


Evaluating A Math Instructor's Clarity: An Analysis Of Low-To-Moderate Items Versus A High Inference Item

Kelly D. Bradley, (E-mail: kdbradley@uky.edu), University of Kentucky
James W. Bradley, Lexington Community College

ABSTRACT

Instructors and students identify communication skills as being essential to successful classroom teaching and student learning. This study investigates the stability of student evaluations with respect to items related to instructor's clarity. The majority of student evaluations completed at the end of a semester at the college level contain a question directly related to clarity. A pilot study was conducted which utilized four finite mathematics courses being taught by the same instructor. After assessing reliability and validity, a follow-up study was conducted utilizing four sections of introductory statistics employing the same design as the pilot. Ninety-four students completed some component of the researcher-constructed evaluation resulting in an overall 96% response rate. Seventy-one students completed the entire evaluation, allowing for a comparison between the high inference and the low-to-moderate inference items used to measure instructor's clarity, resulting in a useable response rate of 72%. After a statistical transformation of the high inference item, descriptive statistics were produced for each type of item. Through a t-test, a statistically significant difference was found between the high and low-to-moderate clarity items. The difference between the types of clarity ratings was then analyzed via ANOVA techniques to explore differences by gender and expected course grade. Males tended to rate the instructor inconsistently, assigning lower ratings for the low-to-moderate inference items, while other ratings appear to demonstrate consistency.

INTRODUCTION

 ver time, clarity has been consistently recognized as a quality that good instructors possess. Often, clarity is the aspect of instruction influencing how students remember an instructor. Encarta defines clarity as the quality of being clearly expressed, clearness in what somebody is thinking, or the quality of being clear in sound or image (Encarta, 2003-2004). In general, clarity can be thought of within one of three realms, expression, thought, or reproduction. Hines (1981) defines clarity in teaching as a collection of "teacher behaviors which relate to and facilitate the communications to students of the subject matter of instruction in such a manner as to make this content intelligible to them thus enabling student learning of that subject matter" (pp. 16-17).

The process of clear instruction can be best realized through various behaviors. It is important for the instructor to outline the objectives for a particular lesson indicating to the students what they are to learn. Also, if an instructor wants students to write something down, the instructor should write the material on the board. Instructor handwriting should be neat and easy to read, and the instructor should speak clearly, at an appropriate pace, and with an enthusiasm that encourages student engagement. Additionally, examples presented in class should be rooted in real world application, related to students' assigned work. Finally, an instructor should pay attention to the verbal and non-verbal cues of the students since it is best to anticipate difficulties at the early stages. As a general rule, to present an effective and clear lecture, the instructor should have direct contact with the students, an effective use of language, the ability to manage time, and a propensity to be entertaining (Gelula, 1997; Herber, 1978; Thielens, 1987).

This paper examines college students' evaluations of instructor's clarity through a two-stage study utilizing a baseline pilot study and a follow-up study. In both studies students enrolled at a four-year institution in a large Midwest city completing their required mathematics sequence were asked to voluntarily complete a survey that evaluated the students' perceptions of the clarity of their mathematics instructor. The instructor was held consistent across classes and throughout the stages of the study. The pilot study focused on students enrolled in finite mathematics, while the follow-up focused on those enrolled in introductory statistics. The purpose of the research was to compare low-to-moderate inference items with an overall high inference item, uncovering similarities and differences among and between student evaluations of instructor of clarity.

THEORETICAL PERSPECTIVE

Many studies have found strong correlations between teacher clarity and student learning (Evans & Guymon, 1978; Hines, 1981; Cruickshank & Kennedy, 1986). Furthermore, much of the literature indicates that an instructor's overall effectiveness depends on a combination of instructor attributes, with major contributors including clarity of the instructor's lectures, the course's organization, and the degree to which the instructor motivates students (Arreola, 1995; Centra, 1993; Feldman, 1976). Given these insights, it is reasonable to argue that classroom students' perception of instructor's clarity is an area deserving attention.

