

AN ABSTRACT OF THE THESIS OF

Keisha A. Villanueva for the degree of Master of Science in Civil Engineering presented on September 14, 2016.

Title: Investigating the Current Teaching Evaluation Practices used in Engineering Programs and their Usefulness to Educators

Abstract approved:

Shane A. Brown

Evaluation of teaching is used to examine and improve teaching quality in higher education. The literature consists of best teaching evaluation practices that have been developed, yet there is little research that has been conducted which investigates teaching evaluation practices in engineering programs. This thesis represents two manuscripts that explore the current teaching evaluation practices that are used in engineering programs at three institutions using exploratory sequential mixed-method design. The data were collected by interviewing and surveying thirty-four educators including course instructors, department heads, and program coordinators. The results showed that there are six approaches to evaluate teaching currently used in engineering programs. These approaches are utilized either for formative or summative evaluations to determine an educators' effectiveness. Participant's perspectives were also examined to identify teaching evaluation practices that were found to be useful to educators. A major finding of this study is that current teaching evaluation practices in engineering programs do not align with evidence-based teaching practices despite a substantial interest in improving teaching evaluation practices. It is important for educators to identify and adopt best teaching evaluation practices that have been shown to enhance teaching effectiveness and student learning.

©Copyright by Keisha A. Villanueva
September 14, 2016
All Rights Reserved

Investigating the Current Teaching Evaluation Practices used in Engineering Programs
and their Usefulness to Educators

by
Keisha A. Villanueva

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Presented September 14, 2016
Commencement June 2017

Master of Science thesis of Keisha A. Villanueva presented on September 14, 2016

APPROVED:

Major Professor, representing Civil Engineering

Head of the School of Civil and Construction Engineering

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Keisha A. Villanueva, Author

ACKNOWLEDGEMENTS

The author expresses sincere appreciation to Dr. Shane A. Brown for his guidance and expertise in assisting to complete this research. I will be forever grateful for giving me this wonderful opportunity to earn my degree. You have been a great, and supportive advisor this entire process. I would also like to thank my graduate committees (Dr. Nicole P. Pitterson, Dr. Ann Sitomer and Dr. David S. Hurwitz) for their help and guidance.

I would like to express my deep gratitude to my parents (Paul, Ronald and Gemma), brother (T.J.) and boyfriend (Dean). They were always there to support and encourage me. I dedicate all my hard work to you.

I thank my entire engineering education research group and all the transportation graduate students in the office for making my time as a graduate student full of fun and excitement. Thank you to the National Science Foundation (NSF) for funding this research.

CONTRIBUTION OF AUTHORS

Keisha A. Villanueva performed most of the qualitative analysis of the interviews and survey used in this research, and wrote the majority of the contents within this thesis. Dr. Shane A. Brown assisted in writing, organizing and editing both the manuscript presented in chapter 2 and 3. Dr. Ann Sitomer assisted in writing the literature review presented in chapter 3 and editing of both manuscript in chapter 2 and 3. Dr. Nicole P. Pitterson assisted in writing and editing of both manuscript in chapter 2 and 3. Dr. David S. Hurwitz assisted in editing the manuscript in chapter 2.

TABLE OF CONTENTS

	<u>Page</u>
CHAPTER 1 INTRODUCTION	1
1.1 Research context	1
1.2 Research questions	2
1.3 Organization of the thesis	3
1.4 Description of research	3
1.5 Terminology.....	4
CHAPTER 2 TEACHING EVALUATION PRACTICES IN ENGINEERING PROGRAMS: CURRENT APPROACHES AND USEFULNESS	6
2.1 Abstract	7
2.2 Introduction	8
2.3 Background	9
2.3.1 What practices are used?.....	10
2.3.2 Formative and summative teaching evaluation practices	12
2.3.3 Call for improvements to teaching evaluation practices.....	13
2.3.4 What practices work for engineering educators?.....	15
2.4 Method	17
2.4.1 Study design.....	17
2.4.2 Participants characteristics.....	17
2.4.3 Data collection	18
2.4.3.1 Semi-structured interviews	19
2.4.3.2 Qualtrics survey	20
2.4.4 Data analysis	22
2.5 Results	23
2.5.1 What teaching evaluation practices are used in engineering programs	23

TABLE OF CONTENTS (Continued)

	<u>Page</u>
2.5.1.1 Classroom observation.....	23
2.5.1.2 Evaluation of classroom materials	25
2.5.1.3 Student end-of-course evaluation	26
2.5.1.4 Student mid-course evaluation.....	27
2.5.1.5 Exit evaluation	27
2.5.1.6 Alumni evaluation.....	28
2.5.1.7 Practices usage summary	29
2.5.2 Which practices are used for formative and/or summative purposes	30
2.5.3 What practices do engineering faculty report as useful?	31
2.5.4 Role of rewards	35
2.6 Discussion	35
2.6.1 Teaching evaluation practices are used for either formative or summative purposes.....	36
2.6.2 Importance, critiques and possible improvement of SET.....	36
2.6.3 Approximately 85% of the evaluators who observe the classroom are not trained	37
2.6.4 Alignment with evidence-based teaching practices	37
2.7 Conclusion	41
2.7.1 Student end-of-course evaluation remains the most popular choice for evaluating teaching.....	41
2.7.2 Teaching evaluation methods are used for making formative and/or summative decisions	42
2.7.3 Formative evaluations and assessments were reported to be useful by educators	42
2.8 Acknowledgment	43
2.9 References	43
CHAPTER 3 INVESTIGATING CURRENT APPROACHES TO ASSESSING TEACHING EVALUATION IN ENGINEERING DEPARTMENTS.....	47
3.1 Abstract	48

TABLE OF CONTENTS (Continued)

	<u>Page</u>
3.2 Introduction	49
3.3 Perspectives from literature	50
3.3.1 Research on the purpose of teaching evaluation.....	51
3.4 Methodology	53
3.4.1 Study design.....	53
3.4.2 Participants (Institutional Overview).....	54
3.4.3 Data collection	54
3.4.4 Data analysis	55
3.5 Results	56
3.5.1 What are the current approaches to assessing teaching evaluation?	56
3.5.2 How are these approaches similar or different based on institution type?	57
3.6 Discussions	59
3.7 Conclusions	60
3.7.1 Student evaluation of teaching surveys are still the most prominent teaching evaluation approach.....	60
3.7.2 Level of usage of teaching evaluation approaches were similar among teaching focused institutions and community colleges.....	61
3.8 Acknowledgment	62
3.9 References	62
CHAPTER 4 CONCLUSION.....	64
4.1 Challenges and Recommendations	65
4.2 Future Research	67

