Effects of Perceived Teacher Personality on Student Class Evaluations: A Comparison between Japanese Instructors and Native English Speaking Instructors

Yoshitaka Tanabe¹ & Setsuko Mori¹

Correspondence: Setsuko Mori, Faculty of Law, Kinki University, 3-4-1 Kowakae, Higashiosaka City, Osaka, Japan. Tel: 81-6-6721-2332. E-mail: setsukomori@mac.com

Received: March 30, 2013 Accepted: April 27, 2013 Online Published: May 17, 2013

Abstract

This study explores how university students' perceptions of their classes and instructor personality contribute to their overall rating of the class. The study also investigates whether the students rate Japanese instructors and native English speaking instructors differently. Data was collected using a questionnaire comprised of instructional rating and teacher personality rating sections. The instructional rating section was based on the Instructional Rating Form (Tomasco, 1980), and European Portfolio for Student Teachers of Languages (Newby et al., 2007) whereas the teacher personality rating section was derived from Murray, Rushton, and Paunonen (1990). The study employed statistical analyses including multivariate analysis of variance (MANOVA), principal components analysis and regression analysis. The results of the study revealed that students' perceptions of the class as interesting, organized and clear, positively influenced their overall rating of the class. On the other hand, their perceptions of their instructor as aggressive, dominant, anxious, and authoritarian negatively affected the overall rating. The findings also indicated that when it comes to Japanese instructors, personality traits such as sociable, ambitious, intellectually curious, intelligent, and gentle influenced students' overall ratings.

Keywords: student evaluations, students' perceived teacher personality traits, instructional ratings, overall impression of class

1. Introduction

1.1 Student Evaluations – A Controversial Issue

A need for university reform has been emphasized throughout the world in the last few decades, and along with its history, it appears that student evaluations as a measure of faculty development have taken root in higher education. In Japan, for instance, 80 percent of universities currently conduct student evaluations according to the Ministry of Education, Culture, Sports, Science and Technology (2010), and it seems unlikely that student evaluations will fade out from university campuses anytime soon. While there has recently been a trend towards augmenting student ratings with other data sources to obtain student feedback on one's teaching quality and effectiveness (Newby et al, 2007), student evaluations seemingly have often been seen as "the most influential measure of instructional effectiveness" (d'Apollonia & Abrami, 1997). Despite student evaluations seeming to be firmly rooted in educational settings, there has always been controversy and discontent over the system. Much of this derives from such concerns as "student evaluations are no more than mere personality contests" (Tomasco, 1980). As such, it is, therefore, significant to ease such apprehension by seeking a better understanding of the nature of student evaluations.

1.2 Literature Review

Along with their history, student evaluations of instruction have always gained much attention among researchers in terms of their reliability, validity and usability. Some researchers have proposed that student evaluations are biased by factors that may be irrelevant to effective teaching. Cohen (1981) analyzed the relationship between student ratings of instruction and student achievement. Although the results provide strong support for the validity of student ratings as measures of teaching effectiveness, he warns of the possibility that higher grades in a course may actually reflect grading leniency, rather than student's perception of amount

¹ Faculty of Law, Kinki University, Osaka, Japan

learned. Greenwald and Gillmore (1997) similarly found that students' evaluative ratings of instruction correlated positively with expected course grades, that is, higher ratings would be expected in the more leniently graded courses. They criticized that the grades-ratings correlation is due to the unwanted influence of instructors' grading leniency on ratings, and indicated a rather negative attitude toward the uncritical use of student evaluations. They then suggested that student ratings should be statistically controlled for grading leniency. There are, however, some studies such as d'Apolloniaand Abrami (1997) and McKeachie (1997) which have questioned the conclusion and the suggestion that Greenwald and Gillmore made. The studies took a critical view of the data that Greenwald and Gillmore used in their analyses. Without being artificially controlled, average grades, course workload and ratings naturally vary by subject and by course level, and these factors are likely to influence the size of the correlation coefficients between student efforts, grades, and ratings. The Greenwald and Gillmore's study failed in removing the influence of the factors, or the range of courses, from their correlation coefficients, which after all made the results of the study questionable. Patrick (2011), in her extensive review of the research done on the possible effects of grading leniency, pointed out that such research is likely to merely prove a minimal relationship even though they reported positive correlations, and cannot be seen as strong evidence that reliable and valid student evaluation instruments are affected a great deal by grades.