The majority of research on clarity of instruction has taken place at the pre-college level, with only a few studies being completed in the higher education settings. College and university literature reveals the significance of clear teaching to both students and instructors. Feldman (1989) reviewed studies on college instructors' classroom behaviors and found that among twenty-two instructional dimensions, 'clarity and understandableness' had the highest importance level as measured by four performance indicators. A study completed at Virginia Commonwealth University during the fall 1998 semester found that students and faculty agreed that factors related to clear communication were important. Similar findings were reported by Spencer and Schmelkin (1995), who concluded that students considered issues of clarity, fairness, and respect to be the most important in determining instructor effectiveness (O'Toole, Spinelli, & Wetzel, 2000). While clarity has been shown to be an important factor to all students, research indicates that clarity is assigned higher importance by students in mathematics and science based courses, as compared to humanities and social sciences at the college level (Feldman, 1976; Hativa and Raviv, 1993).

Research on student evaluations suggests student evaluation surveys are generally consistent, but student biases may exist that affect their responses. A contention is made throughout the literature that student evaluation surveys are generally valid (Arreola, 1995; Centra, 1993; Marsh and Bailey, 1993; Marsh and Roche, 1997; Seldin, 1993). One study conducted at the college level by Hativa and Raviv (1996) suggested that students have a good perception of the nature of clarity, supported by the consistent ratings of clarity across time and courses. There is opposition that postulates the effectiveness of ratings is biased by instructor characteristics and course expectations, such as the grades that students expect to receive or the difficulty of the material presented (D'Appollonia & Abrami, 1997; Wilson, 1998). Nelson and Lynch (1984) found that grading did not have a significant impact on student perceptions of effectiveness; however, this finding was diminished due to the omission of a number of questionnaire items regarding instructor characteristics. In contrast, Robinson (1995) found a significant difference in teacher ratings was associated with differences in academic performance. Gender is another characteristic that is brought out in the literature. Robinson posits that teacher ratings do not vary by gender of student. Based on these findings, a continuation of the study of the relationship between student ratings and expected grade, as well as other teacher and student characteristics should be in place.

Given the affirmation of the importance of clarity in the classroom, along with its connection to an instructor's effectiveness, operationalization of the clarity variable on student evaluations becomes a significant point of interest. When clarity is measured by student ratings on teacher evaluations, it is common for it to be a single item, something comparable to 'rate your instructor's clarity', making it a high inference measure. The way the variable is measured, percentage versus likert-type scale can also compound the issue. Several studies dealing with college instruction have identified low-to-moderate inference behaviors that promote clarity. Using such indicators, asking multiple questions that together define clarity allows for a more direct measure (Hines, 1981; Murray, 1983; Smith &

Land, 1980). Within the area of mathematics-based curriculum, Smith and Land's research performed such a task. As so, the indicators they used will be reproduced in this study.

BACKGROUND

Pilot Study

A pilot study was conducted using four finite mathematics classes to assess the reliability of the instrument, identify characteristics associated with differences between the high inference variable and low-to-moderate inference variables and explore reasonable hypotheses. The four classes were taught by the same professor, who held Assistant rank, having approximately four years of teaching experience at the college level teaching predominately introductory mathematics and statistics courses. This experience coupled with the professor's educational background of a B.S. in accounting and mathematics and a M.S. in statistics classified the individual as a qualified instructor. The four classes were similar in course structure, each being composed of a lecture component and a lab component, both of which were taught by the professor.

Course characteristics and instructional settings were similar. Specifically, the instructor was responsible for all four classes, the course material was consistent, with the exception of the accelerated group's class pace, and all classes had access to information and assignments for the course via the instructor's web page. Student majors and class size were two areas of concern. Each class had a different primary composition of majors; however, they all shared the same curriculum of mathematics courses. An interview with the professor provided the researchers with a better understanding of the structure of the classes. A section letter assigned by the institution denotes the four classes. Class X, the accelerated class, had an enrollment of 35 students, 18 females and 17 males that were primarily business administration majors. Students enrolled in the accelerated program were required to be at least 21 years old and were expected to have access to a home computer and the Internet. The class met on Saturday morning for seven weeks. The professor noted the accelerated class was more attentive, had better test scores, and higher attendance and assignment completion rates than its counterparts. Class P, a full term class, had an enrollment of 20 students, 12 females and 8 males that were primarily accounting majors. The class met on Tuesday and Thursday afternoons. The professor considered this class to be traditional, with the majority of students being between the ages of 18-22, having a broad range of attitudes toward school, and tending to be more talkative than the accelerated class. Another full term class, Class D, had an enrollment of 31 students, 5 females and 26 males, most of whom were computer information systems majors. The class met on Wednesday and Thursday mornings. Again, the professor viewed this class as traditional. Finally, the full term Class Q was composed of business administration majors and had an enrollment of 23 students, 8 females and 15 males. Meetings took place on Monday and Tuesday afternoons. The professor commented that this group was similar to the others; however, this class was presented the material first each week. Given the differences between classes, it was of interest to see if class was a significant variable when studying the difference between the high and low-to-moderate inference items.

