their perceptions of their teachers and courses.

The results of the regression analysis presented in tables 3 to 6 present evidence on the criterion related validity of the specific items to predict the global course item, and of the effect of self-evaluation items on the global course item.

Table 3 presents the results of a stepwise multiple regression of the specific teacher items on the global teacher item. Number of lecture hours each

instructor contributed was also included as a predictor. Examination of the table indicates that each of the specific teacher items accounted for significant variance in the criterion. Thus, eventhough the seven specific teacher items are highly correlated with each other and the ctiterion, each of these items individually contributes to the validity of the global rating. The predictor "number of lecture hours" entered the regression equation last, indicating that this predictor accounted for little new variation in the criterion. The b-weight for this predictor was significantly different from zero ($p < 0.05$), indicating a small but statistically significant amount of incremental validity for this predictor. However, the amount of new variance accounted for by this predictor is so small, it is difficult to give it any practical significance.

Table 4 presents the results of the regression analysis of the specific course items on the global course item. Only three of the six specific course items contributed unique significant variation in predicting overall course rating. Item 9 (subject material of lecture), item 15 (effectiveness of team teaching), and item 5 (exam questions) accounted for 58 percent of the variance in the criterion.

Table 5 presents the results of a stepwise multiple regression analysis of specific course items and self-evaluation items to predict the global course rating. The results indicate that only the same three specific course items used in the previous analysis account for significant variation in the criterion. The incremental validity of the self evaluation items appears not to be significant.

Table 6 presents the stepwise multiple regression of specific course items and mean teacher rating to predict overall course rating. The rationale for including mean teacher rating as a predictor was to allow for the investigation of possible halo effect of teacher rating on course rating. Previous research (Widlak et al., 1973) has indicated that students do no rate courses and instructors separately. An examination of table 6 indicates not only that mean teacher rating is a good predictor of global course rating but it is the second best predictor in the set on predictors. When the incremental validity of mean teacher rating (McNemar, 1962) is assessed by comparing the

equation in table 6 to the equation in table 5, the result is a significant increase in incremental validity ($F=10.58, df=1,133$). This finding is indicative of the halo effect discussed by Widlak et. al. (1973).

Table 7 and 8 present the results of principal component analyses of the seventeen item instrument. Table 7 presents an analysis of the instrument using the instructor as the basic unit of analysis. Table 8 uses the rater as the basic unit of anlaysis. Each analysis represents the result of a principal components solution using an oblique (obliquimax) transformation (see Hofman, 1970). The two solutions are discussed below.

Inspection of table 7 indicates that each of the 17 items loaded highly on only one of the three derived components. The first component contained the eight teacher evaluation items. The second component contained six of the seven course evaluation items. The third component contained three items with high loadings, the two self evaluation items and the item on course level. The first component accounted for 37 percent of the total variance, the second component accounted for 15 percent of the total variance, and the third accounted for 9 percent of the total variance. Intercorrelations of the primary axes indicated that these components do not represent independent dimensions of evaluation. Component I was correlated with component II ($r = .43$). Component III was not correlated with component I ($r = 0.05$) and was slightly correlated with component II ($r = -0.23$).

The results of the second principal components analysis yielded a nearly identical component structure to the first analysis. Component I consists of the teacher evaluation items, component II consists primarily of the course evaluation items, and component III consists of the self-evaluation items and the course level items. In this analysis item 13 did not load highly on any of the three components. Also, item 16 loads highly on components II and III. Component I accounted for 43 percent of the total variance, component II accounted for 14 percent, and component II accounted for 8 percent.

In both analyses the negative loading of item 16 on component II is indicative of the fact that as the level of difficulty of the course increases (a low rating) the student perceptions of their effort increases (a high rating). The fact that item 14 (course level) loaded highly on the third

component rather than on the second could be because the item is more indicative of the students perception of their ability to deal with the level of difficulty of the course rather than their perception of relative difficulty of the course in terms of some ideal level.

The intercorrelation of the primary axes for this solution indicated that components I and II are moderately correlated (r= -0.58) components I and III (r=0.01) are not correlated and components II and III (r=0.24) were only slightly correlated.

