

November 2013

Analyzing E.M.B.A. Student Feedback

Ernest A. Capozzoli

Kennesaw State University, ecapozzo@kennesaw.edu

David Gundersen

Stephen F. Austin State University, dgundersen@sfasu.edu

Follow this and additional works at: <https://digitalcommons.kennesaw.edu/jee>

 Part of the [Business Commons](#), and the [Education Commons](#)

Recommended Citation

Capozzoli, Ernest A. and Gundersen, David (2013) "Analyzing E.M.B.A. Student Feedback," *Journal of Executive Education*: Vol. 2 : Iss. 2, Article 5.

Available at: <https://digitalcommons.kennesaw.edu/jee/vol2/iss2/5>

This Article is brought to you for free and open access by DigitalCommons@Kennesaw State University. It has been accepted for inclusion in Journal of Executive Education by an authorized editor of DigitalCommons@Kennesaw State University. For more information, please contact digitalcommons@kennesaw.edu.

Analyzing Student Feedback: An EMBA Perspective

Ernest Capozzoli and David Gundersen

Abstract

The results of an executive Masters of Business Administration (MBA) program assessment are analyzed and interpreted against the backdrop of increasing competition between universities to attract and retain qualified students. The public education environment is currently in a turbulent state. This turbulence is in part caused by such factors as: reductions in public funding for higher education due to constrained state budgets, student expectations, requirements imposed by accrediting bodies, and other outside constituencies requesting more accountability. The pressure to hold education institutions accountable is increasing at a rapid rate. To provide a measure of accountability and quantify program quality, education institutions have placed a great deal of emphasis on program assessment. Informational results from program assessments influence a myriad of decisions made by many that ultimately impact student enrollment, program support and program funding. Despite questionable psychometric properties and potentially conflicting outcomes, student evaluations of teaching faculty continue to be a primary source of information used in program assessments. This burdens educators and administrators with the task of interpreting and utilizing incomplete and perhaps inaccurate information. Results indicate that student response rates decline with increased numbers of evaluations and influences on teaching quality assessments may be unrelated to content and presentation.

Introduction

According to Marsh and Roche (1993), universities have traditionally had students evaluate professor performance to improve course content and structure and for tenure and promotion decisions. Research on the topic generated more than 2000 studies by 1998 (Wilson, 1998) and the literature is rife with inconclusive outcomes associated with student evaluations. Some studies provide general support for the reliability and validity of student evaluation use (Marlin and Gaynor, 1989; Nimmer and Stone, 1991; Scherr and Scherr, 1990; Byrne, 1992; Tagomori and Bishop, 1995). Other studies indicate that student evaluations suffer from design flaws and cannot accurately capture many aspects of teaching effectiveness (Sheenan, 1975; Cashin, 1983; Rodin and Rodin, 1972; Seldin, 1993; Centra, 1993; Green et al, 1998). Despite the inconsistent results in the literature, a study conducted by the Carnegie Foundation for the Advancement of Teaching indicated that approximately 98% of the universities surveyed used some form of student evaluations (Simpson and Siguaw, 2000). Business schools also feel pressure from accrediting bodies such as the American Assoc-

Ernest Capozzoli, Ph.D. is an Assistant Professor of Accounting & Information Systems at Kennesaw State University.

David Gundersen, Ph.D. is a Professor of Management at Stephen F. Austin State University.

lication of Collegiate Schools of Business (AACSB) where 99.3% use some form of evaluation to monitor performance (Comm and Mathaisel, 1998).

Student evaluation data can be used by administrators to plan strategically. Strategically, evaluation data can be used for realignment of university missions and to alter educational offerings. A more crowded and competitive education market is also forcing universities to promote and adopt a customer-driven approach to course offerings. While having great intuitive appeal, the customer-driven approach has shortcomings. Driscoll and Wicks (1998) suggest that a strong marketing orientation may be a potential threat to program quality. They indicate that a marketing approach assumes that student needs and wants are proper to satisfy, that student customers are aware of their needs and can communicate and use them as a basis to make selections. A distinction between customer-led and market-oriented marketing approaches is also required (Slater and Narver, 1998). One problem in the customer-led approach is defining the customer (Bailey and Dangerfield, 2000). Is the customer the student, the taxpayer, the organization that hires the student graduate or some other third party? While acknowledgment that all the preceding may constitute "customers," it is typically the student providing information on teaching performance.