The instrument used to collect data was constructed of items used to identify teacher behaviors that promote clarity. It included one high inference question, asking the students to rate the overall instructor clarity, followed by 14 low-to-moderate inference items related to clarity that were taken from a study completed by Metcalf (1989) formulated from Smith and Land's (1980) research. Eighty-nine surveys were collected, and all data collected were useable, resulting in an overall response rate of 82%. Class response rates were: class X = 80%, class P = 90%, class D = 81%, and class Q = 78%. Thirty-five of the respondents were female and 54 were male. For the comparative component, seventy-eight students completed the survey in its entirety, allowing for a difference to be calculated.

One of the main objectives of the pilot study was to establish reliability of the instrument. A correlation analysis was conducted for the low-to-moderate clarity items. All correlation coefficients were positive, ranging from 0.076 to 0.764, with the majority of correlations being moderate to high. Forty-three of the seventy-eight pairs had a coefficient of 0.45 or greater. A reliability estimate, Cronbach's alpha, was also calculated for the Likert-type component of the clarity survey to provide a measure of internal consistency. For the 14 items, alpha was calculated to be approximately 0.92. Thus, using the inter-class correlations and the reliability estimate, it is reasonable to argue the items are measuring the same dimension, which has been operationalized as clarity of instructor.

A significant difference was found between the high and low-to-moderate inference items, indicating students' perceptions of the scales are indeed different. Given the variety in classes, it was important to ascertain if class was a variable that should be of importance in the full study. An ANOVA (see Table 1) on the difference of clarity ratings indicated that class was not significant.

Table 1: (Pilot Study) One-Way ANOVA for Difference with Class as the Independent Variable

Source	df	F	p
Class	3	2.25	0.090
Error	74		

Implications for the Follow-Up Study

The pilot study established the reliability of the instrument, as well as support for the unidimensionality of the operationalized low-to-moderate items clarity variable. Moreover, it was established that differences were not detected by class, providing reason to not collect detailed class demographics in the follow-up study. The significant difference between the ratings of clarity lead to a hypothesis, $H_a: \mu_{\text{high}} - \mu_{\text{low/moderate}} = 0$. Finally, based on student comments, one item was added to the final clarity survey and an additional item underwent minor revisions.

METHODOLOGY

Research Questions

This research was directed by specific questions. First, the study was driven by the need to compare the low-to-moderate inference clarity ratings with the high inference clarity rating. It was also important to assess the correlation between the low-to-moderate inference clarity items. In general, clarity skills perceived to be exhibited by the instructor and the agreement between student ratings were analyzed. Whether the students indicated teacher evaluations were useful to the instructor and if they felt their instructor would use the information provided on the survey was a critical component to the validity of the research. Finally, differences in student perceptions of clarity were studied by expected grade and sex.

Participants

A census sample of four introductory statistics classes taught by the same professor at a four-year college in a large Midwest city was utilized. The same professor was used in the pilot and follow-up; however, the professor did have an additional year of experience when conducting the follow-up study. In the pilot study, no significant difference was found by class. Here, the instructor was responsible for all classes, the course material was consistent, and all classes had access to information and assignments via the instructor's web page. Even though a variety of student majors were represented, all students shared the same core of mathematics curriculum. Class sizes were not equal; still, no class would be classified as small or large, resulting in comparable situations. Thus, given the results of the pilot, class was purely a descriptive variable.