Cashin (1995) argued that many variables, including student motivation, class size and such, may bias student evaluations and they should be controlled for by using appropriate comparative data. In response to these concerns about the validity of student evaluations, Harrison, Ryan, and Moore (1996) asserted that students have self-insight, a form of metacognition, into how they make decisions concerning teacher effectiveness since they have an implicit awareness of the relative importance of the factors they are considering. d'Apollonia and Abrami (1997) also claimed that student ratings are not affected by biasing variables since the General Instruction Skill, which they describe as global components of teaching such as delivering instruction, facilitating interactions, and evaluating student learning, is substantially correlated with student learning. Gigliotti and Buchtel (1990) criticize reluctant and negative attitudes of teachers towards student evaluations. They explored how "self-serving" bias affects student evaluations, which is a type of attributional bias referring to people's tendency to take credit for success and avoid blame for failure. Their analyses showed that the "self-serving" bias has minimal or nonexistent effect on evaluations and supported the validity of properly obtained student evaluations. They then claimed that it is the perception of bias by instructors that should be regarded as a dubious issue and that might be unjustifiably hindering teachers in using the evaluations as meaningful clues for improvement.

Several researchers have discussed the relationship between teachers' personality traits and student evaluations of teaching effectiveness. Radmacher and Martin (2001) investigated which of the following factors could be predictors of student evaluations; (a) teachers' age and extraversion traits, and (b) students' course grades, gender, enrollment status, academic abilities and age. The result of their study suggested that extraversion was the only significant predictor of student evaluations even after controlling for other factors. With regard to teacher's disposition of extraversion and student evaluations, Murray, Rushton, and Paunomen (1990) also reported a positive correlation between the two, and so did Patrick (2011) though, in her study, other personality traits such as openness, agreeableness, consciousness and neuroticism are also reported to have an effect on student evaluations. However, Kneipp, Kelly, Biscoe, and Richard (2010) found extraversion not to be significantly predictive of student's perception of instructional quality.

In an attempt to shed light upon this issue of teacher personality, Mori and Tanabe (2012) investigated; (a) if there is any correlation between students' perceived teacher personality and class evaluations, and (b) what teacher personality traits and instructional ratings contribute to the overall impression of the class. The study found four interpretable factors in terms of students' perceived teacher personality; Negative Affect, Extraversion, Achievement and Meekness, and their correlations with 24 items from the instructional rating were investigated. The result showed that all teacher personality factors except Meekness were found to have a considerable degree of influence on class evaluations. Especially, Achievement was most strongly correlated with all of the instructional ratings, and Extraversion also showed strong correlations with the majority of instructional ratings. With regards to the second research question, the analysis found that two personality factors, Extraversion and Negative Affect, and one instructional rating, "interprets clearly," significantly contributed to the overall evaluation of the class. The authors concluded that teacher personality traits influence student ratings and are likely to be a source of bias in evaluations.

1.3 Research Questions

Based on the findings of the previous research, the authors this time hypothesized that the Japanese students rate Japanese instructors and native English speaking instructors differently. Observing the faculty report on the

results of student evaluations, whose details are made referable for students and all faculty members, there seems to be a relatively high tendency that Japanese instructors are rated higher than native English speaking instructors. What factors contribute to the differences in the results of student evaluations between the two types of instructors? In order to better understand this perceived tendency based on the subjective observation of the authors, the following research questions were formulated for this study:

- 1. What instructional ratings contribute to the overall impression of the class?
- 2. What students' perceived teacher personality traits contribute to the overall impression of the class?
- 3. Does the relationship between instructional ratings and overall impression differ based on instructors' first language?
- 4. Does the relationship between students' perceived teacher personality trait and overall impression differ based on instructors' first language?

2. Method

2.1 Participants

The participants in this study were 160 first and second year law major students (74 first year and 86 second year) at a private university in Osaka, Japan. They were in seven different intact English classes that were randomly chosen. The number of students in these classes varied from 13 to 26 with the mean being 23. All of the first year students were taking both English 1 and Communicative English 1 whereas the second year students were taking English 2 and Communicative English 2. English 1 and 2 were taught by native Japanese speakers with a focus on reading and listening while Communicative English 1 and 2 were taught by native English speakers with a focus on oral communication. The same groups of students took these two classes over the course of a year. First year students were placed in their classes based on their performance on the TOEIC Bridge, administered at the beginning of the first semester, while second year students were placed in their classes based on their scores on the TOEIC, administered at the end of the first year. Their proficiency varied greatly from a low score of 110 to high score of 156 on the TOEIC Bridge, and a low score of 180 to a high score of 420 on the TOEIC.