Despite potential shortcomings, a more meaningful question is what happens to the resulting data gathered and how is that data used? In a study by Comm and Mathaisel (1998), 71.8% of schools do not share the results of evaluations beyond the faculty and administration. This calls into question the notion that evaluations are used to satisfy customer demands. Also, the evidence that evaluations are used to improve teaching performance is suspect. Evaluation processes often lack follow-up and quality assurance checks and are often conducted on an annual basis post-facto for data gathering and reporting purposes only (Comm and Mathaisel, 1998).

A research project was initiated to analyze student evaluations associated with an executive MBA program of a large regional university. Response rates and influences on ratings of instruction quality were considered in light of the importance placed on student ratings. A key investigation is to determine whether student evaluations warrant the influence universities place on them for determining program success or failure. In short, are response rates numerous enough to make program decisions based on evaluation results? Do superfluous factors influence student evaluations rendering them less useful for the strategic decisions mentioned earlier?

Program Description

The executive MBA program, at the university where the data was gathered for this study, extends for 18 months. It consists of a kick-off residency session involving 64 hours of contact time over eight days, an international residency requirement of ten days outside the United States, and four courses with approximately 80 hours of contact time. With the exception of the residencies the cohort groups meet once per month on weekends over a four-month period. The courses are designed to be delivered in an integrated format. This format requires two to four hours of instruction, defined as units, across multiple academic disciplines. A typical weekend could have units from accounting, economics, finance, law, leadership, management, marketing, or quantitative areas. The integrated approach is intended to meld and link the various academic disciplines. The unit is delivered by a discipline expert and supported by other discipline experts. All of the integration efforts are supported by a sophisticated course management system.

The course management system uses Lotus Notes with Learning Space. The Lotus system is a very sophisticated application that allows most materials to be delivered electronically to the student. Within Lotus there is

a course schedule and structure listing all of the units to be delivered, a media center that catalogs all electronic material, a course room with multiple message and topic threads, an assessment area where assignments may be given and where feedback instruments are completed for each unit, and a profile area listing all assignments completed and graded as well personal vitas on students and professors. Each unit delivered has a feedback/evaluation instrument available electronically. Lotus compiles the feedback results and makes them available for viewing by the instructor on a real-time basis.

Research Method

Subjects

The subjects participating in the study spanned a period of four years and consisted of six classes of participants totaling 279 individuals in the executive education program leading to the Masters of Business Administration degree at a large regional university. Three classes were comprised solely of employees from a large national employer with class sizes of 38, 25, and 45. The other three classes had employees from a variety of regional employers with class sizes of 52, 59, and 60. Other than class size ($F=1718.585$, $p < .000$) and employers represented in the classes, the two class groups were not demographically or otherwise significantly different. Table 1 presents demographic data on the students.

Executive Education Program Evaluation

As participants begin the executive education program, as described above, they are formed into classes comprised of individuals from one employer or from many employers. Each class is distinct and stays together throughout the entire executive education program. The executive education program includes four

Table 1: Demographic Data

Age Range	27 - 59 years
Average Age (start of program)	35.4 years
Caucasian	69 %
Black	20%
Hispanic	3%
Asian	3%
Mixed-Ethnically	4%
Male	56%
Female	44%
Degrees from different states	35
Degrees from non-U.S. countries	8

distinct courses comprised of different units covering assorted business topics.

The evaluation component of the integrated curriculum is the unit. Unit Instructors in the program provide participants with a standardized statement encouraging unit evaluation for reasons such as improving program quality and providing content responsive to participant needs. Each unit has an electronic feedback/evaluation option where the evaluation can be anonymously recorded within Lotus Learning Space. The feedback is in real-time and can be viewed by the instructor as soon as a student posts it in the system. Procedurally it is asked that the feedback instrument be completed and uploaded within three to five days after the unit is delivered although no negative consequences exist for students who do not evaluate units previously attended.

The total number of possible unit evaluations by a class during the entire life of the program leading to the graduate degree range between ninety-five and one hundred and sixteen. The instrument asks participants to evaluate the unit on a five-point rating scale where 5 equals excellent and 1 equals poor. Exhibit 1 shows an example of a feedback/evaluation instrument.