Instrumentation

The instrument used to collect data was constructed of items used to identify teacher behaviors that promote clarity. It included one high inference question, asking the students to rate the overall instructor clarity, followed by 15 low-to-moderate inference items related to clarity, parallel to that of the pilot study. The high inference variable was recorded as a percentage value between 0 and 100 by the students and later transformed to a 4-point rating scale using the equation, $\frac{3x}{100} + 1$, where x is the student assigned percentage rating. All of the items applied in the survey were Likert-type items recognized consistently throughout the literature. Because of this, only minor edits were made

to the wording of items; however, items having a five-point scale were reduced to a four-point scale, eliminating the neutral category, corresponding to 1 = strongly disagree and 4 = strongly agree. The demographic section consisted of four questions consistently brought forth in the literature. Two questions were related to personal characteristics, asking students to report their sex and expected letter grade in the course. The final questions were related to the students' feeling toward the instrument itself. Specifically, they were asked to indicate if they felt teacher evaluations were useful to the instructor and if they felt their instructor would use the information provided. Finally, students were provided an opportunity to comment on the understandability of the items and any additional concerns related to the instrument.

Procedures

Throughout the literature, it is indicated that instructors and students feel that communication skills are relevant to a successful classroom. Students overwhelmingly feel that it is essential for instructors to have the quality of 'clarity' in order for them to succeed and achieve in the course. Thus, the underlying assumption of this study is that teachers make a difference and that clarity of instruction is an important skill.

Prior to completing the evaluation, students were given assurance their identity would not be revealed in regards to their responses and were informed that participation was voluntary. The questionnaire was administered following the students' final examination. Once the data were collected, the responses were entered into a spreadsheet. No response was found to be out of range, meaning all data supplied were credible. Since students were informed that they were not required to respond to all items, missing data were treated as missing.

The analyses began with the calculation of response rates, followed by a check of representativeness of the sample by gender and major. Then, means, standard deviations, and sample sizes were recorded for all the clarity items and correlation estimates were computed. Using the average scores for the fourteen low to medium inference items, weak and strong clarity skills of the professor were identified as perceived by the students. It was expected that the standard deviation would be approximately 1 for the Likert-type items; thus, items with a standard deviation below 0.5 were considered to have an abnormally high agreement, while items with standard deviations above 1 were considered to have a large disagreement. The overall clarity rating was transformed to a 1 – 4 scale that could be compared to the average of the Likert-type items. Using the equation described above in instrumentation, 0% would be equivalent to 1, 25% = 1.75, 50% = 2.5, 75% = 3.25 and 100% = 4. A difference was calculated by subtracting the average of the low-to-moderate inference items from the rescaled high inference item. A t-test for this difference between the two types of clarity ratings was performed followed by an ANOVA on the difference to explore gender and expected grade effects. When significant differences were found, descriptive statistics were produced for the clarity items based upon the specified variable.

DISCUSSION/RESULTS

Ninety-four students completed some component of the researcher-constructed evaluation resulting in an overall 96% response rate. Seventy-one students completed the entire evaluation, allowing for a comparison between the high inference and the low-to-moderate inference items used to measure instructor's clarity, resulting in a useable response rate of 72%.

The high inference clarity rating is simply the student's overall perception of instructor's clarity on a scale from 0 – 100%. Seventy-one students provided a rating, with an average of 92.41% and a standard deviation of 7.45 indicating a student perception that the instructor is very clear. There was some variability in the ratings; however, it should be noted that the distribution of clarity scores was skewed left, indicating that the majority of the students assigned the instructor a high rating. Means and standard deviations were calculated for the low-to-moderate inference clarity items as well. Students assigned ratings from 1 to 4, where 1 indicated strongly disagree and 4 indicated strongly agree. All items had an average above 3, ranging from 3.14 to 3.82, suggesting the students perceived the instructor to display all the low-to-moderate inference clarity skills. The lowest scoring item was the instructor shows similarities and differences between things", and the highest scoring item was "the instructor works examples and

explains them”. No standard deviation was larger than 1, indicating that students had high agreement on all items. On the other end, none of the items had abnormally low standard deviations.