TABLE OF CONTENTS (Continued)

	<u>Page</u>
BIBLIOGRAPHY.....	68
APPENDIX A - INTERVIEW PROTOCOL.....	72
APPENDIX B - SURVEY PROTOCOL	78

LIST OF FIGURES

<u>Figure</u>		<u>Page</u>
2.1	Participant's breakdown using Carnegie classification of institutions	18
2.2	Interview questions	19
2.3	Formative and summative teaching evaluation practices	31
3.1	Distribution of institution type used in our study	58

LIST OF TABLES

<u>Table</u>		<u>Page</u>
2.1	Number of participants and their role in their department and/or institution	18
2.2	Primary survey questions	21
2.3	Categories in Dedoose and description.....	22
2.4	Teaching evaluation practices usage summary.....	29
2.5	Teaching evaluation practices usage summary by institution type.....	30
2.6	Methods, purposes, and how teaching evaluation is used within engineering departments	30
2.7	Effectiveness of teaching evaluation practices based on participants' perspectives	32
2.8	Reasons to improve the SET process based on sample participants' perspectives.....	33
2.9	Aligning current teaching evaluation practices with EBIPs	40
3.1	Participants' department role	54
3.2	Summary of methods used to evaluate teaching.....	56
3.3	Teaching evaluation practices utilization summary.....	57
3.4	Teaching evaluation practices utilization classified by institution types	58
4.1	Teaching evaluation practices challenges and recommendations.....	66

Chapter 1 – Introduction

1.1 Research context

Teaching evaluation is important because it offers valuable feedback that may help colleges and universities evaluate the efficacy of an educator. Over the years, teaching evaluation has been criticized because it often relies only on student evaluation of teaching (SET) (Mohanty, Gretes, Flowers, Algozzine, & Spooner, 2005). This critique fueled calls for using multiple sources of evidence to ensure that several teaching dimensions are observed when evaluating teaching (National Research Council, 2003). For this reason, best practices for evaluation of teaching were investigated and described (Goldschmid, 1978; Johnson & Ryan, 2000). Despite knowing that best teaching evaluation practices have been described, the approaches used in engineering programs and the quality of these measures is not known due to the lack of research in this area. Therefore, research that investigates the current state of teaching evaluation practices used in this discipline is needed. This research provides a broad spectrum of teaching evaluation approaches that could be useful in improving teaching and student learning in engineering programs.

Postsecondary institutions are under pressure to examine and change their practices and assumptions about teaching, especially in the Science, Technology, Engineering, and Mathematics (STEM) departments. The National Research Council, NRC (2003) calls for thorough evaluations that have demonstrated the improvement of teaching and learning in STEM. However, there are many variations that exist in colleges and universities. For instance, the institutions academic goals, type of programs offered and number of students they serve vary widely. As a result, the evaluation of teaching in higher education has become inconsistent and less rigorous (National Research Council, 2003).

The goal of this research is to investigate teaching evaluation practices in engineering programs. Furthermore, it seeks to understand educators and administrators perceptions regarding the usefulness of evaluation practices. Exploring teaching evaluation in engineering is essential because engineering expertise is measured by the knowledge, skills and attitudes gained from educational experience. One of the goal in engineering education

is to change faculty practices to enhance the quality of teaching (Fink, Ambrose, Wheeler, 2005). So, studies that explore the nature of teaching as well as how teaching is assessed as being effective or not is of much benefit to engineering.

The student end-of-course evaluation and classroom observation are the two most common approaches used to evaluate teaching (Berk, 2005). Although, some educators criticize that using student evaluations of teaching is not an effective way to measure the quality of their teaching because they are based solely on students' opinions. In addition, classroom observations are only effective if the observers are trained in the process for conducting such reviews. Some faculty are not convinced by the fairness of these approaches used for evaluating their teaching effectiveness (National Research Council, 2003). Furthermore, educators are going to adopt and use teaching evaluation approaches, as they identify which technique is most effective to them. There has been no research that examines which teaching evaluation practices educators perceived as useful and valuable. Therefore, it is essential to explore participants' perspectives because this exploration may potentially lead to increased adoption of effective teaching evaluation practices in engineering programs.

1.2 Research questions

The two manuscripts presented in this thesis attempt to address five research questions.

1. What teaching evaluation practices are used in engineering programs?
2. Which practices are used for formative and/or summative purposes?
3. What practices do engineering faculty report as useful?
4. What are the current approaches used in engineering departments/units to assessing teaching evaluation?
5. How are these approaches similar or different based on institution types?

In chapter 2, three research questions (1 to 3) and in chapter 3, two research questions (4 and 5) were explored to achieve our goals.

1.3 Organization of the thesis

This first chapter describes the context and relevance for conducting this research broadly. It also outlines the research questions and summarizes the organization of the thesis with a description of the research and terminologies used to examine the study of interest. The next two chapters contain manuscripts submitted to peer-reviewed publications. The first manuscript was submitted to a peer-reviewed engineering education journal (International Journal of Engineering Education, IJEE) (Brown, Villanueva, Pitterson, Hurwitz, & Sitomer, n.d.). The second manuscript has been accepted in a refereed conference proceedings (Frontiers in Education, FIE) (Brown, Villanueva, Pitterson, & Sitomer, 2016). The final chapter of the thesis (chapter 4) is a conclusion that presents the results of the studies described in chapters 2 and 3, and the implications of these results.

1.4 Description of research

The primary purpose of this research is to investigate current teaching evaluation practices used in engineering programs and, to understand the evaluation practices' usefulness as perceived by educators and administrators. This thesis addresses this goal with the presentation of two manuscripts, which are presented as individual chapters:

Chapter 2-This study's objective is to explore current practices used for teaching evaluation within engineering programs and the perceived usefulness of these practices for improving teaching. Additionally, this paper provides a description of how current teaching evaluation practices could be aligned with evidence-based teaching practices.

Chapter 3-This study is part of the larger study (Chapter 2), which describes the teaching evaluation practices at three types of institutions with the goal of providing stakeholders a repository of teaching evaluation practices.