2.2 Measures

The participants in this study completed two sections of a rating instrument, the instructional rating and teacher personality rating sections. The instructional rating section is comprised of 24 items including one item concerning the overall evaluation of the class. These items were written in Japanese based on the Instructional Rating Form (Tomasco, 1980), and European Portfolio for Student Teachers of Languages (Newby et al., 2007) (See the Appendix for translation). The teacher personality rating section consists of 28 items. All of the items were derived from Murray et al. (1990) and translated into Japanese. Although Murray's measures of personality included 29 items, one item concerned with aesthetical sensitivity was omitted as it was not relevant to the context (See the Appendix for translation). Except for the item asking about students' overall evaluation of the class on a 10 point Likert scale, all the items were on a six point Likert scale with one being strongly disagree and six being strongly agree.

2.3 Procedure

The questionnaire was administered at two different times, the end of the first semester and the end of the second semester. In the first administration, the participants were asked about the classes taught by Japanese instructors (English 1 or 2) and their instructors' personality. In order to create an environment where the participants would be able to answer the questionnaire more freely and comfortably, the instructors were requested to leave the room. Prior to administration, the participants were told that the questionnaire was anonymous and the results would never be exposed to the instructors or used for any other purposes but for research. The questionnaire was conducted and collected by one of the two researchers. In the second administration, the participants were asked about the classes taught by native English speaking instructors (Communicative English 1 or 2) and their instructors' personality. The participants were again reminded that the questionnaire was anonymous and the results would never be exposed to the instructors or used for any other purposes but for research. The questionnaire was conducted and collected by the Japanese instructors in their English 1 or 2 classes. The questionnaire was completed within approximately 15 minutes in each occasion.

3. Results

3.1 Reliability, Descriptive Statistics and Mean Differences

320 sets of responses, 160 regarding Japanese instructors and 160 regarding native English speaking instructors, were analyzed. The internal consistency estimates of reliability for the instructional rating section and teacher

personality section were calculated. Cronbach's Alpha was .97 and .81, respectively, which indicate that both sections of the questionnaire were highly reliable. Tables 1 and 2 show means and standard deviations of both sections of the questionnaire.

Table 1. Descriptive statistics and mean differences of the instructional rating sections

	NJS (n=160)		NI	ES (n=160)
	Mean	S.D.	Mean	S.D.
1. Arousal interest	4.73*	1.10	4.19*	1.15
2. Expands viewpoints	4.48*	1.12	4.11*	1.08
3. Informative lectures	4.81*	0.98	4.18*	1.10
4. Interprets clearly	5.10*	0.97	4.44*	1.14
5. Useful examples	4.80*	1.01	4.41*	1.13
6. Inspire confidence	3.97	1.28	3.65	1.22
7. Encourage initiative	3.99	1.25	3.70	1.23
8. Provides new tools	4.10*	1.30	3.54*	1.15
9. Stimulates thinking	4.20	1.19	3.92	1.14
10. Organized presentation	4.93	0.97	4.58	1.10
11. Uses time effectively	4.81	1.10	4.48	1.09
12. Respects opinions	4.74	1.10	4.49	1.08
13. Sensitivity	4.86	1.03	4.60	1.13
14. Fair examinations	5.15	0.85	4.99	1.02
15. Progress report	4.98*	0.97	4.36*	0.98
16. Class preparation	5.12	0.87	4.81	0.99
17. Challenges students	4.50	1.23	4.10	1.16
18. Motivates students	4.26	1.32	3.94	1.24
19. Good atmosphere	4.84	1.13	4.53	1.13
20. Clear rules	4.96*	1.02	4.54*	1.09
21. Effective materials	5.09*	1.01	4.73*	0.96
22. Clear evaluation	5.00	1.02	4.75	1.15
23. Challengeable assignments	4.62	1.26	4.29	1.11
24. Overall evaluation	8.50*	1.57	7.86*	1.77

^{*}p<.002

Note. NJS=native Japanese speaker, NES=native English speaker

Table 2. Descriptive statistics and mean differences of the personality rating sections