Table 2: Means and Standard Deviations of Clarity Survey Items

Clarity Item	Measure		
	<i>n</i>	Mean	SD
Assigned Percent	71	92.41	7.45
Q1	94	3.55	0.71
Q2	94	3.63	0.80
Q3	93	3.71	0.79
Q4	94	3.53	0.81
Q5	93	3.77	0.77
Q6	93	3.41	0.82
Q7	94	3.74	0.78
Q8	94	3.77	0.75
Q9	94	3.52	0.89
Q10	94	3.39	0.78
Q11	93	3.44	0.84
Q12	94	3.61	0.78
Q13	94	3.45	0.82
Q14	94	3.61	0.83
Q15	94	3.70	0.75

As was the case in the pilot study, all low-to-moderate items were positively correlated with each other, with a range from 0.55 to 0.96. The highest correlation estimate corresponds to the pair of items ‘the instructor writes important things on the board (formulas, charts)’ and ‘the instructor works examples and explains them’. The item ‘the instructor provides time for students to practice/work examples’ was also highly correlated with the above pair. The item “the instructor was clear” was moderately correlated with all other survey items.

Table 3: Pearson Correlations for Low-to-moderate Inference Clarity Items

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14
Q2	0.65													
Q3	0.79	0.77												
Q4	0.75	0.70	0.75											
Q5	0.71	0.85	0.87	0.79										
Q6	0.65	0.68	0.69	0.65	0.66									
Q7	0.71	0.80	0.90	0.73	0.89	0.71								
Q8	0.74	0.85	0.90	0.80	0.96	0.70	0.91							
Q9	0.68	0.61	0.76	0.67	0.68	0.57	0.73	0.70						
Q10	0.69	0.62	0.77	0.72	0.66	0.70	0.70	0.69	0.71					
Q11	0.61	0.67	0.72	0.63	0.72	0.61	0.74	0.73	0.55	0.71				
Q12	0.75	0.76	0.86	0.73	0.83	0.68	0.85	0.85	0.69	0.70	0.71			
Q13	0.60	0.73	0.67	0.65	0.64	0.66	0.64	0.65	0.53	0.68	0.59	0.73		
Q14	0.72	0.68	0.77	0.74	0.76	0.63	0.76	0.81	0.66	0.71	0.75	0.74	0.70	
Q15	0.72	0.80	0.87	0.74	0.89	0.64	0.85	0.93	0.68	0.72	0.77	0.81	0.69	0.81

The literature claims that if students feel that evaluations are important to their instructors that the data collected from such instruments will have increased validity. Thus, students were asked if they felt that teacher evaluations were useful to instructors. All individuals surveyed responded, with 90.22% answering ‘yes’. In order to focus on the instructor, students were asked if they felt their instructor would use the information provided on the

survey. Again, all students responded. Over ninety-three percent felt that the instructor would use the information. Only nine of the ninety-two responses indicated the instructor would not use the results and even less students, six, said that the information was not useful. This is an indication that the students perceived their input as valuable, which is crucial to generating valid results.

Table 4: Frequencies and Percents for Variables Related to Perceptions of Survey Information

Variable	Measure			
	Frequency		Percent	
	Yes	No	Yes	No
Teacher Evaluation	83	9	90.22	9.78
Information Useful	86	6	93.48	6.52

Expected grade and gender were two variables of interest. Thus, descriptive statistics were produced for the difference between the two types of variables. Viewing the descriptive statistics, it is believed that students are assigning higher ratings to the instructor’s clarity when using the high inference variable. Upon further investigation, (see Table 5), it appears that females tend to be quite consistent in their assigned ratings of clarity. However, males appear to be assigning a higher rating on the single question, percentage clarity item. Student’s expected grade has different values; however, when stratified by grade, all differences are positive, indicating students are assigning higher value clarity ratings to the high inference, single item. When reviewing the grades and the average difference between the perceived clarity ratings, it is apparent that students expected to receive a ‘D’ have the least consistent assignment of rating for the instructor by far, with an average mean difference of 0.68. This could be partly due to the fact that only 6 students fell into this category. Attention should also be given to the high standard deviation associated with this measure, indicating little agreement among the students expecting to receive a grade of ‘D’. All other grades have values less than 0.2, with the most consistent assignment coming from those who expect to earn a grade of ‘C’.

Table 5: Descriptive Statistics for the Differences by Gender and Expected Grade

Variable	n	Measure	
		Mean	SD
Gender			
Female	24	-0.003	0.211
Male	43	0.296	0.814
Expected Grade			
A	16	0.190	0.711
B	24	0.168	0.756
C	21	0.072	0.267
D	6	0.680	1.127

A statistical difference was identified for the mean difference of the low-to-moderate inference item ratings and the rescaled high inference rating at the $\alpha = .05$ level (see Table 6). This is a clear indication that students are not using the two different measures of clarity consistently.

