

國立臺灣師範大學教育心理與輔導學系
教育心理學報，民92，35卷，2期，183—200頁

Results of student ratings: Does faculty attitude matter? *

TE-SHENG CHANG

Graduate Institute of Compulsory Education
National Hualien Teachers College

The purpose of this study was to compare the student ratings of faculty with positive attitude and negative attitude towards student ratings. The correlation between faculty attitude towards student ratings and their ratings scores was investigated. The sample consisted of 70 faculty members at National Hualien Teachers College in Taiwan, during the 1999 spring semester. Two instruments, the Attitude towards Student Ratings of Instruction Questionnaire (ASRI) and the Student Ratings of Instruction form (SRI), were administered. The ASRI was used to measure faculty attitude towards student ratings, The SRI was used measure students' perceptions of these teachers' appeal and their teaching effectiveness.

Results indicate that (1) The overall-agreement group was rated statistically higher than was the overall-disagreement group on Preparation/Planning, Material/Content, Assignment/Examination, and the total rating scores; (2) Faculty members with positive attitude towards the publication of student ratings were rated higher than those with negative attitude regarding to Method/Skill, and the total rating scores; (3) The correlation between faculty attitude towards overall agreement on student ratings and their ratings scores was statistically significant.

KEY WORDS: student ratings of instruction, faculty evaluation, attitude

Student ratings have been officially employed in institutions of higher education since 1927. Wagenaar (1995) stated that well over 90 percent of schools currently use student ratings for assessing the teaching staff. Evidence from many other studies also indicates that most universities in North America use student ratings of instruction as part of their evaluation of teaching effectiveness (d'Apollonia & Abrami, 1997). Seldin (1993) found that student ratings have been used as a component of faculty evaluation systems in more than 86% of the 600 humanities colleges surveyed. Calderon, Gabbin, and Green (1995) found that close to 95 % of accounting departments use student ratings of their instruction and as many as 18% rely exclusively on those student ratings in evaluating faculty teaching performance. As Wilson (1998) predicted, all of the

* This research was supported by research grants (NSC 89-2413-H-026-004-S) and (NSC 89-2413-H-026-021-SSS) from National Science Council in Taiwan (國科會) .

colleges would consider student ratings of instruction as a measure of “teaching quality” for their faculty.

Student ratings of instruction are commonly used to provide: (1) formative feedback to faculty for improving teaching, course content and structure; (2) a summary measure of teaching effectiveness for promotion and tenure decisions; and (3) information to students for the selection of courses and teachers (Marsh & Roche, 1993). Why are student ratings commonly used for faculty evaluations? Some realistic reasons for this are as follows: (1) Raters from the student body have closely and recently observed a number of teachers; (2) Students’ frank reactions can be a beneficial aid to the faculty member in refining his/her course structure and teaching styles; (3) Student ratings are seen to be more objective than other approaches such as administrator evaluations, peer evaluation, self-rating, and classroom visitation evaluation; (4) Student ratings are unique in capacity to indicate how students think and feel; (5) Students are a convenient source of rating (Arreola, 1995; Peterson, 1995).

The advent of student consumerism and accountability issues in the last two decades, as some researchers offer to explain, has propelled research in student ratings to new heights (Benson & Lewis, 1994; Van Patten, 1994). Research on student ratings of instruction has often examined issues such as the development and validity of an evaluation instrument (e.g., Marsh, 1987), the validity (e.g., Cohen, 1981), and reliability (e.g., Feldman, 1977) of student ratings in measuring teaching effectiveness, and the potential bias of student ratings (e.g., Centra & Gaubatz, 2000; Chang, 2000; Feldman, 1993). With considerable consistency, research brings to the fore elements of instructor behavior which student’s rate to be exceedingly important. With varying labels and weights depending upon the individual studies, these elements include the factors of learning value, instructor enthusiasm, organization and clarity of explanation, individual rapport, group interaction, ability to stimulate student interest and learning, breadth of coverage of material, and overall fairness in examinations, grading, assignments, and workload difficulty.

Student ratings of instruction have been the subject of numerous studies with much of the research focusing on the validity, reliability, relationship to other variables, and potential biasing factors (e.g., Centra & Gaubatz, 2000; Chang, 2000; Feldman, 1993; Marsh, 1984). Comparatively little empirical investigation has been devoted to the perspectives of the individuals who are the rated, that is faculty. Without faculty cooperation and support, student ratings are useless. From 1927, a lot of research studies have investigated the faculty attitude towards student ratings, especially the use of student ratings. But, why are faculty attitude towards student ratings so important? Most of these “faculty attitude” studies have assumed that faculty attitude towards student ratings have an influence on the results of student ratings scores. Will faculty members with positive attitude towards the policy of student ratings receive higher student ratings scores than those with negative attitude, and vice versa? Is there any statistically correlation between faculty attitude towards the use of student ratings and their student ratings scores? Or, is there any statistically difference in student ratings scores among faculty with different attitude towards the use of student ratings?