	NJS (n=160)		NES (n=160)	
	Mean	S.D.	Mean	S.D.
1. Meek	4.03	1.39	3.91	1.21
2. Ambitious	4.71	1.01	4.39	1.00
3. Sociable	5.14	1.16	5.22	1.07
4. Aggressive	2.19	1.21	2.24	1.34
5. Independent	3.59	1.34	3.64	1.30
6. Changeable	4.58	1.14	4.56	0.99
7. Seeks definiteness	3.55*	1.31	3.04*	1.12
8. Defensive	2.38	1.28	2.16	1.26
9. Dominant	2.25	1.28	2.30	1.32
10. Enduring	4.07	1.10	4.11	1.06
11. Attention-seeking	2.59	1.19	3.05	1.44
12. Harm-avoiding	2.93	1.21	3.00	1.13
13. Impulsive	2.38	1.17	2.55	1.24
14. Supporting	4.69	1.19	4.95	1.00
15. Orderly	4.70*	1.06	4.10*	1.24
16. Fun-loving	4.88	1.11	5.13	1.04
17. Approval-seeking	3.85	1.30	3.83	1.11
18. Seeks help and advice	2.93	1.16	3.01	1.19
19. Intellectually curious	4.41	1.07	4.16	1.03
20. Anxious	2.14	1.17	2.11	1.22
21. Intelligent	4.24	1.25	4.08	1.10
22. Liberal	4.53	1.12	4.48	1.03
23. Shows leadership	4.38	1.13	4.24	1.12
24. Objective	4.91	1.19	4.78	1.21
25. Compulsive	3.17*	1.27	2.65*	1.23
26. Authoritarian	2.36	1.34	2.27	1.23
27. Extraverted	4.31	1.25	4.59	1.16
28. Neurotic	1.89	1.05	2.03	1.15

^{*}p<.0017

A one-way multivariate analysis of variance (MANOVA) was conducted to determine the effect of the two types of instructors (native Japanese speakers and native English speakers) on the 24 instructional ratings as dependent variables. Significant differences were found between the two types of instructors on the dependent variables, Wilks's Λ .=74, F(24, 295)=4.22, p<.01. Analyses of variances (ANOVA) on each dependent variable were conducted as follow-up tests to the MONOVA. Using the Bonferroni method, each ANOVA was tested at the .002 level (.05 divided by 24). The results of ANOVA show that the former group scored significantly higher than the latter group on nine instructional ratings and overall evaluation (See Table 1).

A one-way multivariate analysis of variance (MANOVA) was also conducted to determine the effect of the two types of instructors (native Japanese speakers and native English speakers) on the 28 personality ratings as dependent variables. Significant differences were found between the two types of instructors on the dependent variables, Wilks's Λ .=78, F(28, 291)=3.01, p<.01. Analyses of variances (ANOVA) on each dependent variable were conducted as follow-up tests to the MONOVA. Using the Bonferroni method, each ANOVA was tested at the .0017 level (.05 divided by 28). The results imply that the Japanese instructors were perceived to seek more

definiteness, be more orderly, and more compulsive (See Table 2).

3.2 Significant Predictors of Overall Evaluation

3.2.1 Factor Analyses of the Instructional and Personality Ratings

In order to reduce the instructional items, first principal components analysis was performed. Four criteria were used to determine the number of factors to rotate: a minimum eigenvalues of 1.0, the scree test, a minimum loading of .45, and the interpretability of the factor solution. Based on these criteria, two factors were rotated using a Varimax rotation procedure. The result found four interpretable factors, which accounted for65.22 % variance (See Table 3).

Table 3. Principal components analysis summary for the instructional rating section: Eigenvalues and percent of variance explained

Factor		Initial Eigenvalues	
	Total	% of Variance	Cumulative %
1	13.13	57.08	57.08
2	1.87	8.15	65.22

Table 4. Principal components analysis results for the instructional rating section

	Factor		
	I	II	h^2
1. Arousal interest	0.75	0.40	0.72
2. Expands viewpoints	0.80	0.34	0.75
3. Informative lectures	0.70	0.41	0.67
5. Useful examples	0.56	0.55	0.62
6. Inspire confidence	0.79	0.26	0.70
7. Encourage initiative	0.85	0.21	0.77
8. Provides new tools	0.76	0.20	0.62
9. Stimulates thinking	0.75	0.39	0.71
17. Challenges students	0.75	0.44	0.76
18. Motivates students	0.84	0.33	0.81
4. Interprets clearly	0.52	0.59	0.61
10. Organized presentation	0.22	0.75	0.62
11. Uses time effectively	0.35	0.69	0.60
12. Respects opinions	0.36	0.64	0.55
13. Sensitivity	0.39	0.69	0.63
14. Fair examinations	0.21	0.76	0.62
15. Progress report	0.41	0.65	0.59
16. Class preparation	0.18	0.80	0.66
19. Good atmosphere	0.51	0.60	0.63
20. Clear rules	0.31	0.70	0.58
21. Effective materials	0.28	0.74	0.63
22. Clear evaluation	0.34	0.72	0.64

As Tables 3 and 4 show, Factor 1 accounted for 57.08% of variance. It was labeled as Interest in Class as most of the items loaded on this factor seem to be concerned with how interesting the class is. Factor 2, accounted for

8.15%, was defined as Class Management because many of the items seem to tap into students' perception of how organized and clear the instructions are.