As can be seen from Exhibit 1 multiple questions are asked however, only question three pertains to instructional effectiveness. The questions one and two address physical plant and catering measures. It is estimated that

completing the three questions would take approximately two-three minutes.

Exhibit 1
Feedback Evaluation Instrument

Question 3: Unit 1.3 Overall Effectiveness
Question: Overall, I would rate this unit, in general as:
(0) Excellent (2) Very Good (3) Good (0) Fair (1) Poor

Analysis

A variety of analyses were conducted to assess training-program evaluation quality and differences in evaluation feedback across different classes. Also, the University conducting the program had specific concerns regarding reduced participation in the training program evaluation process as classes progressed through the executive education program. This concern, which was qualitatively determined by Executive Education Program Administrators eyeballing the data, had not been quantitatively analyzed. Consequently, the reduced participation or casualty rates in evaluating training programs were investigated. Statistical analyses included using correlations, multiple regressions, analysis of variance, and various means tests depending on the specific investigation being conducted. Results of the various investigations are provided in the next section.

Results

Evaluation Response Casualty Trend

The declining response rate by participants as a function of the number of units attended was investigated. A correlation analysis

between the number of units attended and the number evaluations completed by the class was conducted to determine if a relationship could be quantitatively found. This analysis revealed that an inverse relationship ($r^2 = -.439$, $p < .01$) existed between how many units participants attended and how frequently they conducted an evaluation of the unit. This inverse relationship shows that as participants attend more sessions, they less frequently evaluate the unit attended.

This trend was further investigated to determine if the declining evaluation response rate differed by class employer composition. In other words, did company dedicated classes differ from classes where many employers were represented? These additional analyses revealed that this trend with single employer classes ($r^2 = -.753$, $p < .01$) and multiemployer classes ($r^2 = -.758$, $p < .01$) showed the inverse relationship. Participants were less likely to evaluate program units as they progressed through the program and completed more units. The six individual classes were also analyzed separately to see if the declining evaluation response rate trend was evident in every class. Results showed that each individual class experienced the same phenomena with correlations ranging from $r^2 = -.735$, $p < .01$ to $r^2 = -.819$, $p < .01$ across the six classes. A graphical example of one of the classes can be seen in Figure 1 below.

***Sessions Attended
Influencing Evaluation Scores***

Another concern was whether or not participant's evaluations changed as they completed more sessions. With fewer participants responding as more units were completed, the concern was whether this had an influence on actual evaluations made. A regression analysis was conducted to determine if the average evaluation score differed significantly depending on the number of previous units completed. Results indicate that the mean score did vary ($F = 6.094$, $p < .014$) as

Figure 1: Response Rates by Cumulative Sessions Attended

a function of completed units although no clear trend was evident due to the large number of cases in a further means analysis.

Finding that evaluations were influenced by the number of sessions attended triggered additional analyses to determine whether group assignment, class designation, course number, unit, and topic influenced evaluation scores. These analyses revealed that class designation, course number, and topic all influenced average course evaluations.

Type of Class Influencing Evaluation Scores

This investigation looked at whether average evaluation scores varied across the six individual classes. An analysis of variance was used where the average score was the dependent variable and the class was the independent variable. Results showed that the

average score did vary by class ($F = 4.508, p < .000$) indicating that classes evaluated the units differently. The range of average scores across the six classes was relatively small however, where the lowest average evaluation was 4.13 and the highest was 4.31.

Variations of evaluation scores within a class were also noted. Higher standard deviations reflect more variability within a class indicating differences in perceptions of quality held by participants. The range of standard deviations across the six classes was from .316 to .431.

Course Designation Influencing Evaluation Scores

Course designation was also considered as an influence on average evaluations conducted by participants. An analysis of variance was used where the average score was the dependent variable and the course was the independent variable. Results showed that the

average score did vary by course ($F = 3.971, p < .008$) indicating that the four courses evaluated resulted in different unit scores. The range of average scores across the four courses also relatively small where the lowest average evaluation was 4.19 and the highest was 4.32.

Variations of evaluation scores within a course were also noted. Higher standard deviations reflect more variability within a course indicating differences in perceptions of quality held by participants. The range of standard deviations across the four courses was from .334 to .423.