The first study (Chapter 2) entitled: *Teaching evaluation practices in engineering programs: current approaches and usefulness*. This paper describes the findings from the data collected from the interview and survey of thirty-four engineering educators. The result showed that there are six approaches for teaching evaluation used in engineering departments which include student end-of-course evaluation, classroom evaluation by peers or non-peers, evaluation of classroom materials, student mid-course evaluation, exit evaluation, and alumni evaluation. These approaches are used for formative and summative evaluations. This study also suggests descriptions of how current teaching evaluation might be aligned with evidence-based teaching practices.

The second study (Chapter 3) entitled: *Investigating current approaches to assessing teaching evaluation in engineering departments*. This paper reports the preliminary analysis of the interviews. This study examines teaching evaluation approaches in engineering departments and whether these approaches are consistent across the different institution types, classified using the Carnegie classification of institutions. The findings from this study suggest that the student evaluation of teaching survey remains the most prominent teaching evaluation approach. Similar teaching evaluation approaches are used across the institution types investigated. Student end-of-course evaluation, peer evaluation and student exit evaluation had comparable usage across all three institution types. Student mid-course evaluation of teaching has a higher usage among teaching-focused universities and community colleges compare to research-intensive universities.

1.5 Terminology

The following terms are used extensively throughout the document and commonly used in teaching evaluation practices literature.

a. Evaluation

Evaluation focused on determining the effectiveness of educational programs, practices, and procedures (Reeves, 2006). The term teaching evaluation is used to describe the collection of evidence of the effectiveness of teaching practices.

b. Assessment

The term assessment is used to measure student learning and person's characteristics such as aptitude (Reeves, 2006). Assessment and evaluation have two distinct meanings, however they are sometimes used interchangeably and are often confused because they both rely on similar data.

c. Formative evaluations

These are evaluations made during the course that are primarily intended to provide feedback so educators can make in-course corrections (Wankat & Oreovicz, 2009). This type of evaluation involves collecting evidence that identifies and improves the quality of teaching (Berk, 2005).

d. Summative evaluations

These evaluations are typically done at the end of the course or after the course is over (Wankat & Oreovicz, 2009). This type of evaluations uses evidence to summarize the overall performance of an educator. Administrators or colleagues used this evaluation for rendering formal personnel decisions such as promotion and tenure, annual review or continued employment and summative data is typically based on a variety of data (Berk, 2005). Summative evaluations may also be used to provide feedback to the educator.

e. Evidence-based instructional practices (EBIPs)

These are teaching approaches that aligned to evidence of student learning that are supported by evidence and research.

f. Active learning

According to Prince (2004), "Active learning is generally defined as any instructional method that engages students in the learning process" (p. 1). Active learning techniques alternatives to traditional teaching methods such as content-centered lectures. Active learning introduces a variety of opportunities for students to interact with content and each other into traditional lectures and promotes student engagement. According to Chi and Wylie (2014), "active learning is learning that requires students to engage cognitively and meaningfully with the materials, to get involved with the information presented, really thinking about it rather than just passively receiving it" (p. 219).

Chapter 2 – Teaching Evaluation Practices in Engineering Programs: Current Approaches and Usefulness

Keisha A. Villanueva, Shane A. Brown, Nicole P. Pitterson, David S. Hurwitz and Ann
Sitomer

International Journal of Engineering Education (IJEE)
(Manuscript under review)

2.1 Abstract

The evaluation of teaching is a critical aspect in higher education. There is substantial knowledge based on best teaching evaluation practices that can be used, yet there is a lack of research regarding teaching evaluation practices that are used specifically in engineering programs. This research characterized teaching evaluation practices within engineering programs across the country, using qualitative interviews and a survey, to understand and assess the current state of practice. Three research questions were explored: (1) *What teaching evaluation practices are used in engineering programs?* (2) *Which practices are used for formative and/or summative purposes?* (3) *What practices do engineering faculty report as useful?* An exploratory sequential mixed-method design, utilizing interview and survey methods was used. In the qualitative phase, data were collected using semi-structured interviews followed by the quantitative phase, which included the development of a survey to more fully understand the evaluation practices of the interviewed participants. Thirty-four educators, including course instructors, department heads, and program coordinators in engineering programs participated in the interview. The result of this study showed that end-of-course student evaluation of teaching (SET) is the most common approach. In addition to SET, other approaches to teaching evaluation include classroom observation by peers or non-peers, evaluation of classroom materials, and student mid-course evaluations. It was also identified that the formative practices used mainly to gather student feedback or to improve faculty teaching are useful to the participants. Although there is substantial interest in improving teaching evaluation practices, generally current practices are still much different from identified best practices. The teaching evaluation system in engineering programs can be improved when educators become aware of and chose to adopt approaches that have been demonstrated to improve teaching and student learning.

Keywords - teaching evaluation, formative evaluations, summative evaluations, engineering, and higher education

2.2 Introduction

A National Research Council (NRC) report proposed that “rigorous evaluation to improve teaching and learning must become integral to Science, Technology, Engineering, and Mathematics (STEM) departmental culture if the broad teaching missions of colleges and universities are to be achieved” (National Research Council, 2003, p. 5). This report indicates that the use of evaluation in higher education to examine and improve teaching quality is increasing. However, evaluation requires an understanding of what to evaluate, how to evaluate, what data to collect and analyze, and how to implement teaching improvements based upon what is learned (Morgan, 2008). Evaluation is one tool that can be used to improve STEM postsecondary teaching. Moreover, evaluation is an important part of instruction and of learning because through evaluations, we can determine the effectiveness of an educator by measuring what students have learned (Dressel, 1960).

Substantial research has been conducted on teaching evaluation in general. Student evaluation of teaching was found to be the most common method for teaching evaluation in higher education (Campbell & Bozeman, 2007). Although it was suggested as a necessary source of evidence of teaching effectiveness, student evaluation of teaching was critiqued for being insufficient when used as the only data source, because it often relies exclusively on the qualitative judgments of students rather than how the teaching supported student learning (Mohanty et al., 2005). This critique has fueled the call for multiple sources of evidence in assessing teaching effectiveness (National Research Council, 2003). Employing additional methods to evaluate teaching effectiveness can help ensure that all dimensions of teaching (e.g., content knowledge, pedagogy knowledge, pedagogical content knowledge), and the course (e.g., lectures, materials, resources, etc.) are explored (Hill, Ball, & Schilling, 2008). Additionally, it is essential to have a clear purpose for the evaluation process used and to ensure that the process is a valid and reliable indicator of educator effectiveness (Morgan, 2008). However, the same method of teaching evaluation may not be valid for all subjects, student groups, levels and disciplines (Nygaard & Belluigi, 2011).

