Integrated Evaluation of Teaching Effectiveness: A Case Study*

YASSER E. IBRAHIM
Engineering Management Department, Prince Sultan University, Saudi Arabia. E-mail: ymansour@psu.edu.sa
On leave: Structural Engineering Department, Zagazig University, Egypt.

Student evaluation of teaching is considered one of the main elements in assessing the quality and effectiveness of teaching of a faculty member. In most cases, the student evaluation is conducted at the end of the semester to assess all aspects of the course in terms of performance of the instructor and adequacy of the teaching resources. However, student evaluation of a course at the end of the semester may not necessarily reflect the actual performance of the instructor, as it may be affected by other factors such as the grading style of the instructor and the grade expected by the student upon completing the course. In this regard, there is a need to better assess the teaching effectiveness, especially in cases where student evaluation is used as the primary criterion for assessing the faculty member. In this research, an approach was adopted to assess the teaching effectiveness whereby the student evaluation at the end of the course is complimented by another one mid-way as well as classroom observation conducted by the department chair. This approach was tested on thirteen undergraduate courses in the Engineering Management program for different program years. A total number of 346 students participated in the evaluation of teaching effectiveness. The results obtained from the three evaluations mentioned above were compared and the grades of the students in these courses were considered to investigate their effect on the students' evaluations. The integrated evaluation that combines and averages the three evaluations minimizes the misleading effect of traditional SET, normally conducted at the end of the course, which is sometimes affected by the grades expected by students. The whole process gives a clear picture and appears promising as a comprehensive assessment tool of teaching effectiveness.

Keywords: teaching effectiveness; class observation; students evaluation; assessment

1. Introduction

There is a need to assess the performance of faculty members when considering administrative decisions such as contract renewal and promotion. In such cases, many aspects are normally considered including the teaching quality and performance, research, community services and professional development. The most important factor among these items is the student evaluation of teaching (SET), which is conducted at the end of the semester for all courses taught by the instructor. SET is considered a valid and reliable tool to assess the performance of instructors [1].

Currently, SET is extensively used in higher education [2, 3]. There is usually ongoing debate about how reliable the SET is to assess the pedagogical performance of the instructor. Some researchers realized that the quantitative nature of the SET results and the ease with which they are obtained led to relying on them as the best measure of the teaching effectiveness [4–8]. However, others demonstrated that there are many issues that may raise concerns around the SET role and value. Many researchers noted that the SET results may be affected by factors that are unrelated to the teaching effectiveness, such as the final grade that the student expects, being reported by the instructor for cheating or plagiarism, the instructor's accent, the rele-vance of the course to the student's major, the student motivation, the instructor's gender, class time, class size and class level [8–10].

Students usually prefer the instructor with a teaching method that is relevant or that suits his/her favorable way of learning. For example, passive learners do not prefer non-traditional methods in teaching, because it challenges their desire of learning just by receiving without any active participation [11]. Students may negatively respond to innovative techniques used by instructors who rely on new approaches of experimentation and trial and error concepts. In fact, relying solely on SET may unfairly affect the career of instructors, which may negatively affect the presumably healthy relationship between the instructor and students over time. Accordingly, different strategies are needed to benefit from student voices in an effective way [12], especially during the course of the semester rather than at its conclusion. The significance of the SET from the students' perspective is important, as they evaluate their instructors. Normally the students who value SET procedures tend to provide higher SET scores [13].

The instructor's attitude and psychology may affect the overall SET. When the course is evaluated at the end of the semester, the way that the instructor deals with students (based on his/her character) plays an important role in the outcomes of the SET.

Many researchers attempted to define the noticeable characteristics of good instructors from the students' perspective. Anderson et al. [12] stated that doctoral students reported that the most important characteristics of instructors are being responsive, student-centered, ethical, professional and enthusiastic. Al-Mohaimeed and Khan [14] reported that medicine students showed their appreciation and admirations for those instructors who show respect to their students, demonstrate expertise in the subject, understand/relate to students and apply good communication skills. On the other hand, having a sense of humor, giving high grades, sharing personal experience and dressing up appropriately were least valued by the students. Xiao and Dyson [15] reported that Chinese students considered the traits of: knowledge, responsibility, effective teaching and encouragement and facilitation of independent, critical thinking to be the most valuable characteristics of accounting instructors.