Literature Review

Faculty Attitude towards Student Ratings

In a study by Ryan, Anderson, and Birchler (1980), it was found that the introduction of mandatory use of student evaluations led to a significant reduction in faculty morale and job satisfaction. It also motivated faculty to reduce standards and workloads for their students and to make examinations easier. A similar study, in which the use of student ratings was not compulsory, found that faculty members were generally satisfied with the process (Baxter, 1991). Ryan and his colleagues also believed that the imposition of mandatory student ratings might contribute to grade inflation. This is what Marsh and Dunkin (1992) called a leniency hypothesis: instructors with more lenient grading standards receive more favorable ratings. Even if this hypothesis is incorrect, faculty may nonetheless believe it to be accurate (Centra, 1993) and reduce standards anyway.

By contrast, Jacobs (1987) found that a majority of faculty responding to her survey did not feel that student ratings have a negative effect on faculty morale, and she believed that the ratings should be required. It is interesting to note, however, that her survey was conducted at a more highly research-oriented university than that of Ryan and his colleagues' study. Rich (1976) found that in research-oriented institutions, where teaching holds a position of lesser importance, faculty were more favorably disposed towards the use of student ratings for summative purposes. Avi-Itzhak and Kremer (1986) found that senior and tenured faculty are most opposed to the use of student ratings for summative purposes, while non-tenured faculty, both tenure-track and non-tenure-track, were most supportive of this use of student ratings. They believed that this is due to the fact that senior faculty members devote most of their time to research and less time to their teaching, and are less student oriented than junior faculty. Spencer and Flyr (1992) found that only 23% of faculty responding to their survey reported making changes to their teaching based on student evaluation results, and when such changes did occur, they usually took the form of altering handouts, presentation habits, and assignments. According to Franklin and Theall's study (1989), faculty members who familiarize themselves with existing research on student ratings of instruction tend to have more positive attitude about their use.

From the above research studies, it seems that most of faculty members agree with that the implement of student ratings have some influences on teacher teaching behavior in the class. Some research studies (e.g., Armstrong, 1998) indicated there is no evidence that these changes are likely to contribute to learning. However, some researchers (e.g., Jacobs ,1987) believed that student ratings should be a requirement for school, especially for the research-oriented schools.

Teacher Variables Thought to Influence Student Ratings

A wealth of research exists in the area of student ratings, ranging from analyses of validity and reliability to studies parceling effects related to course, student, and teacher characteristics. This section provides a review of the findings on teacher characteristics that can conceivably exert an influence on student ratings scores.

Research typically indicated a positive effect of teacher rank on student ratings but a negative effect for age of the faculty member and years of teaching on ratings (Feldman, 1983). Feldman noted that while higher faculty rank is typically associated with higher overall ratings, the relationship can disappear or reverse when

particular dimensions of teaching are examined. Discussion of the effect of teacher gender on student ratings appeared to be quite varied. In a two-part meta-analysis, Feldman (1992, 1993) reviewed existing research on student ratings of male and female teachers in both the laboratory and the classroom setting. In his review of laboratory studies, Feldman (1992) reported that the majority of studies reviewed showed no difference in the global evaluations of male and female teachers. In his review of classroom studies, Feldman (1993) again reported that the majority of studies reported no significant differences between the genders.

Grading standard perhaps generates the most suspicion about the validity of student ratings. Bridgeman (1986) and Owie(1985) compared summary evaluation scores of three student groups, those receiving grades worse than expected, same as expected, and better than expected. Both of them found significant differences among the groups. The lowest evaluations came from the negative discrepancy group; the highest came from the zero discrepancy group for Bridgeman and the positive discrepancy group for Owie. Greenwald and Gillmore (1998) found that teachers can raise their ratings substantially by grading more leniently. They believed lenient grading leads to increased student ratings and is easily performed. On the other hand, Roche and Marsh (1998) disagreed that lenient grading inflates student ratings.

It must be noted here that the mere existence of a correlation between a background variable and rating scores does not necessarily constitute a bias or a threat to the validity of student ratings (Brandenburg, Slinde, & Batista, 1977; Tatro, 1995). For example, if the student expected grade in the course is found to be associated with the rating that student gives to the instructor, it does not necessarily follow that an instructor can obtain higher ratings merely by giving higher grades. Alternative explanations include the possibility that more effective teaching will inspire students to work harder and earn better grades (Watchel, 1998).

Some researchers have tried to find factors, which may be contributed to the ratings scores, including faculty characteristics and their grading systems. On the other hand, some researchers have investigated faculty attitude towards student ratings and the effect of their background variables on the attitude difference. They emphasized that faculty cooperative attitude is an import element to the success of student ratings. However, none research examines the correlation between faculty attitude towards student ratings and their rating scores. Why are faculty attitude towards student ratings so important? Most of these “faculty attitude” studies may assume that faculty attitude towards student ratings have an influence on the results of student ratings scores. Will faculty members with positive attitude towards student ratings receive higher student ratings scores than those with negative attitude, or vice versa?