Principal components analysis was performed with the personality ratings section as well. The same criteria were used to determine the number of factors to rotate for this section. Item 5 (Independent) and item 11 (attention-seeking) were eliminated from the analysis as they did not load on any factor clearly.

Table 5. Principal components analysis summary for the personality rating section: Eigenvalues and percent of variance explained

Factor		Initial Eigenvalues			
	Total	% of Variance	Cumulative %		
1	8.56	32.91	32.91		
2	4.27	16.44	49.35		
3	1.53	5.90	55.25		
4	1.14	4.38	59.62		

Table 6. Principal components analysis results for the personality rating section

		Comp	onent		
	I	II	III	IV	h2
2. Ambitious	0.73	-0.11	0.10	0.20	0.60
3. Sociable	0.65	-0.37	0.20	-0.11	0.61
6. Changeable	0.78	-0.25	0.16	-0.05	0.69
10. Enduring	0.69	0.04	0.07	0.00	0.49
14. Supporting	0.62	-0.38	0.24	-0.12	0.60
16. Fun-loving	0.69	-0.36	0.23	-0.15	0.68
19. Intellectually curious	0.75	-0.03	0.05	0.25	0.62
21. Intelligent	0.65	-0.12	0.06	0.29	0.52
22. Liberal	0.81	-0.05	0.08	0.02	0.67
23. Shows leadership	0.80	0.00	0.00	0.17	0.67
24. Objective	0.63	-0.33	0.10	0.12	0.53
27. Extraverted	0.63	-0.10	0.14	-0.40	0.58
4. Aggressive	-0.13	0.80	-0.06	0.07	0.67
8. Defensive	-0.12	0.81	0.13	0.11	0.70
9. Dominant	-0.11	0.84	-0.05	-0.08	0.73
13. Impulsive	-0.02	0.75	0.09	-0.22	0.62
20. Anxious	-0.27	0.74	0.14	0.00	0.63
25. Compulsive	-0.08	0.55	0.04	0.46	0.52
26. Authoritarian	-0.26	0.71	0.03	0.15	0.58
28. Neurotic	-0.30	0.66	0.20	-0.05	0.57
1. Meek	0.29	-0.28	0.49	0.19	0.43
12. Harm-avoiding	0.04	0.40	0.46	0.08	0.38
17. Approval-seeking	0.30	0.06	0.72	-0.11	0.62
18. Seeks help and advice	0.12	0.36	0.69	0.04	0.63
7. Seeks definiteness	0.14	0.43	-0.12	0.45	0.43
15. Orderly	0.37	-0.16	0.19	0.74	0.74

As Tables 5 and 6 show, Factor 1 accounted for 32.91% of variance. Factor 1 was labeled Extraversion/Achievement as items normally indicative of extraversion (i.e., sociable, fun-loving, and extraverted) and items generally indicative of achievement (i.e., ambitious, intellectually curious, and intelligent) loaded together on this factor. Factor 2, accounted for 16.44% of variance, was interpreted as Negative Affect as high scorers on this factor were perceived by their students as aggressive, dominant, anxious, authoritarian, and neurotic. Factor 3, Meekness, was so termed as the items that loaded on this factor included approval-seeking, and harm-avoiding, and was accounted for 5.90%. Factor 4 obtained loadings from only two items definiteness seeking and orderly, thus defined as Orderliness.

3.2.2 Significant Predictors of Overall Evaluation for All Instructors

To determine which instructional and personality items contributed to the overall evaluation of the class, a multiple regression analysis was performed between the overall rating (item 24 of the instructional rating section) as a dependent variable, and the factor scores of the two instructional components and the factor scores of the four teacher personality traits as independent variables. The linear combination of strength measures was significantly related to the overall rating, F(6, 313)=107.99, p<.00. The sample multiple correlation coefficient was .82, indicating that approximately 67% of the variance of the overall rating in the sample can be accounted for by the linear combination of personality and instructional measures.

Table 7 shows indices to indicate the relative strength of the individual predictors. Notice that the correlation coefficients for both instructional components and one personality trait, Negative Affect, are significant at p<.01. The result suggests that the more students perceive the class as interesting and clear, and the less students see the teacher asaggressive, dominant and authoritarian, the higher their overall evaluation of the class gets.