Topic Influencing Evaluation Scores

The topic of each unit was also considered as an influence on average evaluations conducted by participants. The units were categorized into one of eight topic disciplines for analysis: Accounting (A), Economics (E), Finance (F), Leadership (L), Marketing (M), Quantitative (Q), Strategy (S) and Other (O). An analysis of variance was used where the average score was the dependent variable and the topic was the independent variable. Results showed that the average score did vary by topic ($F = 5.525, p < .000$) indicating that the eight topic disciplines evaluated did produce different unit scores. The range of average scores across the eight topics showed the lowest average evaluation was 3.90 and the highest was 4.36.

A point of interest in this analysis is that two topics, designated L and O, received far more participant evaluations with 130 and 143 respectively. Other topics designated A, E, F, M, Q, and S totaled far fewer total evaluations where the range of participants evaluating was from 45 to 84. Topics L and O had characteristics that motivated participants to evaluate these sessions at a higher rate. This is a particularly interesting finding given that L was ranked 6th and O was ranked 4th in the mean score calculations. Table 2 presents the mean scores by topic discipline.

Variations of evaluation scores within a topic were also noted. Higher standard deviations reflect more variability within a topic indicating differences in perceptions of quality held by participants. The range of standard deviations across the eight topics was from .324 to .422. Topics L and O that were discussed earlier had standard deviations of .398 and .350 and did not indicate anything extraordinary.

Table 2
Mean Evaluation Scores by Discipline

Discipline	Mean Score
Economics (E)	4.360
Finance (F)	4.334
Accounting (A)	4.288
Other (O)	4.250
Marketing (M)	4.215
Leadership (L)	4.210
Strategy (S)	4.154
Quantitative (Q)	3.900

Discussion

In explaining the declining response rate trend, it must be recognized that as in all graduate programs, time is a precious commodity. The respondent group used in this study all had full-time demanding jobs in addition to a substantial academic workload that required much out of classroom work. Job, personal and academic time demands are balanced against time available. As the declining response trend suggests, students prioritized available time and providing meaningful feedback on every unit was not considered a high priority. This view is especially compelling when one considers that no negative consequences occur for failing to complete unit evaluations.

Despite this potential explanation, a note of caution must be acknowledged by university administrators who offer student ratings as the basis for faculty evaluations, curriculum changes,

and graduate program direction. Should student evaluations warrant the influence universities place on them for determining program success or failure? Are response rates numerous enough to make program decisions based on evaluation results? With so few evaluations, do superfluous factors influence results that render them useless for the strategic decisions mentioned earlier? These questions can only be addressed by individual administrations. It is noted however that further research is needed to determine whether results from this single university sample can be generalized to a larger population of graduate programs.

In concert with declining response rates is the finding that as more units were presented overall scores were affected. With no discernable trend identified it is difficult to speculate on why this occurred. One possible explanation is that as fewer participants respond each response carries more influence. In short, a progressively smaller respondent group may undermine the quality of total responses as represented by the overall scores. This notion supports the concern mentioned previously regarding how evaluation scores are used by administrators.

As for class differences, class dynamics, class composition, or other unknown factors may have contributed to differing outcomes for the same unit. For this study, faculty composition was stable in that units, for the most part, were taught by the same instructors. Unit content was also viewed by the investigators as similar from class to class. Presentation delivery was consistent in terms of both multimedia and network support. A final perspective with a diminishing respondent pool is that classroom dynamics became more influential in overall scores. Did respondents with similar perspectives or predispositions continue to rate units as others dropped away from rating?

Finally, evaluation scores did vary across the eight different topics as presented in Table 2. Not surprisingly, some topics and some faculty received better or worse evaluations. The leadership and other categories had the most

responses due primarily to the fact that more units incorporated these categories. This result suggests that unit topic and individual instructor delivery had an impact on evaluation scores.

Summary

This paper discussed the qualities, characteristics and consequences associated with student feedback and presented the findings of an analysis of evaluation and feedback efforts of an Executive MBA program. The findings indicate that as students were exposed to extensive feedback requirements their participation levels fell substantially. In a traditional MBA program a student performs one evaluation per course. The sheer volume of evaluations in an integrated program, (95 to 115 units over 18 weekends) requires substantially more effort than a traditional MBA program.