The perception of good teaching differs among disciplines. In fields with difficult applications, such as engineering, good teaching is perceived through knowledge application and integration [16, 17]. On the other hand, in fields with soft applications, such as social work, stress protocols and procedures are the most valued items for good teaching [18, 19].

A study conducted at a graduate college of education [20] revealed that the year of study in a degree program affects the perception of good teaching from the students' perspective. First-year students considered instruction style to be the most important aspect of good teaching; conversely, the personal relationship between the instructor and the students was the most important for second-year students. Older and more mature students value the long-term student development more than younger students do. Gender, age and the type of the higher education institute all affect the students' perception of good teaching [21].

Peer observation of classrooms is an efficient way to measure the effectiveness of teaching. It can be used besides the SET in order to provide comprehensive assessment of teaching. Peer observation plays an important role in enhancing the teaching through reflections and thoughts of the observer, who should have ample knowledge and experience to play this role. Evaluation of the peer observer should not be rendered personal or subjective and no conflict of interest should be assured [22, 23].

In this research, an integral teaching evaluation process is applied considering 13 undergraduate courses in the Engineering Management Curriculum. The process included three different evaluation items considering (1, 2) SET and (3) peer observation. First, a mid-way SET was conducted using a special survey. At the same time, peer observation was conducted by the department chair using another special survey. Both surveys were developed and adopted by the department. Both the feedback from the mid-way SET and the peer observation were sent to the instructor for his consideration in order to enhance the teaching and learning process. Another SET was conducted for the same courses at the end of the semester using the university standard survey. In this study, the results of the two SETs and peer observation were analyzed alongside the grades achieved by the students in these courses to investigate their effect on SET.

2. Methodology

The common practice of evaluating teaching is to seek students' feedback about their courses at some time during the last two weeks of classes. Typical evaluations include many items and aspects of teaching, such as the clarity of syllabus, assessment methods and course objectives from the beginning, availability of resources, availability of the instructor during office hours, commitment of the instructor in his teaching, class times, use of technology, the enthusiasm of the instructor and the way he/she treats students. Other items may affect the valuation of the instructor, including the relevance of the course to the student's major, expected grade by the student, effort exerted from the student during the course compared to other courses and the innovative techniques used in teaching. Passive students may prefer classical teaching methods that rely on one-way lecturing from the instructor to students and do not like to be involved in teaching/learning processes using small groups or flipped class techniques. Appendix A shows the form used for final SET at the end of courses under consideration in this study.

For these reasons, there is always a need to incorporate other methods of evaluation besides students' feedback such as peer observations, whereby the instructor is evaluated by another faculty member to evaluate aspects not considered or not effectively evaluated by the students. To this end, a teaching observation form was adapted by the Engineering Management Department from the Peer Observation Guidelines and Recommendations for developed initially by the University of Minnesota [24]. The form is presented in Appendix B. The evaluation consists of three main sections: content of the lecture, the delivery of the lecture and the environment of the classroom. The evaluation has fourteen items covering these three sections. The overall evaluation is the average score of these fourteen items. After the classroom observation, the form is filled with a score for each item besides comments addressing the strength and weakness

regarding the teaching effectiveness with some recommendations for enhancement. Another form was developed by the department to seek students' feedback at the mid-way of the course in parallel with the class observation.

The evaluation includes basic and brief teaching aspects such as the pace the material is delivered, fairness in assessment, encouragement to participate in class activities, conciseness of the explanations of the instructor and availability of faculty member during class hours. The students' average is marked for each question. The overall score is then calculated by averaging the average score of each question. The evaluation includes some open-ended questions about the things the student likes most in the course and the improvement areas the instructor need to consider in order to improve performance and any other comments. The instructor receives the overall score along with the summary of students' comments. Appendix C shows the mid-way evaluation form used. By this mid-way evaluation form, instructor can get the feedback of the students in the 7th week and try to address their comments and suggestions to improve his teaching, which may be reflected in his final traditional evaluation at the end of the course. The open-ended questions can open the door for students to provide innovative and creative ideas to the instructor to enhance the teaching style and effectiveness to fit students' style and way of thinking and understanding.

As mentioned earlier, 13 courses were selected for different instructors in several program years from the first to the fourth, covering the full engineering management undergraduate curriculum. The relationship between the SET and the marks received by the students is presented and discussed herein to investigate if the high marks correlate with the course evaluation. Also, the three assessment items are compared to each other in order to provide a more realistic view on the performance of the faculty member, which is used in his annual evaluation that plays a significant role in his contract renewal.