While there are many teaching evaluation methods available for implementation, little research has been conducted that investigates teaching evaluation practices, particularly in engineering programs (Fink, Ambrose, & Wheeler, 2005; National Research Council, 2003). The purpose of our research is to explore teaching evaluation practices in one discipline – engineering – at three different types of institutions. These are associate’s colleges, master’s colleges and universities, and doctoral universities. This study seeks to understand the perceptions of educators and administrators regarding the usefulness of evaluation practices. Investigating teaching evaluation in engineering is important since engineering expertise is measured by the knowledge, skills and attitudes gained from educational experience. Fink, Ambrose and Wheeler (2005) posit that, “the ultimate goal of changing engineering education is to change faculty practices to improve the quality of student learning” (p. 191). Consequently, studies that explore the nature of teaching in engineering as well as how teaching is assessed as being effective or not is of much benefit to the discipline. The importance of preparing high-quality engineers, an important aspect of ABET accreditation, is a pertinent reason for conducting teaching evaluation research in engineering programs. Therefore, this study seeks to share teaching evaluation practices used in engineering departments with the broader community with the hope that engineering stakeholders will benefit from its findings.

2.3 Background

The purpose of teaching evaluation in higher education has changed significantly over the years (Ory, 2000). Teaching evaluation in higher education has evolved from primarily relying on a department chair’s evaluation to a more systematic approach, including a variety of methods such as student end-of-course evaluations, peer reviews, peer visits, self-evaluations, document reviews, and evidence of achievement. In higher education, there have been multiple calls for evidence-based teaching practices (EBIPs), such as active learning, that align with evidence of student learning. “There is extensive evidence that active learning works better than a completely passive lecture. Despite this evidence, adoption of these evidence-based teaching practices remains low” (Eddy, Converse, & Wenderoth, 2015, p. 1). The NRC (2003) recommended that, “teaching effectiveness

should be judged by the quality and extent of student learning, and that there are various teaching styles and methods that are likely to be effective” (p. 118). Still, most colleges and institutions tend to focus more on evaluations that are easy to use and measure because a more robust evaluation system requires a larger investment (Reeves, 2006).

“The terms assessment and evaluation are commonly used interchangeably although they have two distinct meanings. Both terms are often confused because the same data can be used for both assessment and evaluation” (Reeves, 2006, p. 305). For example, final exams can be used to assess student learning in courses. Furthermore, it can be utilized in the evaluation of the effectiveness of a course. In this paper, the term evaluation is defined as a collection of evidence regarding teaching practices, whereas assessment is defined as the activity of measuring student learning. The distinction between formative and summative evaluation is described in Section 2.3.2 later in this paper. We also use the term educator to describe a person who provides instruction in a classroom.

2.3.1 What practices are used?

“Educational technology and research have produced tools, and strategies to evaluate and enhance teaching effectiveness, as well as insights into the underlying processes and mechanisms” (Goldschmid, 1978, p. 221). “The literature on the evaluation of college teaching varied from developing effective faculty evaluations to very specific research investigating a variety of methods for evaluating teaching” (Johnson & Ryan, 2000, p. 109). The most frequently referenced methods include student end-of-course evaluations, peer observation, portfolios, evaluation of course materials, and self-evaluation by educators. However, there is no general agreement on how to evaluate teaching in higher education. There is no distinct formula to effectively evaluate teaching because there are considerable variations that exist across institutions and disciplines (National Research Council, 2003). The authors contend that the utilization of teaching evaluation practices is affected by context and culture. The NRC (2003) reports there are differences in cultures of research and teaching in higher education, as well as the criteria used in the evaluation of teaching by disciplines were examined. For example, teaching effectiveness is less valued when

compared to research productivity by many colleges and universities. In community colleges, the evaluation of teaching and learning has received greater attention (National Research Council, 2003). Therefore, it is proposed that effective teaching evaluation stems from a combination of multiple sources of evidence. For example, students (current, graduating seniors, and alumni), graduate teaching assistants, departmental and other faculty colleagues could provide evidence of effective teaching through evaluation and assessment. This group of individuals could provide critical information related to the educator's role in students' learning, approaches to teaching, mentoring, the currency of the materials the educator presents, level of student engagement or participation and leadership in improving undergraduate education (National Research Council, 2003, pp. 3–4).

Student evaluation of teaching is the most common method used for evaluating teaching and courses (Chen & Hoshower, 2003; Goldschmid, 1978). This method is used in higher education to summarize the students' overall perceptions at the end of a term or semester. “Still, student’s learning processes remain underexposed using this quantitative data analysis because the data collected focus on a narrow range of teaching behaviors, unreflective of the cognitive, and affective structures that they construct a partial image of teaching and learning” (Nygaard & Belluigi, 2011, p. 658). In addition, there also appears to be little evidence that SET forms and procedures measure or contribute to teaching quality (Brandenburg, Braskamp, & Ory, 1984). In response to these problems, Broder and Taylor (2015) conducted a survey of teaching evaluation in agricultural economics and related departments. They had three objectives, which include developing a theoretical model of the teaching evaluation process, examining departmental policies and assessments of teaching evaluation, and identifying factors associated with effective teaching evaluations. The data used for this study were taken from a survey conducted in 1988 where department heads in the United States ($n=51$) and Canada ($n=8$) were asked to complete a questionnaire on teaching evaluation in their department and to submit a sample of copies of teaching evaluation forms and statistical summaries. From this study, “empirical models were estimated which characterize faculty reliance on SETs and faculty

satisfaction with SET forms. The faculty was found to have specific preferences for SET form and content, and improvements in SETs are more likely to result from increasing the quality rather than the quantity of SET questions” (Broder & Taylor, 2015, p. 153).