Issues for Student Ratings

To investigate the effects of faculty attitude towards student ratings on their ratings scores, this study is focused on the two important issues related to student ratings. The first issue is the overall agreement on the student ratings policy. It means faculty members agree or disagree with the school, where they serve, to take student ratings as a requirement for all faculty members. The second issue is the application of student ratings. Some important issues of student ratings have been discussed by various research studies, such as the purposes of student ratings (e.g., Spencer, 1994), the concerns about ratings (e.g., Abbott, et. al, 1990), the components of teaching effectiveness, the negative effects of student ratings (e.g., Wachtel, 1998), the application of student ratings (e.g. O, 1996), and the procedures of student ratings (e.g., Centra, 1981). However, the application of student ratings is the most important issue in the field than other issues because it influences faculty personnel decision, such as promotion, tenure, and awards. Besides, the majority of faculty

members have a very high agreement on the purposes of student ratings, the concerns about ratings, the components of teaching effectiveness, the negative effects of student ratings, and the procedures of student ratings but the application of student ratings (Chang, 2002).

According to Chang (2002), there are six different applications of student ratings. They are (1) a source of feedback to individual instructor, (2) input towards the teacher and teaching awards, (3) input for the administrative evaluation of job contracts, (4) input for the administrative evaluation of promotions, (5) assistance for students in selecting courses and instructors, and (6) publication in school journals. This study takes these six different applications and the overall agreement on the student ratings as independent variables and the resultant ratings scores as the dependent variables. It is assumed that faculty with positive attitude towards student ratings receive higher ratings scores than do faculty with negative attitude.

Method

The Measurements

Two instruments, the Attitude towards Student Ratings of Instruction Questionnaire (ASRI) and the Student Ratings of Instruction form (SRI), were employed in the study. ASRI was used to measure faculty attitude towards student ratings. It had 33 items rated on a 5-point Likert scale ranging from “strongly agree” (5-point) to “strongly disagree” (1-point). It included five issues (Items 1 - 32) and an overall item (Item 33). The five issues were: the purposes of student ratings (Items 1 - 7), the concerns about student ratings (Items 8 - 11), the components of student ratings (Items 12 - 20), the negative effects of student ratings (Items 21 - 26), and the application of student ratings (Items 27 - 32). Item 33 asked for an overall agreement on student ratings. In addition to the 33 items, faculty also an opportunity to provide anonymous, written comments about student ratings on the ASRI survey. A panel of 15 evaluation experts for content validity validated all items. Due to the purposes of this study, only the application of student ratings (Items 27 - 32) and Item 33 (an overall agreement on student ratings) are used as independent variables in the study.

SRI was used to measure students’ perceptions of these teacher appeal and teaching effectiveness during the last two weeks of classes. The rating form was composed of 13 questions rated on a 5-point Likert scale ranging from “strongly agree” (5-point) to “strongly disagree” (1-point). It yielded four factors: Preparation/Planning (Item 1, 2, 3), Material/Content (Item 4, 5, 6), Method/Skill (Item 7, 8, 9, 10), and Assignment /Examination (Item 11, 12, 13). The sum of these four factors was considered as a total rating score for a faculty member.

Table 7 in Appendix A presents the correlation between item and subscale and internal consistency for four subscales of SRI. The (coefficients of internal consistency reliability were .857, .917, .933, and .927 for Preparation/Planning, Material/Content, Method/Skill, and Assignment/Examination, respectively. These coefficients confirmed that the questionnaire was a reliable instrument.

Principal components analysis was applied for each factor and the whole instrument separately. Factor loadings for items designed to measure each factor were consistently large, between .885 and .944. Each Principal components analysis had only one eigenvalue greater than one, which indicated the items were pure indicators for their own factors. The four factors accounted for 87% of the total variance. Table 8 in Appendix A shows the result of principal components analysis for each subscale and total scale of SRI.

The SRI also posed an open-ended question to elicit students’ narrative reactions and their specific

suggestions for achieving a better class: “What do you think could be done to improve this class?” Due to the limitation of the paper, the result of this item will not be discussed in this paper.

Data Source

The data for this investigation came from an aggregation of SRI and ASRI. The ASRI questionnaire, with a cover letter explaining its purpose, was mailed to a random sample of 120 faculty members from National Hualien Teachers College in Taiwan between April and June of 1999. An overall response rate of 65.0% ($N = 78$) was achieved. Due to missing data, the rate of usable questionnaires was further reduced to 58.3% ($N = 70$) with 49 male (70.0%) and 21 female (30.0%) faculty members. These 70 faculty members consisted of 9 (12.9%) full professors, 31 (44.3%) associate professor, and 30 (42.9%) lecturers.

The aggregation of SRI data adopted from the same school in the end of the 1999 spring semester was matched with ASRI data based on faculty identification number. Thus, the data set compromised the resultant ratings scores plus attitudinal data for each subject.