3. Results

Table 1 shows the course level and number of students who filled the mid-way and final students' evaluation forms.

Table 2 presents the results of the students at the end of the course and the grade obtained. A weight factor was given to each grade in order to get the grade average of the course out of 5.0. Grade "A" is given 5.0, grade "B" is given 4.0, grade "C" is given 3.0, grade "D" is given 2.0 and grade "F" is given 1.0. The grade average, GA, is calculated by the following equation:

$$GA = \frac{\sum n * w}{N} \qquad (1)$$

Where n is the number of students who got a certain grade, w is the weight factor of this grade; N is the total number of students registered this course.

Table 3 presents results of both the SET of each course and the chair's observation of the faculty performance in the classroom. It shows also the grade average of each course calculated from the final grades of students. Fig. 1 shows the relationship between the SET and the chair observation when conducted at the half-way of the course. Fig. 2 shows the relationship between the SET at the end of the course and the grade average. Fig. 3 presents the relationship between the SET at mid-way and at the end of the course and the chair observation.

Careful analysis of the results presented in Fig. 1 to Fig. 3 leads to the following observations:

- There is a good agreement between the results of SET conducted at the mid-way of the course and the parallel chair observation. Although the nature of the two surveys is different, both evaluations provided close results, which increases the confidence in the evaluation outcomes. Another important observation is the absence of the effect of the students' grade, an outcome most welcomed by teachers.
- Although the grades that the students scored had some effect on the final SET of some courses, there is no clear or constant trend between the final SET and the students' grades. For example, courses number 5 received high SET compared to chair observation while the average grade was considerably high (4.22 out of 5.0). On the other hand, courses number 3 and 4 received chair observation higher than SET and the average grade was 3.66 and 3.53, respectively, which are considered relatively low compared to other

**Table 1.** Course level and number of students involved in the surveys

Course No.	Year	No. of students surveyed	
		Mid-way	Final
1	First	22	37
2	Second	23	28
3	Second	26	29
4	Second	17	30
5	Second	18	36
6	Third	26	26
7	Third	31	33
8	Third	11	16
9	Third	28	29
10	Fourth	11	12
11	Fourth	15	17
12	Fourth	24	30
13	Fourth	12	23

Table 2. Students' grades and grade average

| Course No. | No. of students achieving the final grade | | | | | Grade Average (Out of 5) |
	A	B	C	D	F	
1	11	15	11	1	2	3.80
2	8	9	18	11	3	3.16
3	5	11	12	0	1	3.66
4	12	18	16	10	1	3.53
5	19	14	7	0	1	4.22
6	1	7	9	9	1	2.93
7	7	0	26	0	0	3.42
8	4	6	7	0	0	3.82
9	4	8	11	7	1	3.23
10	4	6	2	1	0	4.00
11	3	12	3	0	0	4.00
12	10	5	8	11	0	3.41
13	16	7	0	0	0	4.70

Table 3. Results of SET and chair observations

| Course No. | SET (Out of 5) | | Chair Observation (Out of 5) | Grade Average (Out of 5) | Overall SET | Integrated Evaluation | % Difference from Final SET |
	Mid-way	Final					
1	4.64	4.25	4.5	3.80	4.45	4.46	5.02
2	4.3	4.34	4.36	3.16	4.32	4.33	0.15
3	4.4	4.36	4.79	3.66	4.38	4.52	3.59
4	3.8	3.82	4.5	3.53	3.81	4.04	5.76
5	4.5	4.53	4.21	4.22	4.52	4.41	2.58
6	3.7	3.78	3.86	2.93	3.74	3.78	0.00
7	4.3	3.7	4.43	3.42	4.00	4.14	11.98
8	3.7	4.32	4.07	3.82	4.01	4.03	6.71
9	4.1	3.88	4.36	3.23	3.99	4.11	6.01
10	4.9	4.44	4.43	4.00	4.67	4.59	3.38
11	4.6	4.67	4.14	4.00	4.64	4.47	4.28
12	3.4	3.42	3.36	3.41	3.41	3.39	0.78
13	4.7	4.67	4.57	4.70	4.69	4.65	0.50
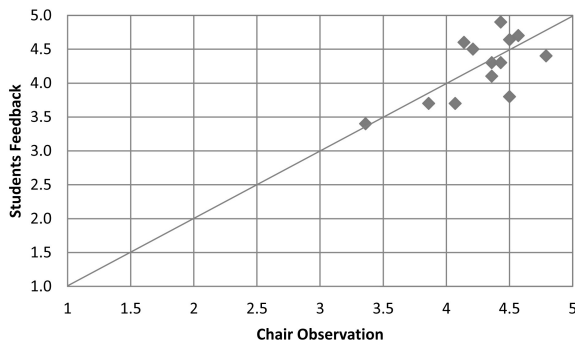
Fig. 1. Relationship between mid-way survey and chair observation (Conducted at same time).
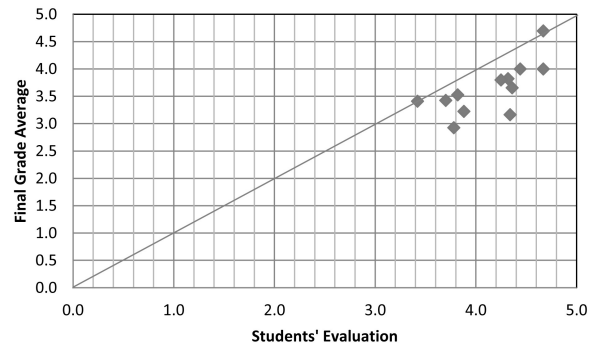
Fig. 2. Relationship between the SET at the end of the course and the grade average.