The most common method, and in some cases the only method, used for the evaluation of teaching was student end-of-course teaching evaluation (Chen & Hoshower, 2003; Goldschmid, 1978). Several studies suggest that student end-of-course evaluations are reliable, and have a high level of validity because students are in the best positioned to judge particular aspects of teaching, and the classroom (Felder & Brent, 2004) However, many studies find that evaluations by students invalid and unreliable when these student end-of-course evaluations are the only method used. Regardless of being reported to be useful, evaluation by students was found to be neither the sole nor the best way to evaluate a course (National Research Council, 2003).

2.3.2 Formative and summative teaching evaluation practices

Generally, there are two types of evaluation practices: formative and summative. Formative evaluation incorporates evidence from faculty or students during and after the course to improve teaching and student learning. This type of information can be used to make in-course changes. Examples of formative evaluations include conversing with students informally during the course to determine what is or is not working, formal weekly meetings with a group of class representatives or dialoguing with Teaching Assistants (TAs) (Wankat & Oreovicz, 2009). Summative evaluations are done after the course and are used for a variety of purposes. These evaluations conclude the educator's overall performance or status and are used to make decisions about annual performance review, promotion and tenure. For example, summative follow-up evaluations by alumni can provide feedback as to what course material has proven to be particularly useful in the industry (Wankat & Oreovicz, 2009).

A variety of formative and summative teaching evaluation practices are used in the evaluation of teaching. While “SET are typically used summatively and exclusively”

(Henderson, Turpen, Dancy, & Chapman, 2014, p. 14), other measures of teaching effectiveness such as mid-course evaluation, exit evaluation, and alumni evaluation have the potential to be used formatively. Student end-of-course evaluation is the most influential measure of educator performance, used for both formative and summative decisions (Berk, 2005). Exit and alumni ratings provide information about the quality of teaching, courses, curriculum, admissions, and other topics on a programmatic level (Berk, 2005). In addition, graduating students and alumni could provide evidence about the educator's role in their learning (National Research Council, 2003, p. 60). Additional sources of evidence for teaching evaluation could provide information gathered from peers and non-peers (e.g., administrators). For example, peer review is an alternative practice, and composed of two activities: (a.) peer observation of in-class teaching performance and (b.) peer review of the written documents used by an educator in a course (Berk, 2005). Peer ratings of teaching performance and materials were found to be the most complementary source of evidence to student end-of-course evaluations. These practices cover the aspects of teaching that students are not in a position to evaluate (Berk, 2005). Lastly, the educator could provide self-assessment of his or her teaching strengths and areas for improvements through self-evaluation by submitting a teaching portfolio.

2.3.3 Call for improvements to teaching evaluation practices

Critiques of current evaluation practices have fueled the call for multiple sources of evidence in assessing teaching effectiveness during the last decade (National Research Council, 2003). Wankat et al. (2009) suggest that students are not qualified to evaluate the educator's teaching, and they often miss the richness of ideas, which can be obtained using other evaluation procedures. Student evaluation of teaching may confuse the evaluation of the educator with an evaluation of the course (National Research Council, 2003). There are also concerns that factors other than teaching quality, including the size of the class, course grade distributions, and whether it was being taken as an elective or a requirement can influence students' responses on such evaluations and, for this reason, student end-of-course evaluations may not be directly comparable to courses or educators (National Research Council, 2003).

In terms of evaluation, one single data source, whether from students, peers or mentors, or one single collection method such as interviews, questionnaires, or discussion, may provide insight on only one aspect of teaching. For instance, a study by Iqbal (2013) showed that a peer review, which usually consists of observation and review of materials, along with student evaluations might be useful for informing the evaluation of teaching because students are best at providing feedback on the quality of student-educator interactions. On the other hand, colleagues are better positioned to comment on matters such as content expertise, instructional design, and methods of assessment (Iqbal, 2013). Several different sources and methods should be utilized to provide a more comprehensive overview and to ensure a range of teaching processes are explored. Each of these sources can provide unique information, but when used alone each has limitations (National Research Council, 2003). A combination of techniques can make up for the deficiencies of student end-of-course evaluations. Evaluations of teaching by students need to be supplemented by other methods, for example, peer observation or review of course artifacts, such as syllabi, course activities, and assessments (Ohanlon & Mortensen, 1980). Colleagues in the discipline are able to determine whether course activities are appropriately challenging and accurate (Felder & Brent, 2004). By drawing upon three or more different sources of evidence, the strengths of each source can compensate for the weaknesses of the other sources, thereby, converging on a decision about teaching effectiveness that is more accurate than one based on any single source (Berk, 2005; Ohanlon & Mortensen, 1980). For instance, Berk (2005) proposed a unified conceptualization of teaching effectiveness where evidence must be collected from a variety of sources to define the construct and to make decisions about its attainment. Each source can supply unique information, but can be weak, usually in a way different from the other sources.

Several comprehensive models of faculty evaluation have been proposed. They include multiple sources of evidence with a greater weight attached to student and peer input and less weight attached to self-evaluation, alumni, and administrators. All of these models can be used to arrive at formative and summative decisions (Berk, 2005). “For any formative or summative evaluation, it was found that assessment based on a single teaching activity

(i.e., classroom presentation) or dependent on information from a single source (i.e., student evaluation forms) is less reliable, useful, and valid than an evaluation of an educator's strengths and weaknesses based on multiple sources" (National Research Council, 2003, p. 51). Comprehensive evaluations of teaching are more accurate, particularly when based on the views of current and former students, colleagues, and the educator or department being reviewed. The process of evaluating teaching has been found to work best when all faculty members in a given department play a strong role in developing policies and procedures. "This is the case because evaluation criteria must be clear, well known, and understood, scheduled regularly, and acceptable to all who will be involved with rendering or receiving evaluation" (National Research Council, 2003, pp. 51–52).

"Research suggests that some combination of formative and summative evidence about student learning can be helpful in evaluating and improving an educator's teaching" (National Research Council, 2003, pp. 2–3). It is known that there are practices that can contribute to effective teaching and learning criteria for assessing teaching performance (National Research Council, 2003). Despite knowing that such teaching evaluation practices already exist, it is still unclear what quality of the measures provides evidence of teaching effectiveness and student learning. Surprisingly, little is known about what teaching evaluation practices are used, especially in engineering programs, and what practices are actually working for engineering educators in their department or institution.