Table 6: T-Test for the Mean Difference

Variable	n	T	P
Difference	71	2.32	0.023*

*p < .05

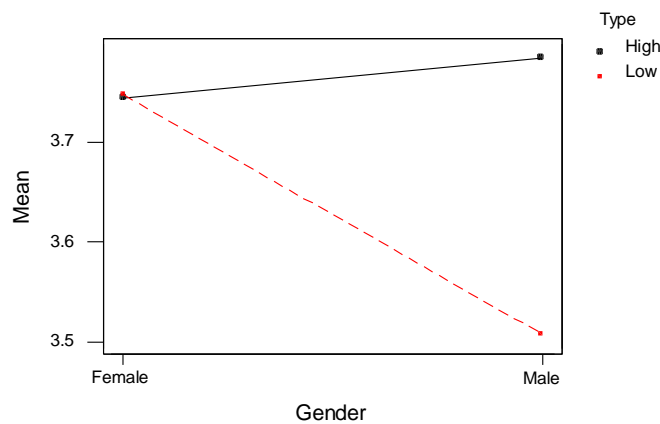
As a final step, an ANOVA was carried out with the mean difference as the dependent variable and gender and grade as the independent variables. The ANOVA (see Table 7) produced a non-significant interaction term. No difference was identified by student’s expected grade for the difference between the clarity ratings, likely due to the majority of students receiving the grade of A, B and C, which were much more consistent than their counterpart grade of D. Gender was found to be significant at the $\alpha = .05$ level. The gender effect is clearly evident in the graphical summary pictured below.

Table 7: General Linear Model for Difference with Gender and Grade as Independent Variables

Source	df	F	p
Gender	1	4.57	0.037*
Grade	3	1.43	0.243
Gender*Grade	3	0.32	0.814
Error	59		

*p < .05

Plot of Means for the Two Types of Rating Scales



In the above, females mean ratings for average of the low-to-moderate clarity items is similar to the rescaled average high inference item. This is not the case for males, as their inconsistency is demonstrated by the gap between the lines. The phenomenon that is occurring could have multiple explanations. It could be that males are paying more attention to the breakdown of items, responding deliberately, where females, in general, is simply providing an overall rating, even when asked for different information. It might be that females have already associated the low-to-moderate items with the high inference item, without even being told to do so. Why males are less consistent than females is unclear but should be given attention in future studies.

CONCLUSION

Throughout the literature, it is indicated that instructors and students feel that communication skills are relevant to a successful classroom. Students overwhelmingly feel that it is essential for instructors to have clarity in order for them to succeed and achieve in the course. What is lacking is an understanding in how clarity is measured and if students are consistent with the ratings when considering low-to-moderate items versus high inference items.

Overall, the instructor in this study was deemed to possess all clarity characteristics with relatively high averages, with the overall average clarity rating assigned by the students being 92.41% out of a possible 100% for the high inference variable and an average 3.5881 out of 4 on the low-to-moderate inference items. It was found that the

low-to-moderate inference items were all positively correlated with one another, with the majority of items being correlated moderately, demonstrating the items tap into the same construct, clarity. A significant difference was identified for the average difference of the clarity ratings, indicating that high inference and low-to-moderate inference items are not easily distinguishable by all students. Here specifically, it seems that males are the group assigning inconsistent ratings.

This study sets a foundation for further research to be done on measuring student's perceptions of clarity, specific to classroom evaluation. There is evidence to support that not all students assign consistent ratings when evaluating clarity if measured in different manners. If it holds that males are indeed less consistent in evaluation ratings of clarity, assigning higher values to high inference items, ramifications could exist for the typical classroom evaluation. An instructor or professor teaching a classroom having a majority of male students could be at an advantage or disadvantage when it comes to teacher evaluations, depending on how the item for clarity was measured.