Analytic Strategy

The Statistical Package for the Social Science (SPSS) program was utilized to tabulate data and compute the statistical tests. The sample was divided into two groups for each independent variable based on their ASRI scores of the application of student ratings (Items 27-32) and the overall agreement (Item 33). For example, faculty members who showed the overall attitude score higher than or equal to 4 were considered as an overall-agreement group. On the other hand, faculty members with the overall attitude score equal to or below 3 were regarded as the overall-disagreement group. The same procedure was applied to the six different applications of student ratings. Therefore, there were seven independent variables (categories), and each variable had two different groups.

The independent t test method was applied for testing the mean difference in student ratings between the overall-agreement group and the overall-disagreement group. Similarly, the independent t test method was applied for testing the mean difference in the results of student ratings between faculty with positive attitude and negative attitude towards the application of student ratings as a source of teaching improvement, teaching awards, job contract, promotion decision, course selection, school journal publication, respectively.

Each of the seven independent variables was correlated with each of the five student rating scores, including Preparation/Planning, Material/Content, Method/Skill, Assignment/Examination, Overall.

Results

Ratings Scores of Preparation/Planning

Table 1 lays out the summary of independent t tests and correlation tests for ratings scores of Preparation/Planning by faculty with positive attitude and negative attitude towards the use of student ratings. Among six different uses of student ratings, fifty-eight (83%) faculty members agree that student ratings scores can be used as a source of feedback to individual instructor. There are 39 (56%) faculty members agree that student ratings score can be considered as an input towards the teacher and teaching awards. Only 15 of 70 faculty members agree that the result of student ratings can be published in the school journals.

The overall-agreement group was rated statistically higher than was the overall-disagreement group on

Preparation/Planning. The correlation between the overall agreement on student ratings and the ratings scores of Preparation/Planning is also statistically significant but the agreement on the use of student ratings is not. It seems that the attitude towards the use of student ratings does not correlate with the ratings scores. However, faculty attitude towards the policy of student ratings correlates with the rating scores of Preparation/Planning.

Ratings Scores of Material/Content

Table 2 shows the summary of independent *t* tests and correlation tests for ratings scores of Material/Content by faculty with positive attitude and negative attitude towards the use of student ratings. Like the finding on Preparation/Planning, the overall-agreement group was rated statistically higher than was the overall-disagreement group on Material/Content. The correlation between faculty overall agreement on student ratings and their ratings scores on Material/Content is .325 which is statistically significant. Faculty attitude towards the use of student ratings do not correlate with the ratings scores. Faculty attitude towards the policy of student ratings correlates with the rating scores of Material/Content but their attitude towards the use of student ratings does not.

Ratings Scores of Method/Skill

Table 3 summarizes the independent *t* tests and correlation tests for ratings scores of Method/Skill by faculty with positive attitude and negative attitude towards the use of student ratings. As for the rating scores of Method/Skill, faculty members with positive attitude towards the use of student ratings as a publication in school journals receive higher ratings scores than those with negative attitude. However, the rest of two-different groups in other application of student ratings receive the non-significant ratings scores. The correlation between faculty attitude towards student ratings and their ratings scores of Method/Skill is not statistically significant.

Ratings Scores of Assignment/Examination

The summary of independent *t* tests and correlation tests for ratings scores of Assignment/Examination by faculty with positive attitude and negative attitude towards the use of student ratings is showed in Table 4. The overall-agreement group nets significantly higher ratings scores on Assignment/Examination than does the overall-disagreement group. The correlation between the overall agreement on student ratings and the ratings scores of preparation/planning is also statistically significant. However, faculty attitude towards the use of student ratings does not correlate with the ratings scores of Assignment/Examination.

The Total Ratings Scores

Table 5 indicates the summary of independent *t* tests and correlation tests for ratings scores of the total ratings scores by faculty with positive attitude and negative attitude towards the use of student ratings. As for the total ratings scores, faculty members with positive attitude towards the use of student ratings as a publication in school journals receive significantly higher scores than do faculty with negative attitude. Again, the overall-agreement group is rated statistically higher than is the overall-disagreement group on the total ratings scores. The correlation between the overall agreement on student ratings and the total ratings scores is statistically significant. And this is only the significant correlation in Table 5.