2.3.4 What practices work for engineering educators?

There are a variety of teaching evaluation practices available for implementation. However, there are far fewer that are frequently used. The adoption of evidence-based practices is a complicated process that involves social, cultural, and individual factors. The perception of value is an important aspect of the adoption process; what educators believe is most effective is influential on what is adopted by the faculty. However, research has not identified which practices educators perceive as effective. We do not argue that what educator believe is effective actually is, but rather that effectiveness, as noted above, is an

elusive goal, and perceived effectiveness may lead to increased adoption of effective and holistic practices would be of benefit to the community.

Research in the constructivist paradigm suggests knowledge is constructed and adopted within a community when it is seen as “useful, practical and adaptive” (Ferguson, 2007). It can be argued that when each individual member has the opportunity to share in the development, implementation, or enactment of a particular practice embodied by the community more valued is placed on said practice (Crotty, 1998). With respect to perspectives on teaching evaluation practices, participants’ knowledge of these practices can be deemed as constructed within the social context of their institutional mission as it relates to how important effective teaching practices and evaluation are. Consequently, since human practices are constructed in and out of their interaction with each other within an essential social context, it is necessary to gauge the perspective of these individuals (Crotty, 1998). Building on the idea of constructivism, the authors posit that studying participants’ perspectives on teaching evaluation practices utilized within their engineering programs is an important and valid area of research.

The objective of this research study is to characterize best practices for assessing teaching evaluation within engineering departments across the country. To meet this goal, three research questions must be addressed.

1. What teaching evaluation practices are used in engineering programs?
2. Which practices are used for formative and/or summative purposes?
3. What practices do engineering faculty report as useful?

2.4 Method

2.4.1 Study design

This research study is based on a constructivist paradigm, which “recognizes the importance of the subjective human creation of meaning, but does not reject outright some notion of objectivity” using a descriptive case study with multiple embedded units (Baxter & Jack, 2008, p. 549). This case study was used to describe the data collected as they occur in various institutions. The data were gathered using an exploratory sequential mixed-method design to answer the research questions, which involves collecting both qualitative and quantitative data (Creswell, 2014). The qualitative data were collected through semi-structured interviews; the quantitative data were gathered from the survey that was developed after the preliminary analysis of the data collected from the interview. The survey was developed to further understand the practices of the research participants and to gather additional information that was not captured from the interview. Collectively, the interview and survey data provided a holistic overview of evaluation practices used by the participants and the value of such practices, as reported by the participants. This study was reviewed and approved by the Institutional Research Board (IRB).

2.4.2 Participants characteristics

Individuals with knowledge regarding the teaching evaluation in their department and/or institution were invited via email to participate in this study. Participants include course instructors, department heads, and program coordinators in engineering programs at doctoral universities, master’s colleges and universities, and associate’s colleges based on the Carnegie classification of institutions. Invited participants contact information was found from their college or university website and through personal contacts. Snowball sampling was used to identify other colleagues who were recruited and whose contacts were provided by the existing participants. Upon expressing interest in participating in the research study, the individuals were contacted again to set up a telephone interview and were provided with the IRB approved consent form. The participants’ location varied, but a large proportion of the participants were from the western region of the United States ($n=23$). The rest of the participants were from Midwest ($n=4$), Southwest ($n=2$), Northeast

($n=1$) or Southeast regions ($n=4$). Table 2.1 shows the number of participants and their role in their department and/or institution. Figure 2.1 provides the breakdown of participant by institution types. Each participant mentioned having more than one role in their department, though the table below shows the participants' primary role.

Table 2.1: Number of participants and their role in their department and/or institution

Participant's department role	Number of respondents
Department Chair	15 (44%)
Associate Department Chair	1 (3%)
Full Professor	1 (3%)
Associate Professor	6 (18%)
Assistant Professor	3 (9%)
Engineering Instructor ^a	6 (18%)
Program Coordinator	2 (6%)

^a These include educators from associate's colleges and non-tenured-track educators from universities

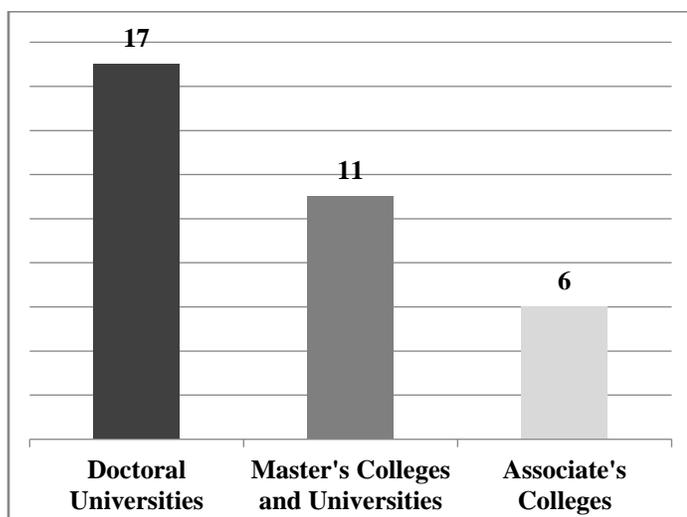


Figure 2.1: Participant's breakdown using Carnegie classification of institutions

2.4.3 Data collection

Data were collected using semi-structured interviews and survey of teaching evaluation practices.

2.4.3.1 Semi-structured interviews

In the qualitative phase of this study, data were collected through semi-structured interviews conducted over the phone. An interview protocol was created to question the participants about the teaching evaluation practices in their department. The participants were initially asked about their background followed by the main question, “*What practices have you used or are currently using to evaluate teaching in your department?*” Overall, they were asked to report what, and how teaching evaluation practices are conducted in their department. However, participants also reported their *own* practices in conducting teaching evaluations. Additional questions were asked based on the participant’s responses in order to elicit more details about particular practices the participants mentioned. Figure 2.2 shows examples of the questions asked in the interview protocol. The duration of each interview ranged between 15-60 minutes in length.

