Validating the Use of Teaching Effectiveness Data in the Improvement of Teaching

Jess Godbey
Jacksonville State University
jgodbey@jsu.edu

Terry Marbut
Jacksonville State University
tmarbut@jsu.edu

Abstract

Using data as a catalyst for initiating changes to improve program performance is critical in any effective assessment strategy. Successful assessment programs typically measure performance and implement improvement strategies in the identified areas of student learning, student services, and to an in increasing degree, teacher effectiveness. In spite of the long running debate over their usefulness [1], student evaluation of teaching (SET) has long been a major component in most assessment strategies involving teaching effectiveness. The US Department of Education's Educational Resources Information Center system cites more than 2,000 articles on research that focuses on student evaluations. Assuming that the primary goal of such instruments is to actually improve teaching effectiveness, this paper explores one department's use of statistical analysis of faculty teaching ratings from graduating seniors in applied engineering and technology programs. This information was used by the department chair to develop appropriate professional development activities and provide performance feedback to individual faculty.

Introduction

Successful academic programs must include performance assessments to ensure the implementation of program improvements. Seldin [2] noted that colleges and universities were moving to concerted and sustained efforts to improve teaching in programs. Since most academic accreditation standards require both student evaluation of the program and application of improvement strategies for teaching effectiveness, continual improvements related to these areas continue to be important to institutions of higher learning. Halloran [3] even pointed out that government in some countries demand that universities be judged on their performance in teaching to meet both regulatory and funding requirements.

The existing research regarding issues that impact teaching effectiveness is extensive. One such issue is faculty recognizing the need for improvement. Blackburn [4] found in interviews with almost 300 college teachers, that 92% of the teachers believed their own teaching was above average. It may be that within a department, faculty members hold these same beliefs and do not actively engage in efforts to improve their teaching effectiveness. While it might be argued that these faculty members could indeed be above average if compared with all college teachers or with the institutional average, it is a fact that "within a department" all cannot be above average.

Another issue is the relationship between teaching styles and learning styles. There is ample research that suggests students learn best when they are taught in ways that match their style of learning [5-8]. There is also significant research that reveals no such correlation [9, 10, 11]. Some of this research suggests that the best correlation is actually between content and teaching style and contends that students learn a given content better through a particular teaching style even if the style does not match the students' preference [11]. Renaldi and Garung [10] suggests that including various teaching styles in a course should provide all students with something that best matches their learning preference and increase the students' engagement and therefore enhance learning and student satisfaction.

While this project does not attempt to further investigate any of the noted research, the project was based on the realization that there are very likely faculty members in a given technology and engineering department that utilize more effective teaching styles than others in the same department. Faculty members' ratings on senior surveys were statistically compared to identify any such faculty with the intent of developing improvement strategies to enhance student learning within this particular technology and engineering department.

Survey Description

The graduating senior survey was strongly encouraged but not a mandatory requirement for graduation. The actual response rate was above 90%. The rating scale for the teaching effectiveness of a faculty member ranged from A (excellent) to D (very poor). There was also an E rating (no opinion) if the student did not have the faculty member as an instructor. In order to compare the mean rating of each faculty member, points of 4, 3, 2, and 1 were assigned for answers A through D respectively. The rating of E, no opinion, was considered a non-rating and was not counted. Since an ordinal rating scale was used, the distinction must be made that the subjects taking the survey may not have felt that the difference between each step was equal. An A rating may have been considered twice as good as a B rating to some subjects but only marginally better to others. For the purpose of this comparison, it was assumed that each step in the ratings scale was equal.

Statistical Analysis Performed

Minitab was selected as the software package for data analysis. An analysis of variance (ANOVA) determined if there was a statistically significant different in the average ratings of the faculty members. T tests were utilized to compare faculty members' older ratings to their newer ratings to determine if differences over time were significant. Finally, the variances in ratings for faculty members were compared.

Comparison of Mean Scores and Analysis of Variation

To determine if any faculty member had a significantly higher rating than the other faculty members, a one-way analysis of variance (ANOVA) was performed. Groupings were made using the Tukey method with a 95% simultaneous confidence interval. Results of this analysis are presented in Table 1. Table 1 shows the means and grouping by faculty member. The listed means of faculty members that do not share a letter were significantly different.

Table 1. Grouping information using Tukey method

Faculty	N	Mean	Grouping
3	189	3.7460	A
7	167	3.6886	A B
1	144	3.5139	B
5	133	3.2105	C
4	105	3.2000	C
2	185	3.1838	C
6	66	3.0909	C

Based on these Tukey groupings, faculty members 3 and 7 belonged to group A and were not significantly different. Faculty members 1 and 7 belong to group B and were not significantly different, but faculty members 1 and 3 were significantly different. The rest of the faculty members, 2, 4, 5, and 6, all belonged to group C and were not significantly different from each other but were significantly different from 1, 3, and 7.