In summary, the principal components analysis indicated that students rate instructors and courses on separate but not independent dimensions. They appear to evaluate themselves on a separate and independent dimension. The correlation between the teacher dimension (component I) and the course dimension (component II) gives further support to the concept of halo effect. The independence of the "self-evaluation" dimension is consistent with the regression analysis findings that the self-evaluation items do not account for significant variation in the criterion of overall course rating.

Discussion

The present results in many ways parallel the findings of Widlak et al (1973) that were based on data gathered in single instructor courses. As in the Widlak study, the global rating of instructor was found to be higher than the overall rating for courses. However, the strong relationship exhibited between these two variables discounts any general statement concerning the meaning of this difference.

While the instructor rating items are highly intercorrelated with each other and also highly correlated with the global rating of instructor, it is not clear that the homogenity of these items represents a halo effect. Though the factor analysis indicates that the instructor rating items are rated along one dimension, the regression analysis of these items indicates that each item accounts for unique variance in global instructor ratings. Anderson and Jacobson (1965) present evidence that person perception is the result of averaging of stimulus components. An interpretation along these lines may be more satisfactory than that of a halo effect.

While the level II instructor items were good predictors of the level I teacher ratings, the level II course items were less effective predictors of the level I course items. The lower predictive efficiency of the course items may be due to an inadequate sampling of course characteristics, the vagueness of both level I and level II course items, or the result of the halo effect of teacher impression on course evaluation. It is clear from the regression analyses and the principal components analyses that the students did not rate teachers and courses independently. With several instructors in a single course and the presence of halo effect of instructor impression over course rating, the meaningfulness of the course rating becomes blurred.

There is little doubt that student ratings can yield reliable, valid, and useful information about the quality of courses and instruction (Costin, Greenough, and Menges, 1971), however the findings of this study suggest that the use of level I and level II items as defined by Smock and Crooks (1973) for course evaluations in multiple instructor courses are of questionable value. The finding of this study also suggest that level II instructor evaluation items appear to have some validity for predicting level I instructor evaluations in the multiple instructor setting.

References

Anderson, N.H., & Jacobson A. Effect of stimulus inconsistency and discounting instructions in personality impression formation. Journal of Personality and Social Psychology, 1965, 2, 531-539.

Brown, D.L. Faculty ratings and student grades: A large-scale multivariate analysis by course sections. Unpublished doctoral dissertation, University of Connecticut, 1974.

Bureau of Institutional Research, University of Connecticut, UCRSI- University of Connecticut Rating Scale for Instruction, 1971.

Costin, F., Greenough, W.T., & Menges, R.J. Student ratings of college teaching: reliability, validity, and usefulness. Review of Educational Research, 1971, 41, 511-535.

Gessner, P.K. Evaluation of instruction. Science, 1973, 180, 566-570.

Grush, J.E. & Costin, F. The student as consumer of the teaching process. American Educational Research Journal, 1975, 12, 55-66.

Hofman, R.J. The obliquimax transformation. Unpublished doctoral dissertation, State University of New York at Albany, 1970.

Hoyt, D.P. Instructional effectiveness: 1 Measurement of effectiveness. Kansas State University Office of Educational Research Report #6, Nov., 1969.

McNemar, Q. Psychological Statistics (3rd ed.), N.Y.: Wiley, 1962 Chapter 11

Rodin, M. & Rodin, B. Student evaluations of teachers, Science, 1972, 177, 1164-1166.

Rosenshine, B. Evaluation of Classroom instruction, Review of Educational Research, 1970, 40, 279-300.

Rous, S.N., Bamford, J.C. Gromisch, D., Rubin, S., & Sall, S., The improvement of faculty teaching through evaluation: a preliminary report Journal of Surg. Research, 1971, 11, 311.

Rous, S.N., Bamford, J.C., Gromisch, D., Rich, H., Rubin, S., & Sall, S. The improvement of faculty teaching through evaluation: a follow-up report. Journal of Surg. Research, 1972, 13, 262.

Smock, H.R., & Crooks, T.J. A plan for the comprehensive evaluation of college teaching. Journal of Higher Education, 1973, 44, 577-586.