<p>Participant’s Background</p> <ul style="list-style-type: none"> • What is the name of your university? • What is the name of your department? • Could you please tell me your role in your department? <p>What practices have you used or are currently using to evaluate teaching in your department? (e.g., student end-of-course evaluation, classroom observation, and evaluation of classroom materials)</p> <ul style="list-style-type: none"> • What is the purpose of this teaching evaluation practice? • Could you please describe the process for this method? <p>Based on the participant’s response, clarification of the teaching evaluation practices may be needed. Examples 1 and 2 provide some follow-up questions that may be asked to the participants if they indicated using student evaluation of teaching and peer review.</p> <p>Example 1: Student Evaluation of Teaching (SET) Questions</p> <ul style="list-style-type: none"> ▪ What is the purpose of the SET in your department? ▪ How it is administered, online or paper-based and why? ▪ Who determines what questions are on the SET? ▪ Are the questions on SET standard? ▪ Please provide some examples of the questions asked on SET? ▪ Can students write comments on the SET? ▪ Can instructors add their own questions on SET? ▪ When do students usually take the survey? ▪ Can students still take the survey after grades have been posted? Can students still take the survey after grades have been posted? ▪ Are the evaluations published for students and for public to see? Explain

Figure 2.2: Interview questions

Example 2: Peer Review Questions

- What is the purpose of the peer review in your department?
- What does a single peer review consist of?
 - Observation of single or more lectures
 - Review of syllabus, course notes, assignments, exams, etc.
- Who is the evaluator (peer from the same and/or different department)?
 - Does the evaluator get any form of training?
 - Does he use a checklist or standard form during an evaluation?
- What documentation is provided to the faculty member?

Does your institution have a Center for Teaching and Learning or equivalent?

- What services does your CTL provide related to teaching evaluation and assessment?

Participants' Perspectives

- What do you think about the teaching evaluation practices in your department? Are they effective or ineffective and please explain why?
- What changes would you like to see regarding the teaching evaluation practices in your department?

Could you provide us any materials that are used for teaching evaluation in your department?**Figure 2.2:** Interview questions (Continued)*2.4.3.2 Qualtrics survey*

The interview protocol was adapted throughout the study to add questions based on what was learned during previous interviews. For this reason, a supplemental survey was developed to gather more information and to clarify some of the findings that emerged during the interview with the participants. The survey questions contained similar questions to those from the interview, but it further explored the practices of the interview participants. The survey was distributed to interview participants via email. The time necessary for participants to complete the survey ranged between five and 30 minutes; they were given one month to complete the survey.

There were four different sections and 13 primary questions on the survey. The first section of the survey was intended for the collection of background information, which includes institution and department name, average number of students who complete the engineering program annually, and degrees that the institution offers. The second section of the survey consisted of questions related to teaching evaluation practices gathered from either the faculty member's peers or administrators (i.e., classroom observation and

evaluation of classroom materials). The third part asked for teaching evaluation practices gathered from the students (i.e., student end-of-course evaluation, student mid-course evaluation, exit evaluation, and alumni evaluation). Finally, the survey asked whether the participant's institution has a Center for Teaching and Learning (CTL) or equivalent. Finally, participants were asked whether any other teaching evaluation practices not mentioned in the survey are used. Upon indicating the use of a particular method, sub-questions asked participants to elaborate on that method. However, if the participant indicated that their department does not use a certain method, the survey skipped the sub-questions as they would not apply to these participants. Table 2.2 shows the 13 primary questions asked on the survey protocol.

Table 2.2: Primary survey questions

1. Context (Background)	What is the name of your college or university?
	What is the name of your department or school?
	On average, how many undergraduate students complete your engineering program annually?
	On average, how many graduate students complete your engineering program annually?
	What degrees are offered?
2. Teaching Evaluation Practices gathered from either Peers or Non-Peers	Does your school or department use classroom observation?
	Does your school or department use classroom materials evaluation?
3. Teaching Evaluation Practices gathered from Students	Does your school or department use student end-of-course evaluation?
	Does your school or department use student mid-course evaluation?
	Does your school or department use exit evaluation?
	Does your school or department use alumni evaluation?
4. Other Teaching Evaluation Practices (Programs such as Center for Teaching and Learning)	Does your college or university have a Center for Teaching and Learning or equivalent that provides support and services for faculty members related to teaching evaluation and/or teaching improvement?
	Does your school or department use any other teaching evaluation practices that are not mentioned?

2.4.4 Data analysis

The audio recordings were transcribed either via a professional transcription company or by graduate researchers. Transcriptions were uploaded into Dedoose (2015), a web application for managing; analyzing qualitative and mixed methods research data. The participants' transcriptions were coded and arranged into four categories: method, purpose, procedure, and participant's perspective. For example, the number of participants who used student end-of-course evaluation were grouped together and counted to identify the total number of this approach. The same analysis was completed with other teaching evaluation practices using the transcription data. Then, these data were combined with the data collected from the survey to summarize and describe the overall data set that represents the study sample.

While the data were initially separated into three categories based on institution types, the unequal distribution of participants in each type meant it was not possible to provide meaningful comparisons of teaching evaluation practices across institutions types. Half ($n=17$) of the total participants came from doctoral universities, 11 of the participants were from master's colleges and universities, and the rest came from associate's colleges ($n=6$). Instead, the data were grouped based on their purposes and usefulness for further analysis. Table 2.3 provides a short description of the four main categories.

Table 2.3: Categories in Dedoose and description

Categories in Dedoose	Description of what is included in each category
Method	Variety of approaches that are currently used to evaluate teaching in engineering departments and by individual participants such as student end-of-course evaluation, student mid-course evaluation, classroom observation (peers and non-peers), alumni evaluation, exit evaluation, review of classroom materials and classroom assessment techniques
Purpose	Formative evaluation, summative evaluation, or both. For instance, promotion, tenure, teaching improvement, accountability, feedback, annual review, disciplinary action, newly hired assessment
Procedure	Process for how evaluation methods were enacted in the department. For example, voluntary or mandatory, formal or informal
Participant's Perspective	Perspective regarding what teaching evaluation practices were found to be useful by our participants

2.5 Results

2.5.1 *What teaching evaluation practices are used in engineering programs?*

The participants commonly referenced the following approaches classroom observation by peers and/or non-peers, evaluation of classroom materials, student end-of-course evaluation, student mid-course evaluation, exit evaluation, and alumni evaluation. Results gathered from exit and alumni evaluations were sometimes used to provide formative feedback to educators, but more frequently for students to provide feedback about their experience in the engineering program. Some participants mentioned additional methods such as teaching portfolios, self-assessment, and informal in-class assessment techniques. However, the use of these methods to evaluate teaching was not explicitly discussed. Shown below are more detailed accounts of the practices used at participating institutions.