Table 7. The bivariate and partial correlations of the predictors with the overall rating of all instructors

Predictors	Correlation between each predictor and the overall rating	Correlation between each predictor and the overall rating controlling for all other predictors
Interest in class	0.61**	0.63**
Class management	0.54**	0.50**
Extraversion/Achievement	0.57	0.05
Negative Affect	-0.29**	-0.18**
Meekness	0.13	0.09
Orderliness	0.15	-0.06

^{**}p<.001

3.2.3 Significant Predictors of Overall Evaluation for Japanese Instructors

In order to investigate whether the students reacted to Japanese instructors and native English instructors differently, a multiple regression analysis was performed separately between two groups. Table 8 shows students' responses toward Japanese instructors. The linear combination of strength measures was significantly related to the overall rating, F(6, 153)=53.43, p<.00. The sample multiple correlation coefficient was .82, indicating that approximately 68% of the variance of the overall rating in the sample can be accounted for by the linear combination of personality and instructional measures.

Table 8 shows indices to indicate the relative strength of the individual predictors. The correlation coefficients for both instructional components and three personality traits, Extraversion/Achievement, Negative Affect, and Meekness are significant at p<.01. The fact that Meekness correlates positively with the overall rating, there is a possibility that traits normally indicative of meekness such as harm-avoiding and approval and advice-seeking may not be regarded as unfavorable by this population of the participants.

Table 8. The bivariate and partial correlations of the predictors with the overall rating of Japanese instructors

Predictors	Correlation between each predictor and the overall rating	Correlation between each predictor and the overall rating controlling for all other predictors
Interest in class	0.60**	0.47**
Class management	0.53**	0.34**
Extraversion/Achievement	0.61*	0.19*
Negative Affect	-0.42**	-0.29**
Meekness	0.27*	0.17*
Orderliness	0.20	0.05

^{*}p<.05

3.2.3 Significant Predictors of Overall Evaluation for Native English Instructors

A multiple regression analysis was also performed with the native English speaking instructors separately between two groups. The linear combination of strength measures was significantly related to the overall rating, F(6, 153)=56.35, p<.00. The sample multiple correlation coefficient was .82, indicating that approximately 68% of the variance of the overall rating in the sample can be accounted for by the linear combination of personality and instructional measures.

Table 9 shows indices to indicate the relative strength of the individual predictors. The correlation coefficients for both instructional components and one personality trait, Negative Affect, are significant at p<.01. Interestingly enough, the result suggests that all but one personality traits count when it comes to Japanese instructors whereas only students' negative affect can predict their overall evaluation of native English speaking instructors.

Table 9. The bivariate and partial correlations of the predictors with the overall rating of native English speaking instructors

Predictors	Correlation between each predictor and the overall rating	Correlation between each predictor and the overall rating controlling for all other predictors
Interest in class	0.60**	0.68**
Class management	0.52**	0.56**
Extraversion	0.55	0.00
Negative Affect	-0.20*	-0.17*
Meekness	0.02	0.09
Orderliness	-0.01	-0.12

4. Discussion and Conclusion

The results of this research suggest some major conclusions concerning the possible existence of bias that may affect student evaluations. First, in the comparison on the instructional rating scores between Japanese instructors and native English speaking instructors, the former group has been found to score higher than the latter on all the 24 items including the overall evaluation of the class, and among them, nine instructional ratings and overall evaluation showed a significant difference. The result, therefore, turned out to support the authors' observation. Then, what background factors lie behind this Japanese instructors' superiority in student ratings? One possible explanation would be the amount of target language, English in this context, that instructors mainly use in class. Communicative English classes are expected to be taught entirely in English, whereas English 1 and 2 are taught by Japanese instructors who frequently use their first language Japanese in such cases as when they explain complicated concepts like grammar or when their focus is not on the use of the target language or the

^{**}p<.01

contents of lesson's topic but on instructions and assignments during the class. Polio and Duff (1994) argue the benefits of first language use in language class and one of them is reducing stress and anxiety arising from a lack of understanding about the target language. It could be easily assumed that students, especially ones with low English ability, are often placed under pressure and stress when they fail to clearly understand what their teacher says in English. In association with this, most of the students are not accustomed to the communicative approach adapted in the Communicative English classes, which is often described as student-centered learning style, because English classes are commonly taught in lecture or teacher-centered style in high schools in Japan. Cortazzi (1990) analyzed that Japanese students prefer drilling and memorizing the materials presented by their teachers in contrast to American students who consider a classroom to be a place for developing and discussing their critical ideas. English native instructors often expect their students to proactively participate in class activities and to take initiative to be outspoken though Japanese students are often labeled as being shy and not being risk-taking (Paul, 2003). As Cortazzi (1990) suggests, many students in the current study may also be confused about the educational and cultural expectations of their Communicative English instructors. In sum, the inevitable uncertainty derived from use of the target language and such anxiety and frustration of students attached to different educational and cultural expectations may have caused the difference in rating.