In an integrated MBA program where subject matter is woven from several disciplines to deliver a course, it would seem intuitive that specific individual topic and faculty performance data be gathered. However, the results indicated that the intent of gathering data at a very specific level has had the opposite effect and reduced overall feedback participation. This finding would suggest that caution should be taken when requiring extensive feedback at a very granular level. The decline in evaluation completion calls into question the very purpose of the evaluations.

Should evaluations be a multipurpose tool? Can that tool be simultaneously used to evaluate teaching effectiveness and curriculum requirements? This study described the results of an intensive feedback and evaluation process.

Study Limitations and Areas for Further Study

Caution should be exercised in the interpretation of this study. EMBA programs by their nature can have a high degree of variability from program to program and would therefore limit the generalizability of the findings to other programs. The study also highlights areas for

further study. Further examination of the motivations of faculty and administrators in the use of the existing evaluation system should be performed to validate the purpose and use of the current evaluation system. Further investigation of past and current participants in the program should be conducted to determine their motives and explanations or for such a dramatic drop in evaluation levels, and to solicit suggestions on how to improve the overall feedback and evaluation system.

References

- Bailey, Jeffrey J., Dangerfield, Byron, (2000), Applying the distinction between market-oriented and customer-led strategic perspectives to business school strategy," Journal of Education for Business, Jan/Feb, Vol. 75, Issue 3
- Byrne, C. J., (1992), "Validity studies of teacher rating instruments: Design and interpretation," Research in Education, 48, (November), pp: 42-54.
- Cashin, W. E., (1983), Concerns about using student ratings in community college. In Evaluating Faculty and Staff. New Directions for Community Colleges, edited by A. Smith. San Francisco, CA: Jossey-Bass.
- Centra, J. A., (1993), Reflective Faculty Evaluation, San Francisco, CA: JosseyBass.
- Comm, Clare L., and Mathaisel, Dennis F. X., (1998), "Evaluating teaching effectiveness in America's business schools: Implications for service marketers. Journal of Professional Services Marketing Quarterly, 16, (2), pp: 163-70.
- Driscoll, Cathy, Wicks, David, (1998), "The Customer-Driven Approach in Business Education: A Possible Danger?, Journal of Education for Business, Sep/Oct98, Vol. 74, Issue 1.
- Green, B. P., Caleron, T. G., and Reider B. P., (1998), "A content analysis of teaching evaluation instruments used in accounting departments," Issues in Accounting Education, 13, pp: 15-30.
- Marlin, J. E., Jr., and Gaynor, P., (1989), "Do anticipated grades affect student evaluations? A discriminant analysis approach," College Student Journal, 23, (Summer), pp: 184-192.
- Marsh, H. W., and Roche, L., (1993), "The use of students' evaluations and an individually structured intervention to enhance university teaching effectiveness," American Educational Research Journal, 30, (Spring), pp: 217-251.
- Nimmer, J. G., and Stone, E. F, (1991), "Effects of grading practices and time of rating on student ratings of faculty performance and student learning," Research in Higher Education, 32, (April), pp: 195-215.
- Rodin, M. and Rodin, B., (1972), "Student evaluations of teachers: students rate most highly instructors from whom they learn least," Science 177, September, pp: 1164-1166.
- Scherr, F., and Scherr, S. S., (1990), "Bias in student evaluations of teacher effectiveness," Journal of Education for Business, 65, (May), pp: 356-358.
- Seldin, P., (1993), "The use and abuse of student ratings of professors," The Chronicle of Higher Education, (July), pp: A40.
- Sheenan, D.S., (1975), "On the invalidity of student ratings for administrative personnel decisions," Journal of Higher Education, Vol. 46, No. 6, pp: 687-700.
- Slater, S. F., and Narver, J. C., (1998), "Customer-led and market-oriented: Let's not confuse the two," Strategic Management Journal, 19, pp:1001-1006.

Simpson, Penny M., and Siguaw, Judy A, (2000), "Student evaluations of teaching: An exploratory study of the faculty response," Journal of Marketing Education, December, Vol. 22, Iss. 3; pp: 199-113.

Tagomori, H. T., and Bishop, L. A., (1995), "Student evaluation of teaching: Flaws in the instruments," Thought & Actions: The NEA Higher Education Journal, 11, pp: 63.

Wilson, Robin, (1998), "New research casts doubt on value of student evaluations of professors," The Chronicle of Higher Education, 44, (19), pp: A12-A14.