Table 1 Summary of Independent *t* Test and correlation test for ratings scores of Preparation/ Planning by faculty with positive attitude and negative attitude towards the used of student ratings

	<i>N</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>r</i>
A source of feedback to individual instructor					
Agree group	58(83%)	4.026	.244	.048	.024
Disagree group	12(17%)	4.023	.328		
Input towards the teacher and teaching awards					
Agree group	39(56%)	4.016	.296	-.307	.016
Disagree group	31(44%)	4.034	.204		
Input for the administrative evaluation of job contracts					
Agree group	31(44%)	4.011	.311	-.390	-.026
Disagree group	39(56%)	4.036	.210		
Input for the administrative evaluation of promotions					
Agree group	32(46%)	4.060	.301	.171	.01
Disagree group	38(54%)	4.027	.233		
Assistance for students in selecting courses and instructors					
Agree group	35(50%)	4.034	.278	.279	.027
Disagree group	35(50%)	4.016	.240		
Publication in school journals					
Agree group	15(21%)	4.119	.256	1.603	.178
Disagree group	55(79%)	4.000	.255		
Overall agreement on student ratings					
Agree group	57(81%)	4.058	.245	2.325*	.369**
Disagree group	13(19%)	3.880	.271		

Note: * $p < .05$; ** $p < .01$; *t* means independent *t* test; *r* means Pearson correlation.

Table 2 Summary of Independent *t* Test and correlation test for ratings scores of Material/Content by faculty with positive attitude and negative attitude towards the used of student ratings

	<i>N</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>r</i>
A source of feedback to individual instructor					
Agree group	58(83%)	4.064	.236	.331	.062
Disagree group	12(17%)	4.036	.366		
Input towards the teacher and teaching awards					
Agree group	39(56%)	4.025	.299	-1.223	-.105
Disagree group	31(44%)	4.101	.198		
Input for the administrative evaluation of job contracts					
Agree group	31(44%)	4.016	.313	-1.233	-.157
Disagree group	39(56%)	4.093	.206		
Input for the administrative evaluation of promotions					
Agree group	32(46%)	4.046	.3015	-.370	-.057
Disagree group	38(54%)	4.070	.224		
Assistance for students in selecting courses and instructors					
Agree group	35(50%)	4.075	.261	.503	-.010
Disagree group	35(50%)	4.043	.261		
Publication in school journals					
Agree group	15(21%)	4.137	.257	1.327	.073
Disagree group	55(79%)	4.038	.259		
Overall agreement on student ratings					
Agree group	57(81%)	4.090	.242	2.170*	.325**
Disagree group	13(19%)	3.921	.299		

Note: * $p < .05$; ** $p < .01$; *t* means independent *t* test; *r* means Pearson correlation.

Table 3 Summary of Independent *t* Test and correlation test for ratings scores of Method/Skill by faculty with positive attitude and negative attitude towards the used of student ratings

	<i>N</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>r</i>
A source of feedback to individual instructor					
Agree group	58(83%)	4.005	.196	.398	-.033
Disagree group	12(17%)	3.982	.209		
Input towards the teacher and teaching awards					
Agree group	39(56%)	3.993	.194	-.383	-.032
Disagree group	31(44%)	4.011	.204		
Input for the administrative evaluation of job contracts					
Agree group	31(44%)	3.993	.205	-.295	-.082
Disagree group	39(56%)	4.007	.193		
Input for the administrative evaluation of promotions					
Agree group	32(46%)	4.004	.209	.097	-.016
Disagree group	38(54%)	3.999	.189		
Assistance for students in selecting courses and instructors					
Agree group	35(50%)	4.002	.196	.026	-.004
Disagree group	35(50%)	4.000	.201		
Publication in school journals					
Agree group	15(21%)	4.088	.182	1.969*	.144
Disagree group	55(79%)	3.977	.230		
Overall agreement on student ratings					
Agree group	57(81%)	4.008	.200	.583	.125
Disagree group	13(19%)	3.972	.189		

Note: * $p < .05$; ** $p < .01$; *t* means independent *t* test; *r* means Pearson correlation.

Table 4 Summary of Independent *t* Test and correlation test for ratings scores of Assignment/ Examination by faculty with positive attitude and negative attitude towards the used of student ratings

	<i>N</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>r</i>
A source of feedback to individual instructor					
Agree group	58(83%)	4.033	.251	-.652	-.038
Disagree group	12(17%)	4.088	.331		
Input towards the teacher and teaching awards					
Agree group	39(56%)	4.034	.295	-.301	.001
Disagree group	31(44%)	4.053	.226		
Input for the administrative evaluation of job contracts					
Agree group	31(44%)	4.037	.308	-.150	-.011
Disagree group	39(56%)	4.047	.228		
Input for the administrative evaluation of promotions					
Agree group	32(46%)	4.060	.301	.523	.075
Disagree group	38(54%)	4.027	.233		
Assistance for students in selecting courses and instructors					
Agree group	35(50%)	4.074	.276	.985	.051
Disagree group	35(50%)	4.011	.253		
Publication in school journals					
Agree group	15(21%)	4.127	.264	1.410	.129
Disagree group	55(79%)	4.019	.263		
Overall agreement on student ratings					
Agree group	57(81%)	4.072	.253	2.001*	.271**
Disagree group	13(19%)	3.913	.288		

Note: * $p < .05$; ** $p < .01$; *t* means independent *t* test; *r* means Pearson correlation.