Further research might include following these students over time through their math sequence. This would help identify relationships between instructor clarity and academic performance or satisfaction over time. It is also necessary to study various instructors and subject areas. Mathematics type courses were chosen here since the literature consistently recognizes that students enrolled in math classes assign clarity great importance. With various professors teaching multiple sections of the same course, practice effects and class size could also be investigated. Finally, a qualitative study, possibly using focus groups, of the students whose high and low-to-moderate clarity ratings differed by +/- 1 standard deviation would be insightful. Such a follow-up could provide insight into inconsistency in student ratings of instructor characteristics, specifically clarity.

REFERENCES

1. Arreola, R.A. (1995). *Developing a comprehensive faculty evaluation system*. Boston: Anker.
2. Centra, J.A. (1993). *Reflective faculty evaluation*. San Francisco: Jossey-Bass.
3. Cruickshank, D.R. & Kennedy, J.J. (1986). Teacher clarity. *Teacher and Teacher Education*, 2(1), 43-67.
4. D'Appollonia, S. & Abrami, P.C. (1997). Navigating student ratings of instruction. *American Psychologists*, 52(11), 1198-1208.
5. *Encarta World English Dictionary* (North American Addition). (2003-2004). Microsoft Corporation. Bloomsbury Publishing Plc.
6. Evans, W. & Guymon, T. (1978). Clarity of explanation: A powerful indicator of teacher effectiveness. Paper presented at the annual meeting of the American Educational Research Association, Toronto, Canada (ERIC Document Reproduction Service No. ED 151 321).
7. Feldman, K.A. (1976). The superior college teacher from the student view. *Research in Higher Education*, 5, 243-288.
8. Feldman, K.A. (1989). The association between student ratings of specific instructional dimensions and student achievement. *Research in Higher Education*, 30(6), 583-645.
9. Gelula, M.H. (1997). Effective lecture presentation skills. *Surg Neurol*, 47, 201-204.
10. Hativa, N. & Raviv, A. (1993). Using a single score for summative teacher evaluation by students. *Research in Higher Education*, 34(5), 625-646.
11. Hativa, N. & Raviv, A. (1996). University instructor's ratings profiles: Stability over time and disciplinary differences. *Research in Higher Education*, 37(3), 341-365.
12. Herber, H. L. (1978). *Teaching Reading in Content*. Prentice-Hall.
13. Hines, C. (1981). A further investigation of teacher clarity: The observation of teacher clarity and the relationship between clarity and student achievement and satisfaction. Unpublished doctoral dissertation, The Ohio State University.
14. Marsh, H.W. & Bailey, M. (1993). Multidimensional students' evaluations of teaching effectiveness: A profile analysis. *Journal of Higher Education*, 64(1), 1-18.
15. Marsh, H.W. & Roche, L.A. (1997). Making students' evaluations of teaching effectiveness effective: The critical issues of validity, bias, and utility. *American Psychologists*, 52(11), 1187-1197.
16. Metcalf, K. (1989). An investigation of the efficacy of a research-based regimen of skill development on the instructional clarity of preservice teachers. Unpublished doctoral dissertation, The Ohio State University.

17. Murray, H.G. (1983). Low-inference classroom teaching behavior and student ratings of college teaching effectiveness. *Journal of Educational Psychology*, 75(1), 138-149.
18. Nelson, J.P. & Lynch, K.A. (1984). Grade inflation, real income, simultaneity, and teaching evaluations. *Journal of Economic Education*, 15(Winter), 21-37.
19. O'Toole, D., Spinelli, M., & Wetzel, J. (2000). The important learning dimensions in the school of business: A survey of students and faculty. *Journal of Education for Business*, 75(6), 338-342.
20. Robinson, D. (1995). A critical analysis of students' teacher ratings. *Person. Individual. Diff.*, 19(6), 945-947.
21. Seldin, P. (1993). The use and abuse of student ratings of professors. *The Chronicle of Higher Education*, (July 21), A40.
22. Smith, L.R. & Land, M.L. (1980). Student perception of teacher clarity in mathematics. *Journal for Research in Mathematics Education*, 11, 137-146.
23. Spencer, K.J. & Schmelkin, L.P. (1995). Student perspectives on teaching and its evaluation. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.
24. Thielens, W. (1987). The disciplines and undergraduate lecturing. Paper presented at the annual meeting of the American Educational Research Association, Washington, April.
25. Wilson, R. (1998). New research casts doubt on value of student evaluations of professors. *The Chronicle of Higher Education*, (January 16), A12.