courses. It is worth mentioning that the courses that had distinct effect of students' grade on the SET were in the early stage of the curriculum (second year). This indicates the effect of student level in the curriculum on their evaluation. Apparently, senior students care more about the delivery of the course than their grades.

● In some courses, it was noticed that the mid-way SET and chair observation were very close while the final SET was considerably low, which are obvious in course number 7 in the third year. This

is also can be related to the low grade average obtained (3.42 out of 4.0)

● Integrating the three evaluations presented above can form a better platform towards a more comprehensive assessment of the instructor and can be more reflective of the teaching effectiveness; this can help avoid the misleading effect of the grades on the SET. This integrated method is recommended, especially if the contract renewal and promotion of the faculty member depend heavily on the SET. The integrated evaluation is
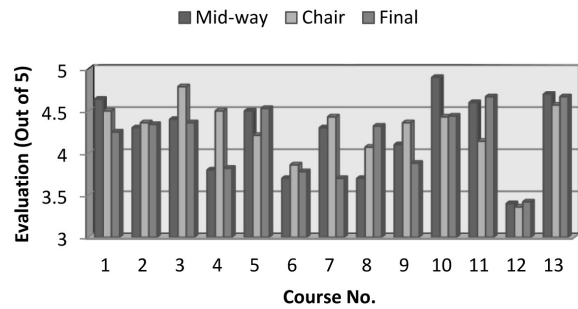
Fig. 3. Relationship between the SET at mid-way and at the end of the course and the chair observation.

calculated by averaging the three evaluations; mid, final and chair observations as shown in Table 3. The integrated evaluation differs from the traditional SET value by around 12% in one of the courses. Although this course didn't experience the lowest grade average (3.42 out of 5.0) but it was somehow affected by the relatively low grade average. Table 3 shows also the overall SET, which is calculated by averaging the traditional SET at the end of the course and the mid-evaluation. This overall SET gives more rational evaluation. The difference between the SET value between the mid and final evaluation reached around 12% in one of the courses.

4. Conclusion

Student evaluation of teaching, SET, is a tool used for assessing the teaching effectiveness and instructor's evaluation for performance assessment. An integrated approach for the valuation of the teaching effectiveness is introduced for better assessment. The approach considers the SET at mid-way and at the end of the course as well as peer observation using well designed surveys. The study is based on thirteen undergraduate courses in the Engineering Management Department. The results showed the effectiveness of the proposed approach and how it can avoid the misleading effect of the final students' grades. Moreover, the mid-way evaluation is considered as preliminary assessment of the teaching effectiveness and it helps the instructor to take corrective steps to fix the weak aspects that are early detected and raised by students.

Based on this approach and results obtained, a mathematical model can be developed to capture the effect of grades obtained by students and hence corrections factors can be obtained in order to adjust the evaluation of the instructors conduced at the end of the course, as being practiced everywhere in higher education, so that the effect of high grades and level of the course is eliminated or minimized. This can be further investigated through

future research, however it needs more statistical data through large number of courses in order to attain high-accuracy and reliable mathematical model.