Table 5 Summary of Independent *t* Test and correlation test for the total ratings scores by faculty with positive attitude and negative attitude towards the used of student ratings

	<i>N</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>r</i>
A source of feedback to individual instructor					
Agree group	58(83%)	4.032	.211	.005	.006
Disagree group	12(17%)	4.032	.294		
Input towards the teacher and teaching awards					
Agree group	39(56%)	4.017	.251	.612	-.033
Disagree group	31(44%)	4.050	.189		
Input for the administrative evaluation of job contracts					
Agree group	31(44%)	4.014	.265	.575	-.074
Disagree group	39(56%)	4.046	.190		
Input for the administrative evaluation of promotions					
Agree group	32(46%)	4.035	.257	.117	.006
Disagree group	38(54%)	4.029	.196		
Assistance for students in selecting courses and instructors					
Agree group	35(50%)	4.046	.230	.520	.019
Disagree group	35(50%)	4.018	.222		
Publication in school journals					
Agree group	15(21%)	4.118	.240	1.695*	.142
Disagree group	55(79%)	4.008	.217		
Overall agreement on student ratings					
Agree group	57(81%)	4.057	.214	2.006*	.306**
Disagree group	13(19%)	3.921	.245		

Note: * $p < .05$; ** $p < .01$; *t* means independent *t* test; *r* means Pearson correlation.

Discussions and Conclusions

The purpose of this study has been to ascertain whether there is any difference between faculty with positive attitude and faculty with negative attitude towards student ratings regarding to their ratings scores. Table 6 is the summary of the independent *t* tests and correlation tests for faculty attitude towards student ratings and their ratings scores.

Table 6 Summary of Statistically Significant Independent *t* Test and Correlation Test for Faculty Attitude towards student Ratings and Their Ratings Scores

Dependent variables	Preparation/ Planning		Material/ Content		Method/ Skill		Assignment/ Examinationl		Tota	
	<i>t</i>	<i>r</i>	<i>t</i>	<i>r</i>	<i>t</i>	<i>r</i>	<i>t</i>	<i>r</i>	<i>t</i>	<i>r</i>
Independent variables										
A source of feedback to individual instructor										
Input towards the teacher and teaching awards										
Input for the administrative evaluation of job contracts										
Input for the administrative evaluation of promotions										
Assistance for students in selecting courses and instructors										
Publication in school journals					√				√	
Overall agreement on student ratings	√	√	√	√			√	√	√	√

Note: *t* means independent *t* test; *r* means Pearson correlation; √ means statistically significant.

As show in Table 6, faculty attitude towards the use of student ratings has little correlation with their ratings scores. However, there are two significantly differences (the ratings scores of Method/Skill and the total ratings scores) between faculty with positive attitude and those with negative attitude towards the use of student ratings as a publication in school journals. A possible explanation for this finding may be that faculty with positive attitude in publication teach more effectively than the other group. Therefore, they receive higher ratings scores and they feel more confident about their teaching. They may be willing or would like to have their ratings scores published in the school journals. That is teaching effectiveness, especially teaching method and skill, can be one of the main reasons for the faculty who feel positive about the publication of their ratings scores, and they likely receive higher ratings scores.

As faculty attitude towards the implement of student ratings policy, the overall- agreement group is rated statistically higher than is the overall-disagreement group on Preparation/Planning, Material/Content, Assignment/Examination, and the total rating scores. The coefficients of correlation between the overall agreement and these four ratings scores are statistically significant. The results indicate that the more positive faculty feel about the implement of student ratings, the higher ratings scores they receive, especially on Preparation/Planning, Material/Content, Assignment/Examination, the total ratings scores.

It should be noted here that although the correlation between faculty attitude and their ratings scores is correlated, it does not necessarily mean faculty attitude causes the range of ratings scores, or vice versa. As the explanation above, there may be another factors that influence both faculty attitude and their ratings scores. This will be another study for the future.

This study provides a starting point for investigating the effect of faculty attitude towards student ratings on their rating scores. It suggests that the more faculty member accept the policy of student ratings, the higher student ratings scores they receive. It is obvious that faculty with positive attitude towards student ratings have higher ratings scores than do faculty with negative attitude. A possible explanation may be that faculty

with positive attitude feel more confident about their teaching performance, therefore, they receive relatively high ratings scores. It will be very important for the school administrator to build up faculty's confidence about their teaching performance. In other words, it is essential and crucial to promote faculty teaching effectiveness. The more effectively faculty teach, the more confident they feel about their teaching. In the end, they will feel more positive about the student ratings because of their high ratings scores.

In summary, this study has pointed to the important of faculty attitude towards student ratings when studying the factors that explain variance in the results of student ratings scores. Although the results are informative, they should be taken as preliminary. The generalizability of the findings would be strengthened with replications over different instructors in different disciplines and with different evaluation instruments.

Acknowledgment

This article was presented at the annual meeting of American Psychological Association, held in Chicago, IL, August, 22-25, 2002. The author would like to thank the three anonymous reviewers for their comments.