Secondly, a multiple regression analysis shown in 3.2.2 has revealed that one personality trait, Negative Affect, significantly contributed to the overall evaluation of the class. The result is consistent with Mori and Tanabe (2012) to a certain extent. The results provide some support for the aforementioned claim that teachers' personality traits, as judged by students, affect student evaluations. It, thus, seems reasonable to suggest that one cannot and should not regard student ratings as a bias-free instrument to evaluate the instructional effectiveness of a teacher. Having said that, the multiple regression analysis has at least warranted the validity of student evaluations as both instructional components, Interest in class and Class management, were proven to correlate positively with the overall rating.

In relation to this point, the results shown in 3.2.3 have indicated that the same pattern was found for both types of teachers in terms of relationship between instructional ratings and overall impression. The relationship was found not to differ regardless of the first language of the instructor. When it comes to the relationship between students' perceived teacher personality traits and overall impression, on the other hand, all but one personality traits of Japanese instructors affected the overall rating while Negative Affect was the only significant predictor with regards to native English speaking instructors. Japanese instructors are evidently more likely to be influenced than native English speaking instructors. With the data collected, it remains unclarified why the personality-related predictors of overall rating differ between the two types of instructors. Further studies are needed to identify a possible explanation.

In conclusion, the current study confirmed that the Japanese students rate Japanese instructors and native English speaking instructors differently. It also asserts that students' perceived teacher personality traits are more than likely be one of the possible sources of bias that may affect student evaluation. Although the current study was conducted only in a limited environment and the results may not reflect the dynamics of student's responses, the authors believe that the results discussed here have given a meaningful insight in better understanding the fundamental nature of student evaluations which need to be examined from a multidimensional perspective.

References

- Cashin, W. E. (1992). Student ratings: The need for comparative data. *Evaluation and Faculty Development*, 12, 1-6.
- Cohen, P. A. (1981). Student ratings of instruction and student achievement: A meta-analysis of multisection validity studies. *Research in Higher Education*, *51*, 281-309.
- Cortazzi, M. (1990). Cultural and educational expectations in the language classroom. In B. Harrison (Ed.), *Culture and the Language Classroom. ELT Documents 132* (pp. 54-65). London: Macmillan. Modern English Publications and the British Council.
- d'Apollonia, S., & Abrami, P. C. (1997). Navigating student ratings of instruction. *AmericanPsychologist*, *52*(11), 1198-1208. http://dx.doi.org/10.1037/0003-066X.52.11.1198
- Gigliotti, R. J., & Buchtel, F. S. (1990). Attributional bias and course evaluations. *Journal of Educational Psychology*, 82(2), 341-351. http://dx.doi.org/10.1037/0022-0663.82.2.341
- Greenwald, A. G., & Gilmore, G. (1997). Grading leniency is a removable contaminant of student ratings. *American Psychologist*, 52(11), 1209-1217. http://dx.doi.org/10.1037/0003-066X.52.11.1209
- Gross, R. B., & Small, A. C. (1979). A survey of faculty opinions about student evaluations of instructors.