References

1. P. Hooper and J. Page, Measuring teaching effectiveness by student evaluation, *Issues in Accounting Education*, 1986, pp. 56–64.
2. S. Darwin, What contemporary work are student ratings actually doing in higher education? *Studies in Educational Evaluation*, **54**, 2016, pp. 13–21.
3. J. D. Newton, Using student evaluation of teaching in administrative control: the validity problem, *Journal of Accounting Education*, **6**(1), 1988, pp. 1–14.
4. B. L. Steward, S. K. Mickelson and T. J. Brumm, Continuous Engineering Course Improvement through Synergistic use of Multiple Assessment, *International Journal of Engineering Education*, **21**(2), 2005, pp. 277–287.
5. K. A. Villanueva, S. A. Brown, N. P. Pitterson, D. S. Hurwitz and A. Sitomer, Teaching Evaluation Practices in Engineering Programs: Current Approaches and Usefulness, *International Journal of Engineering Education*, **33**(4), 2017, pp. 1317–1334.
6. D. Chalmers. A review of Australian and international quality systems and indicators of learning and teaching. Strawberry Hills: The Carrick Institute for Learning and Teaching in Higher Education, 2007.
7. A. C. M. Moskal, S. J. Stein and C. Golding, Can you increase teacher engagement with evaluation simply by improving the evaluation system? *Assessment and Evaluation in Higher Education*, **41**(2), 2017, pp. 286–300.
8. P. Spooren and W. Christiaens, I liked your course because I believe in (the power of) student evaluations of teaching (SET). Students' perceptions of a teaching evaluation process and their relationships with SET scores, *Studies in Educational Evaluation*, **54**, 2017, pp. 43–49.
9. P. A. Koushki, A. Christoforou, A. M. Larkin and A. Al-roomi, Engineering Students' Attitudes towards Teaching and Teachers, *International Journal of Engineering Education*, **21**(3), 2005, pp. 424–433.
10. H. G. Murray, Student evaluation of teaching: Has it made a difference? Annual meeting of the society for teaching and learning in higher education, Charlottetown, Prince Edward Island, 2005.
11. P. Spooren, B. Brockx and D. Mortelmans. On the validity of student evaluation of teaching the state of the art, *Review of Educational Research*, **83**(4), 2013, pp. 598–642.
12. M. Anderson, J. M. Ingram, B. J. Buford, R. Rosli, M. L. Bledsoe and A. J. Onwuegbuzie, Doctoral students' perceptions of characteristics of effective college teachers: A mixed analysis, *International Journal of Doctoral Studies*, **7**, 2012, pp. 279–309.
13. D. Kember and A. Wong, Implications for evaluation from a study of students' perceptions of good and poor teaching, *A. Higher Education*, **40**(1), 2000, pp. 69–97.
14. A. A. Al-Mohaimeed and N. Z. Khan, Perceptions of Saudi medical students on the qualities of effective teachers: A cross-sectional study, *Saudi Medical Journal*, **35**(2), 2014, pp. 183–188.
15. J. Z. Xiao, and J. R. Dyson, Chinese students' perceptions of good accounting teaching, *Accounting Education*, **8**, 1999, pp. 341–361.
16. R. Neumann, Disciplinary differences and university teaching, *Studies in Higher Education*, **26**, 2011, pp. 135–146.
17. R. Neumann, S. Parry and T. Becher, Teaching and learning in their disciplinary context, *Studies in Higher Education*, **27**, 2002, pp. 405–417.
18. T. Becher, Quality assurance and disciplinary differences, *Australian Universities' Review*, **37**(1), 1994, pp. 4–7.
19. T. Becher and P. Trowler, Academic tribes and territories: Intellectual enquiry and the culture of disciplines, 2nd ed.

Ballmoor and Phil: The Society for Research into Higher Education & Open University Press, 2001.

20. M. Miron and M. Mevorach, The good professor as perceived by experienced teachers who are graduate students, *Journal of Education and Training Studies*, **2**(3), 2014, pp. 82–87.

21. F. N. Alhija, Teaching in higher education: Good teaching through students' lens, *Studies in Educational Evaluation*, **54**, 2017, pp. 4–12.

22. R. Berk, Survey of 12 Strategies to Measure Teaching Effectiveness, *International Journal of Teaching and Learning in Higher Education*, **17**(1), 2005, pp. 48–62.

23. Z. S. Siddiqui, C. Diana and E. Sandra, Twelve tips for peer observation of teaching. *Med. Teach.*, **29**, 2007, pp. 297–300.