The author would like to thank National Hualien Teachers College in Taiwan for their willingness to participate in this research.

References

- Abbott, R. D., Wulff, D. H., Nyquist, J. D., Ropp, V. A., & Hess, C. W. (1990). Satisfaction with processes of collecting student opinions about instruction: The student perspective. *Journal of Educational Psychology, 82*, 201-206.
- Armstrong, S. (1998). Are student ratings of instruction useful? *American Psychologist, 53*(11), 1223-1224.
- Arreola, R. A. (1995). *Developing a comprehensive faculty evaluation system*. Bolton, MA: Anker Publishing Company, Inc.
- Avi-Itzahak, T., & Kremer, L. (1986). An investigation into the relationship between university faculty attitudes toward student rating and organizational and background factors. *Educational Research Quarterly, 10*, 31-38.
- Baxter, E. P. (1991). The TEVAL experience 1983-1988: The impact of a student evaluation of teaching scheme on university teachers. *Studies in Higher Education, 16*, 151-178.
- Benson, D. E., & Lewis, J. M. (1994). Students' evaluation of teaching and accountability: Implications from the Boyer and ASA reports. *Teaching Sociology, 22*(2), 195-199.
- Brandenburg, D. C., Slinde, J. A., & Batista, E. E. (1977). Student ratings of instruction: Validity and normative interpretations. *Research in Higher Education, 7*, 67-78.
- Bridgeman, W. J. (1986). Student evaluation viewed as a group process factor. *Journal of Psychology, 120*, 183-190.
- Calderon, T. G., Gabbin, A. L., & Green, B. P. (1995). *Promoting and evaluating effective teaching: The report of the American Accounting Association teaching and curriculum section's committee on promoting and evaluating effective teaching*. Sarasota, FL: American Accounting Association.
- Centra, J. A. (1981). *Determining faculty effectiveness*. San Francisco: Jossey-Bass.
- Centra, J. A. (1993). *Reflective faculty evaluation*. San Francisco: Jossey-Bass.

- Centra, J. A., & Gaubatz, N. B. (2000). *Is there gender bias in student evaluations of teaching?* *Journal of Higher Education*, 70(1), 17-33.
- Chang, T. (2000). *Student ratings: What are teachers college students telling us about them?* Paper presented at the meeting of the American Educational Research Association, New Orleans, LA.
- Chang, T. (2002). Teachers college faculty attitudes toward student ratings: The comparison between required policy and optional policy. *Journal of Ping-Dung Teachers College*, 16, 205-230.
- 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.
- d'Apollonia, S., & Abrami, P. C. (1997). Navigating student ratings of instruction. *American Psychologist*, 52(11), 1198-1208.
- Feldman, K. A. (1977). Consistency and variability among college students in their ratings among courses: A review and analysis. *Research in Higher Education*, 6, 223-274.
- Feldman, K. A. (1983). Seniority and experience of college teachers as related to evaluations they receive from students. *Research in Higher Education*, 18, 3-124.
- Feldman, K. A. (1992). College students' views of male and female college teachers: Part I - Evidence from the social laboratory and experiments. *Research in Higher Education*, 33, 317-375.
- Feldman, K. A. (1993). College students' views of male and female college teachers: Part II-Evidence from students' evaluations of their classroom teachers. *Research in Higher Education*, 34, 151-211.
- Franklin, J., & Theall, M. (1989). *Who reads ratings: Knowledge, attitude and practice of users of student ratings of instruction*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.
- Greenwald, A. G., & Gillmore, G. M. (1998). How useful are student ratings? Research to comments on the current issues section. *American Psychologist*, 53(11), 1228-1229.
- Jacobs, L. C. (1987). *University faculty and students' opinions of student ratings*. Bloomington, IN: Bureau of Evaluative Studies and Testing. (ERIC Document Reproduction Service No. ED291291)
- Marsh, H. W. (1984). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases, and utility. *Journal of Educational Psychology*, 76, 707-754.
- Marsh, H. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research*, 11, 253-388.
- Marsh, H. W., & Dunkin, M. J. (1992). Students' evaluations of university teaching: A multidimensional perspective. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research* (pp. 143-233). New York: Agathon Press.
- Marsh, H. W., & Roche, L. (1993). The use of students' evaluations and an individually structured intervention to enhance university teaching effectiveness. *American Educational Research Journal*, 30, 217-251.
- O, T. Y. (1996). *American and Korean student's perceptions of student ratings of instruction: A comparative study*. Unpublished doctoral dissertation, University of Kansas.
- Owie, I. (1985). Incongruence between expected and obtained grades and students' ratings of the instructor. *Journal of Instructional Psychology*, 12, 196-397.
- Peterson, K. D. (1995). *Teacher evaluation: A comprehensive guide to new directions and practices*. Thousand Oaks, CA: Corwin Press, Inc.