References (Continued)

Villano, M.W., & Rosenstock, E.H. A decade with the course attitude questionaire: A factorial study. A paper presented at the annual convocation of the Northeastern Educational Research Association, Ellenville, N.Y., 1973.

Widlak, F.W., McDaniel, E.D., & Feldhusen, J.F., Factor analysis of an instructor rating scale. A paper presented at the annual meeting of the American Educational Research Association, New Orleans, La., 1973.

TABLE 1

MEANS AND STANDARD DEVIATIONS OF RATING SCALE
N=159

Item	Mean	std. dev.
(Teacher evaluation)		
1. Apparent Knowledge of Subject matter	7.19	1.68
2. Presentation of material	5.62	1.68
3. Balance of breadth and detail	5.71	2.11
4. Enthusiasm for subject	6.28	1.90
5. Fairness in marking	6.10	2.02
6. Attitude toward student	6.07	2.22
7. Personal mannerisms	6.73	1.96
8. Overall summary as teacher	6.16	1.87
(Course Evaluation)		
9. Overall educational experience in this course	4.80	2.05
10. Subject material of lectures	5.15	1.85
11. Assignments	4.50	2.10
12. Textbooks	4.04	2.61
13. Exam Questions	3.35	2.23
14. Course level	6.01	1.42
(Rescaled)	2.69	1.12
11. Effectiveness of Team Teaching	5.22	1.95
(Self-Evaluation)		
16. How much effort did you put into the course?	7.09	1.39
17. How well do you feel you have done?	4.62	1.91

TABLE 2

INTERCORRELATION MATRIX OF ITEM SCORES

Item	1	2	3	4	5	6	7	8	TXP*	9	10	11	12	13	14	15	16	17
1.		28	33	48	39	37	36	46	08									
2.			78	51	56	64	59	77	-49									
3.				47	55	57	59	73	-41									
4.					51	60	48	66	-10									
5.						69	52	68	-24									
6.							61	77	-29									
7.								68	-32									
8.									-30									
Avg. 8										58	44	46	40	53	07	57	14	16
9.											73	42	38	50	23	51	21	27
10.												49	51	44	15	40	26	23
11.													53	47	11	40	-02	11
12.														43	11	37	03	09
13.															26	30	-08	32
14.																05	-21	46
15.																	05	19
16.																		-11
17.																		

*TXP - Instructor Lecture Hours
Correlations greater than 0.16 are significantly different from zero ($p<0.05$)

TABLE 3

STEPWISE MULTIPLE REGRESSION OF TEACHER ITEMS ON OVERALL TEACHER RATING

Step	Variable	Final Step			Cumulative Step		
		bwt	SE	F-Value	Mult-R	SE	F-Value
1.	Presentation of material	0.1912	0.0293	43.3513**	0.7758	1.1804	854.4407**
2.	Attitude toward student	0.1864	0.0255	53.4261**	0.8492	0.9887	729.5667**
3.	Enthusiasm for subject	0.1415	0.0256	30.6719**	0.8683	0.9296	575.3364**
4.	Personal Mannerisms	0.1528	0.0253	36.4614**	0.8821	0.8835	492.9303**
5.	Balance of Breadth of material	0.1659	0.0281	34.8572**	0.8907	0.8533	431.0600**
6.	Fairness in Marking	0.1240	0.0251	24.3307**	0.8963	0.8333	381.3958**
7.	Apparent Knowledge of subject matter	0.0808	0.0249	10.5753**	0.8991	0.8232	337.1118**
8.	Number of Lecture hours	0.0083	0.0040	4.3489*	0.8999	0.8207	297.2836**

*p < 0.05
**p < 0.01

intercept= -0.3886
N = 566

TABLE 4

STEPWISE REGRESSION OF COURSE VALIABLES
ON OVERALL COURSE RATING

Step	Variable	Final Step			Cumulative Step		
		bwt	SE	F-Value	Mult-R	SE	F-Value
1.	Subject material of lectures	0.6071	0.07554	64.5882**	.7109	1.4483	145.1547**
2.	Effectiveness of team teaching	0.2440	0.0648	14.1817**	.7462	1.3758	88.6003**
3.	Exam Questions	0.1505	0.0615	5.9821*	.7626	1.3417	64.8656**
4.	Course level	0.1977	0.1055	3.5140	.7693	1.3298	50.4028**
5.	Textbooks	-0.0338	0.0536	0.3969	.7703	1.3320	40.2951**
6.	Assignments	-0.0173	0.0685	0.0635	.7705	1.3366	33.3620**