The next statistical test that was performed compared the variation associated with each faculty member's student ratings. A standard deviation test was performed to identify significant differences. The results of the test are given in Table 2 and Table 3. Analysis of faculty standard deviations revealed that faculty member 3 had the lowest variation in student scores and faculty members 2 and 6 seemed to have more variation than most of the others. The standard deviations of faculty members 2 and 6 were significantly higher than the standard deviation of faculty member 3.

Hypothesis Tests: Comparing Past and Present Effectiveness

The final tests were designed to determine if faculty member ratings were different now in comparison with ratings from earlier years. Two-Sample T-Tests were conducted for each faculty member comparing the last two years' ratings with their earliest two years of ratings in the sample. The number of years between present and earliest ratings varies depending on an individual faculty member's years of service. The hypothesis tested for each faculty member was that the rating means for past and present were equal. The tests were all performed at the alpha equal to the 0.05 level. These tests revealed that most faculty members had lower ratings in recent years as compared to earlier ratings. However, rating means for faculty members 4, 5, and 6 over the last two years were not statistically different than their earliest rating means. The other faculty members did have different rating means for the two time periods and all appeared to be somewhat lower.

Summary of Statistical Analysis

One faculty member had a mean rating that was significantly ($\alpha = .05$) higher than those of all other faculty members and the smallest variation in rankings. Comparison of most recent ratings with ratings from earlier years generally revealed that the ratings have decreased over time but the decrease was not statistically significant for most faculty members.

Table 2. Faculty Rating Variation

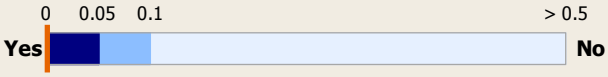
Standard Deviations Test for Faculty 1, Faculty 2, Faculty 3, Faculty 4,...
Descriptive Statistics Report

	Statistics			
Sample	Sample Size	Mean	Standard Deviation	Individual 95% CI
Faculty 1	117	3.4957	0.63821	(0.56341, 0.73526)
Faculty 2	130	3.0615	0.85139	(0.76905, 0.95698)
Faculty 3	134	3.7164	0.52904	(0.44057, 0.64470)
Faculty 4	76	3.2895	0.74504	(0.63830, 0.89265)
Faculty 5	97	3.2680	0.66951	(0.60149, 0.76059)
Faculty 6	66	3.0909	0.85444	(0.72882, 1.0324)
Faculty 7	123	3.6341	0.59047	(0.48154, 0.73578)

Table 3. Summary report of faculty rating variation significance

Standard Deviations Test for Faculty 1, Faculty 2, Faculty 3, Faculty 4,...
Summary Report

Do the standard deviations differ?



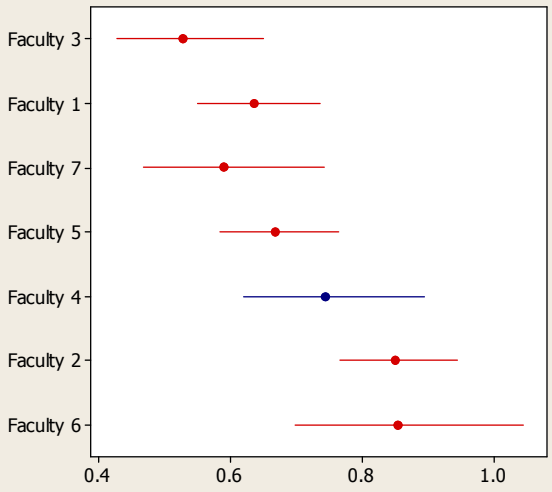
P = 0.000

Differences among the standard deviations are significant (p < 0.05).

Which standard deviations differ?

#	Sample	Differs from
1	Faculty 3	6 7
2	Faculty 1	6
3	Faculty 7	6
4	Faculty 5	6
5	Faculty 4	
6	Faculty 2	1 2 3 4
7	Faculty 6	1

Standard Deviations Comparison Chart
Red intervals that do not overlap differ.



Comments

You can conclude that there are differences among the standard deviations at the 0.05 level of significance. Use the Comparison Chart to identify standard deviations that differ. Red intervals that do not overlap indicate standard deviations that differ from each other. Consider the size of the differences to determine if they have practical implications.

Use of Findings

The data clearly indicated that one faculty member received higher ratings from students than other faculty members in the Department of Technology and Engineering. A check of grading patterns for this faculty member revealed that the grades students received were no better than grades received from other faculty members within the department. This fact negates any tendency to suggest that the higher ratings might be the result of easier grading. Upon extensive review, after analysis of the data, it was apparent that the highest-rated faculty member incorporated teaching strategies that relate well to departmental students. The consistency of this faculty member's ratings (small standard deviation) was especially intriguing. With the idea that this highly rated faculty member might be able to share some insight with the other members of the department, the department head asked this faculty member to develop and deliver a "Best Practices in Teaching" seminar for the department. The faculty member developed and delivered the seminar and feedback was solicited from faculty members to assess its usefulness. The seminar centered around three key points: course structure and organization, feedback to students, and real-world application.