- Teaching of Psychology, 6(4), 216-219. http://dx.doi.org/10.1207/s15328023top0604 7
- Harrison, P. D., Ryan, J. M., & Moore, P. (1996). College students' self-insight and common implicit theories of rating of teaching effectiveness. *Journal of Educational Psychology*, 88(4), 775-782. http://dx.doi.org/10.1037/0022-0663.88.4.775
- Kneipp, L. B., Kelly, K. E., Biscoe, J. D., & Richard, B. (2010). The impact of instructor's personality characteristics on quality of instruction. *College Student Journal*, 44(4), 901-905.
- McKeachie, W. J. (1997). Student ratings: The validity of use. *American Psychologist*, *52*(11), 1218-1225. http://dx.doi.org/10.1037/0003-066X.52.11.1218
- Ministry of Education, Culture, Sports, Science and Technology. (2010). *Daigakuniokerukyoikunaiyonado nokaizenjokyonitsuite* [Report on educational reforms of universities]. Retrieved from http://www.mext.go.jp/a_menu/koutou/daigaku/04052801/__icsFiles/afieldfile/2010/05/26/1294057_1_1.pd f
- Mori, S., & Tanabe, Y. (2012). Effects of perceived teacher personality on student class evaluations. *Kinki University Center for Liberal Arts and Foreign Language Education Journal (Foreign Language Edition)*, 3(1), 163-181.
- Murray, H., Rushton, P., & Paunonen, S. (1990). Teacher personality traits and student instructional ratings in six types of university courses. *Journal of Educational Psychology*, 82(2), 250-261. http://dx.doi.org/10.1037/0022-0663.82.2.250
- Newby, D., Allan, R., Fenner, A., Jones, B., Komorowska, H., & Soghikyan, K. (2007). European portfolio for student teachers of languages: A reflection tool for language teacher education. Retrieved from http://archive.ecml.at/mtp2/publications/C3 Epostl E internet.pdf
- Patrick, C. L. (2011). Student evaluations of teaching: Effects of the Big Five personality traits, grades and the validity hypothesis. *Assessment & Evaluation in Higher Education*, 36(2), 239-249. http://dx.doi.org/10.1080/02602930903308258
- Paul, D. (2003). Teacing English to children in Asia. Hong Kong: Longman Asia ELT.
- Polio, C., & Duff, P. (1994). Teachers' language use in university foreign language classrooms: A qualitative analysis of English and target language alternation. *Modern Language Journal*, 78, 313-326. http://dx.doi.org/10.1111/j.1540-4781.1994.tb02045.x
- Radmacher, S. A., & Martin, D. J. (2001). Identifying significant predictors of student evaluations of faculty through hierarchical regression analysis. *The Journal of Psychology*, 135(3), 259-268. http://dx.doi.org/10.1080/00223980109603696
- Tomasco, A. (1980). Student perceptions of instructional and personality characteristics of faculty: A canonical analysis. *Teaching of Psychology*, 7(2), 79-83. http://dx.doi.org/10.1207/s15328023top0702_3

Appendix

Translation of Instructional Rating and Teacher Personality Sections of the Questionnaire

In this questionnaire, you are asked about this class and your impression on the instructor. When answering the questions, please keep in mind the following two points:

- *The results of the questionnaire will never be exposed to the instructor. Your personalinformation will not be provided to the instructor.
- *Although the results will be statistically analyzed and reported at a conference and/or in a journal, the responses to individual items will not be disclosed.

Read each statement and circle the number to indicate the extent to which you agree or disagree with each statement.

1 Strongly disagree 2 Somewhat disagree 3 Disagree

4 Agree 5 Somewhat agree 6 Strongly agree

A. About this class

- 1. This class aroused my interest.
- 2. This class expanded my viewpoints.
- 3. This class provided me with useful knowledge.
- 4. This class was clear and easy to understand.
- 5. This class offered useful examples.
- 6. This class inspired my confidence.
- 7. This class encouraged me to take an initiative in learning.
- 8. This class provided me with new learning tools.
- 9. The content of this class stimulated my intellectual curiosity.
- 10. This class was clearly organized.
- 11. Time was effectively used in this class.
- 12. The instructor respected students' opinions.
- 13. The instructor tried to meet students' needs.
- 14. The examinations were fair.
- 15. The objects of the class were clear and reasonable/
- 16. The teacher was well prepared for the class.
- 17. The content of this class was challenging.
- 18. This class motivated me to study English.
- 19. The class atmosphere was appropriate.
- 20. This class was disciplined.
- 21. Materials and instruments used for this class were effective.
- 22. Evaluation criteria for presentations and assignments were clearly explained.
- 23. Presentation tasks and assignments were challenging.
- 24. Give your overall evaluation to this class on a scale from 1 to 10.

B. About the instructor

- 1. unassertive and conformable
- 2. ambitious
- 3. friendly and sociable
- 4. argumentative and gets angry easily
- 5. avoids restraints and enjoys being free
- 6. flexible, and likes new and different ideas
- 7. does not like ambiguity
- 8. suspicious and takes offense easily
- 9. forceful and attempts to control environment
- 10. patient and enduring
- 11. attention-seeking

- 12. careful and avoids risks
- 13. impulsive and reckless
- 14. supporting and gives sympathy
- 15. neat and organized
- 16. fun-loving
- 17. works for approval of others
- 18. desires support and sympathy from others
- 19. intellectually curious and insightful
- 20. nervous and unstable
- 21. bright and intelligent
- 22. liberal
- 23. shows leadership
- 24. fair and free of bias
- 25. meticulous and perfectionistic
- 26. authoritarian and opinionated
- 27. extraverted and optimistic
- 28. moody and constantly worried things will go wrong