- Rich, H. E. (1976). Attitude of college and university faculty towards the use of student evaluations. *Educational Research Quarterly, 1*, 17-27.
- Roche, L. A., & Marsh, H. W. (1998). Workload, grades, and students' evaluations of teaching: Clear understanding sometimes requires more patient explanations. *American Psychologist, 53*(11), 1230-1231.
- Seldin, P. (1993). The use and abuse of student ratings of professors. *The Chronicle of Higher Education, 46*, 40.
- Spencer, K. J. (1994). *Student perspectives on course and teacher evaluations*. Unpublished doctoral dissertation, Hofstra University, Hempstead, NY.
- Spencer, P. A., & Flyr, M. L. (1992). *The formal evaluation as an impetus to classroom change: Myth or reality?* (Research/Technical Report). Riverside, CA: Riverside Community College.
- Tatro, C. N. (1995). Gender effects on student evaluations of faculty. *Journal of Research and Development in Education, 28*, 169-173.
- Van Patten, J. (1994). *The politics of assessment of the professoriate*. Paper presented at the annual meeting of the Southeastern Philosophy of Education Society, Atlanta.
- Wachtel, H. K. (1998). Student evaluation of college effectiveness: a brief review. *Assessment and Evaluation in Higher Education, 23*(2), 191-211.
- Wagenaar, T. C. (1995). Student evaluation of teaching: Some cautions and suggestions. *Teaching and Sociology, 64*, 64-68.
- Wilson, R. (1998). New research casts doubt of student evaluation of professors. *The Chronicle of Higher Education, 44*(19), 12-14.

收 稿 日 期：2002 年 11 月 11 日

接 受 刊 登 日 期：2003 年 4 月 25 日

Appendix A

Table 7 The Correlation between Item and Subscale and Internal Consistency for Four Subscales of SRI

Dimension	Rating Items	γ	α
Preparation/	1. is concerned about the effectiveness of his/her teaching	.845***	.857
Planning	2. provides a detailed course syllabus	.893***	
	3. states course objectives for each class section	.912***	
Material/	4. relates the material of this course with other areas of knowledge	.924***	.917
Content	5. demonstrates knowledge and makes it clear how each topic fits into the course	.951***	
	6. is aware when students are having difficulty in understanding a topic	.915***	
	7. establishes and maintains an interaction	.903***	
Method/	8. keeps the course moving rapidly enough for the material	.945***	.933
Skill	9. explains material clearly	.945***	
	10. is helpful with difficulties	.931***	
Assignment/	11. gives good comments on written work	.928***	.927
Examination	12. gives fair grades	.956***	
	13. gives exams and papers appropriate for the course	.936***	

Note. γ means correlation between individual rating item and subscale; α means Cronbach α for internal consistency; *** $p < .001$

Table 8 The Result of Principal Components Analysis for each Subscale and Total Scale of SRI

Scale	Item number	Factor loading	Eigenvalue	Variance explained %
Preparation/ Planning	1	.885	2.390	79.7
	2	.880		
	3	.912		
Material/ Content	4	.912	2.531	84.4
	5	.938		
	6	.905		
Method/ Skill	7	.937	3.530	88.3
	8	.936		
	9	.943		
	10	.942		
Assignment/ Examination	11	.918	2.571	85.7
	12	.944		
	13	.915		
Total	Preparation/Planning	.895	3.478	87.0
	Material/ Content	.932		
	Method/Skill	.963		
	Assignment/Examination	.939		

Bulletin of Educational Psychology, 2003, 35(2), 183-200
National Taiwan Normal University, Taipei, Taiwan, R.O.C.

「學生評鑑教師教學」之結果： 教師態度有關嗎？

張 德 勝

國立花蓮師範學院
國民教育研究所

本研究主要的目的在探討教師對「學生評鑑教師教學」態度及其在評鑑得分之相關，並分析正面態度的教師與負面態度的教師，其評鑑結果的分數的差異情形。研究對象為八十七學年第二學期度國立花蓮師範學院七十位教師。研究工具有兩種，第一種是教師對「學生評鑑教師教學」態度的調查問卷(the Attitude towards Student Ratings of Instruction Questionnaire, ASRI)，主要是了解教師對「學生評鑑教師教學」的看法；另一種工具是「國立花蓮師範學院教學意見反映調查表」(the Student Ratings of Instruction form, SRI)，主要是了解學生對教師教學的滿意度。這項工具共包含四個層面，分別是「準備與計畫」、「教材與內容」、「教法與技巧」、「作業與評鑑」。

研究結果顯示如下：(一)對「學生評鑑教師教學」持正面態度的教師，在「準備與計畫」、「教材與內容」、「評鑑與作業」與整體得分的評鑑分數顯著的高於對負面態度的教師；(二)同意將評鑑結果公佈在學校期刊的教師，在「教法與技巧」與整體性方面顯著高於不同意的教師；(三)教師對「學生評鑑教師教學」整體的同意度與其被學生評鑑結果之分數有顯著的相關。

關鍵詞：學生評鑑教師教學、大學教師評鑑、態度