The second significant discovery was that the scores for the highest-rated faculty members have decreased slightly over a period of years. Each of these faculty members was made aware of these findings and asked to examine their practices to determine if some change in their approach to teaching may be contributing to the lower ratings. During these discussions it was noted that our student population has changed significantly from non-traditional older working students to traditional age college students during this time period. This may be a contributing factor explaining, at least in part, these slight decreases in faculty ratings. Faculty members were encouraged to include teaching methodologies that match the learning styles of these younger students. As a result, the department acquired 25 iPads that are provided to students for portions of selected classes to increase the engagement of the students as active learners.

The third finding was that the scores for the lowest-rated faculty members were not changing significantly up or down. These faculty members were made aware of the findings by the department head to encourage them to explore ways to improve teaching effectiveness, including implementation of methods discussed in the previously mentioned teaching seminar. The department head also reminded faculty members that excellence in teaching is a major criterion for success in a regional institution.

Feedback from Faculty Members

A short survey was administered to all faculty members following the seminar to help assess the effectiveness of the professional development opportunity. Some questions were designed to gather information about the overall effectiveness of the information presented, the format of the seminar, and the time allotted for discussion. One question asked whether faculty members were likely to incorporate an idea presented in the seminar into a class. Survey results were encouraging. All faculty members agreed, and the majority strongly agreed that the seminar should be helpful in improving teaching effectiveness and that they were likely to incorporate at least one idea from the seminar into a class that they teach.

Conclusion

Preliminary results from the project indicate that systematic review and analysis of available data can be useful in developing effective strategies for improving teaching effectiveness. A longitudinal study will be needed to determine if teaching effectiveness in the department was actually strengthened as a result of this project.

References

[1] Centra, J.A. (1993). *Reflective Faculty Evaluation: Enhancing Teaching and Determining Faculty Effectiveness.* San Francisco: Jossey-Bass.

[2] Seldin, P., (1995), *Improving College Teaching.* Bolton, MA: Anker Publishing Company.

[3] Halloran, P. (2010). Using Case Studies as a Lens to Observe Teaching Evaluations. *Journal of Computing and Information Technology*, *18*, 133-139.

[4] Blackburn, R.T., Bober, A., O'Donnell, C., & Pellino, G. (1980). *Project for Faculty Development Program Education: Final Report.* Ann Arbor, Ml: University of Michigan, Center for the Study of Higher Education.

[5] Lovelace, M. K. (2005). Meta-Analysis of Experimental Research Based on the Dunn and Dunn Model. *The Journal of Education Research*, *98*(3), 176-183.

[6] Mahlios, M. C. (2001). Matching Teaching Methods to Learning Styles. In B. H. Stanford & K. Yamamoto (eds.), *Children and Stress: Understanding and Helping*. (pp. 65-73). Olney, MD, US: Association for Childhood Education International.

[7] Ogden, W. R. (2003). Reaching All the Students: The Feedback Lecture. *Journal of Instructional Psychology, 30(*1), 22-27.

[8] Stanberry, A. M., & Azria, E. M. (2001). Perspectives in Teaching Gerontology: Matching Strategies with Purpose and Context. *Educational Gerontology*, *27*(8), 639-656.

[9] Garton, B. L., Spain, J. N., Lamberson, W. R., & Spiers, D. E. (1999). Learning Styles, Teaching Performance, and Student Achievement: A Relational Study. *Journal of Agricultural Education*, *40*(3), 11-20.

[10] Rinaldi, C. & Gurung, R. (2008, 26 October). Should Teaching and Learning Styles Match? *Teaching Forum*. Retrieved from http://www.uwosh.edu/programs/teachingforum/public_html/?module=displaystory&story_id=648&format=html

[11] Pashler,H., McDaniel, M., Rohrer, D., & Bjork, R. (2009). Learning Styles: Concepts and Evidence. *Journal of Psychological Science in the Public Interest*, *9*(3), 106-116.

Biographies

JESS GODBEY is program coordinator of the Occupational Safety and Health Program at Jacksonville State University. He earned a Ph.D. in Industrial & Systems Engineering from Auburn University and M.S. and B.S. degrees from the University of Michigan. Dr. Godbey also has experience in the automotive industry, working with both Ford Motor Company and General Motors. Dr. Godbey can be reached at jgodbey@jsu.edu.

TERRY MARBUT serves as the department head for Technology and Engineering at Jacksonville State University. He earned B.S. and M.S. degrees in Electrical Engineering from the University of Alabama in Birmingham. He has engineering experience working for Goodyear Tire and Rubber Company and has been teaching at Jacksonville State University for the past twenty six years. Mr. Marbut can be reached at tmarbut@jsu.edu.