24. University of Minnesota, Peer Observation Guidelines and Recommendations, Office of Human Resources, 2013: http://www1.umn.edu/ohr/teachlearn/resources/peer/guidelines/

Yasser E. Ibrahim is an associate professor and chairman of the department of Engineering Management at Prince Sultan University since 2015. He served as the chairman of committee for the teaching award of College of Engineering for last two years. He has several conference and journal publications in the area of engineering education. His other research area includes seismic resistant structures and structural response under different dynamic loads. He got his PhD in Civil and Environmental Engineering from Virginia Tech. in 2005.

Appendix A: Final Student Evaluation

A. Instructor:	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1. Was well prepared for class					
2. Made effective use of class time					
3. Explained concepts and ideas clearly					
4. Answered questions in helpful ways					
5. Willing to meet students out of class					
6. Has knowledge of the course contents					
7. Was able to explain difficult concepts/topics in easy way					
8. Exposed us to most recent development of the course					
9. Encourages to explore the content of the course beyond what is required of the text books requirements					
10. Overall, I was satisfied with the quality of my instructor					
B. Course Objectives:	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
11. Course Objectives are stated at the beginning of the term					
12. Course Objectives are achieved at the end of the term					
13. I was able to achieve the course learning outcomes					
C. Course Work (exams, quizzes, projects, presentations, etc.)	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
14. The actual course work meet the Course Objectives					
15. The amount of work I am expected to do to achieve the Course Objectives					
16. The time allocated to complete the course work is suitable					
D. Course Learning Environment	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
17. I enjoy learning together with my friends in this course					
18. I am satisfied with the overall learning resources (e.g., course materials, books, learning aids) provided to support my learning activities					
E. Course Assessment	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
19. Different variety of assessment (exams, quizzes, projects, presentations, etc.) was employed in the course					
20. Grades assigned are based on my performance in the course					
21. Overall, I was satisfied with the quality of this course.					

Appendix B: Peer Observation

Instructor	**Dr.**	Date
Course	**EM**	1 = Strongly Disagree, 2 = Disagree, 3 = Neither Agree or Disagree, 4 = Agree, 5 = Strongly Agree

1 2 3 4 5 NA   Comments

Content		
	1	Clearly articulates the objectives of the class.
	2	Provides the significance/importance of the information to be learned.
	3	Provides context—links to previous and future topics, classes, and courses.
	4	Provides concrete, real life, practical examples.

Delivery		
	1	Effectively uses a variety of instructional techniques.
	2	Provides clear explanations.
	3	Uses a variety of means to get feedback on the level of student understanding.
	4	Presents the material at a good pace and depth/level.
	5	Speaks clearly—volume, tone, grammar.

Environment		
	1	Begins and ends class promptly and in a well-organized manner.
	2	Holds the attention of, and has the respect, of the students.
	3	Practices effective time and classroom management.
	4	Creates a class environment that invites learning and encourages interaction.
	5	Treats students with respect.

Comment

Observer and Evaluator

Developed by EM Chair: Prof. Allen Greenwood, Ph.D.,   P.E. 12 Oct 2014

Adapted from: Peer Observation Guidelines and Recommendations, University of Minnesota, [24].

Appendix C: Mid-way Student Evaluation

- Your anonymous feedback will help all of us improve the education provided by PSU.
- All responses are **confidential**. The forms will be maintained by the Chairman and faculty will receive summarized results, not the individual forms.

Part A: Please score the following based on:

1 = "Strongly Disagree", 2= "Disagree", 3 = "Neither Agree or Disagree", 4 = "Agree", and 5 = "Strongly Agree"

	1	2	3	4	5
1. The material is being taught at an appropriate pace.	1	2	3	4	5
2. Grades for assignments and exams are provided in a timely manner.	1	2	3	4	5
3. Assignments, exams, etc. are graded fairly.	1	2	3	4	5
4. The instructor encourages in-class participation.	1	2	3	4	5
5. The instructor offers help outside of class time.	1	2	3	4	5
6. The instructor speaks clearly and audibly.	1	2	3	4	5
7. The instructor provides clear and concise explanations.	1	2	3	4	5

Part B: Short Answer.

1. How much time outside of class do you spend working on this course?
2. Identify one or two specific things that you like about this course.
3. Describe one or two specific things that the instructor could do to improve student learning in this course.
4. Other Comments.