*p < 0.05
**p < 0.01

intercept = -0.4301
N = 143

TABLE 5

STEPWISE MULTIPLE REGRESSION OF COURSE AND SELF RATINGS ON OVERALL COURSE RATINGS

Step	Variable	Final Step			Cumulative Step		
		bwt	SE	F-Value	Mult-R	SE	F-Value
1.	Subject Material of lectures	0.5773	0.0815	50.2343**	.7246	1.4320	153.6831**
2.	Effectiveness of team teaching	0.2290	0.0656	12.1783**	.7567	1.3634	92.4325**
3.	Exam Questions	0.1644	0.0624	6.9369**	.7739	1.3277	67.8332**
4.	Course level	0.2035	0.1177	2.9876	.7778	1.3202	52.0909**
5.	How much effort put into course	0.1633	0.0872	3.5043	.7845	1.3075	43.2106**
6.	Textbooks	-0.0273	0.0533	0.2620	.7851	1.3108	35.6847**
7.	How well do you feel you have done?	0.0256	0.0671	0.1453	.7853	1.3150	30.5834**
8.	Assignments	0.0008	0.0691	0.0001	.7853	1.3200	26.5593**

*p < 0.05
**p < 0.01

intercept = -1.6678
N = 143

TABLE 6

STEPWISE MULTIPLE REGRESSION OF COURSE ITEMS AND AVERAGE TEACHER RATING ON OVERALL COURSE RATING

Step Variable	Final Step			Cumulative Step		
	bwt	SE	F-Value	Mult-R	SE	F-Value
1. Subject material of lectures	0.5757	0.0787	53.5632**	.7249	1.4229	156.2103**
2. Average teacher rating	0.3836	0.1257	9.3104**	.7788	1.3004	167.9052**
3. Course Level	0.1340	0.0864	2.4070	.7856	1.2871	74.7418**
4. Effectiveness of team teaching	0.1377	0.0695	3.9302*	.7912	1.2768	57.7776**
5. Exam Questions	0.0937	0.650	2.0762	.7937	1.2746	46.6708**
6. How much effort	0.1090	0.0852	1.6595	.7968	1.2707	39.4374**
7. Textbooks	-0.0332	0.0516	0.4134	.7979	1.2723	33.8135**
8. How well you feel you have done	0.0428	0.0643	0.4429	.7988	1.2748	29.5323**
9. Assignments	-0.0186	0.0673	0.0763	.7989	1.2792	26.0785**

*p < 0.05
**p < 0.01

Intercept = -2.7242
N = 143

Table 7

PRIMARY PATTERN MATRIX FOR DERIVED
PRINCIPAL COMPONENT SOLUTION USING OBLIQUIMAX TRANSFORMATION*

	COMPONENT		
Item	1	2	3
1.	53		
2.	84		
3.	80		
4.	75		
5.	76		
6.	92		
7.	77		
8.	91		
9.		81	
10.		88	
11.		75	
12.		76	
13.		52	
14.			72
15.		66	
16.			-59
17.			73

* All loadings have been multiplied by 100
All loadings less than .40 have been deleted

TABLE 8

PRIMARY PATTERN MATRIK FOR DERIVED
PRINCIPAL COMPONENT SOLUTION USING OBLIQUIMAX TRANSFORMATION*
(based on mean teacher ratings per student)

	COMPONENT		
Item	1	2	3
1.	70		
2.	81		
3.	73		
4.	92		
5.	86		
6.	100		
7.	61		
8.	85		
9.		78	
10.		92	
11.		71	
12.		77	
13.			
14.			76
15.		74	
16.		44	-59
17.			74

*All loadings have been multiplied by 100
All loadings less than .40 have been deleted