2.5.1.1 *Classroom observation*

Classroom observation by peers and/or non-peers:

A. Classroom observation by peers

With this method, the evaluator is usually a senior colleague or tenured faculty member. About 60% ($n=20$) of peer evaluators were from the same department. This process is typically voluntary where the educator asks his or her peer to come visit a single lecture. The purpose varied by departments; it could be used for promotion purposes or to gather information for formative feedback as part of a program review, for example. However, this method is usually used to obtain constructive feedback to help improve the individual educator's teaching. This method was also used as part of a mentorship program to assist newly hired educators.

B. Classroom observation by non-peers

A non-peer is considered a dean, program director, tenure committee member, department head or an external evaluator, such as someone who works with the institution's Center for Teaching and Learning (CTL), who is knowledgeable about pedagogical strategies or the content being taught. Non-peer classroom observation

was typically evaluated by the following: department chairs ($n=14$), CTL personnel ($n=8$), deans ($n=7$), committee members ($n=3$), and other external teaching evaluator ($n=1$). This process could be voluntary or mandatory depending on the department. The information gathered may be used to provide constructive feedback to the educator. However, it was usually conducted for promotion purposes, annual review, or in response to student complaints or low scores on student evaluation of teaching.

Overall, the process for conducting a classroom observation was similar when conducted by both peers and non-peers. Twenty-nine out of 34 participating institutions used classroom observation by peers or non-peers to evaluate an educator's teaching. The process usually consists of three parts: pre-observation, actual classroom visit, and post-observation. During the pre-observation, the educator, and the observer meet to clarify expectations for the evaluation and the best way to conduct the observation. The observer may also request classroom materials for reviews, such as syllabus, exams, projects, assignments or quizzes. Then, the observer visits an educator's lecture. The observer evaluates the educator's teaching performance by focusing on the educator's content knowledge, delivery, teaching methods and learning activities. The actual observation is followed by a post-observation where the observer and educator meet again to discuss what went well during the observation and items that may need attention to help improve the educator's teaching performance.

Fourteen participants revealed that educators only receive a single classroom observation per course. The number of observations or whether an educator gets evaluated using this method depends on their rank, position, and employment status (e.g., full-time or part-time faculty). For instance, at doctoral universities, master's colleges and universities, tenure-track and adjunct instructors were sometimes the *only* ones who get evaluated or they receive more evaluations than tenured faculty members. The tenured faculty members either do not get evaluated or if they did, they were not evaluated as regularly as tenure-

track and adjunct educators. Some associate's college participants said that they receive very few observations once they achieved continuous employment.

Eight participants indicated that their department used a standard form or checklist during the classroom observation. Twenty-one participants indicated that their department does not have a standard form or checklist for evaluators to use during the observation. Some evaluators were provided a template as a guide for writing their observations while most evaluators just wrote their observations without a template. About 85% ($n=29$) of the total participants stated that the evaluator or observer does not receive any sort of training or if they do, it was minimal. If training occurs, it usually consists of a workshop or ongoing training provided by the CTL.

2.5.1.2 Evaluation of classroom materials

Participants revealed that classroom materials were typically requested and evaluated as part of the classroom observation process. Fifty percent ($n=17$) of the total participants indicated that their department collects and evaluates classroom materials to assess faculty teaching. Different types of documents, such as course materials (syllabus, and assignments) and/or result of student works (exams, projects, and quizzes), were evaluated. The evaluation of course materials usually occurred when the purpose was to provide formative feedback to the educator. Certain types of documents were additionally requested when the information gathered from this type of evaluation was intended to be used for promotion purposes, annual review, and/or ABET accreditation. In engineering disciplines, ABET accreditation provides evidence that the program met the essential standards needed to ensure that graduates are ready to transition to work as an engineer. It was found that a peer from the same department was the most common choice to evaluate these materials. However, non-peers such as deans, department chairs, an ABET reviewer, or a tenure committee member might evaluate these materials. Many of the individuals evaluating course materials were not trained and did not have a standard form or checklist to review these materials.

2.5.1.3 Student end-of-course evaluation

All study participants indicated that they used student end-of-course evaluations, yet they did not all conduct this practice in the same way. There were two methods of student evaluations of teaching: online and traditional paper format. About 68% ($n=23$) of the participants used an online student end-of-course evaluation to which students were provided a link, while 26% ($n=9$) of the participants used the traditional paper-based format to survey their students. The remaining participants ($n=2$) indicated using both methods. The traditional paper-based format survey was given in an educator's class. Depending on the department, the student end-of-course evaluation results can also be seen and accessed by the dean, program chair, review committee, or tenure committee. Many of the participants mentioned that their department had switched from paper-based evaluations to online evaluations. However, our data show that as institutions considered moving to online evaluations, educators' concerns about response rates emerged.

Some participants (6%, $n=2$) indicated there were different forms of the student end-of-course evaluations from which they could choose that depend on the class level or size of the class that they are teaching. The majority of the participants (94%, $n=32$) said that there was only one form that they could use. The questions on the student end-of-course evaluations were standardized either by the institution, by the college or school of engineering, or by the department only (e.g., civil, mechanical, chemical, electrical). Nearly 62% ($n=21$) of the participants stated that their institution used a set of standardized questions on the student end-of-course evaluations. Four participants discussed that their department developed their own questions for the student end-of-course evaluations, and three participants indicated that the entire college or school of engineering has their own questionnaire for student end-of-course evaluations. Half of the participants ($n=17$) said that they were able to add supplemental questions on the SET to help them improve their teaching and to measure student learning. However, some participants indicated that although adding questions was possible, not many educators took advantage of this opportunity.

2.5.1.4 Student mid-course evaluation

Approximately 74% ($n=25$) of the participants used student mid-course evaluations to assess their own teaching. This evaluation was completed in various ways depending on the educator's preference. Overall, educators undertook student mid-course evaluations voluntarily, that is, the information gathered from this type of evaluation was only seen and used by the educator. There were different ways to administer this type of evaluation. The educator conducting his or her own student mid-course evaluation was the most common way. For example, some of the participants specified using different classroom techniques such as minute paper, muddiest point paper, or having their students write to gather feedback about their teaching and to assess student learning. Only a few of the participants specified asking another, such as a person from a Center for Teaching and Learning, to come and perform mid-course evaluations, providing feedback on how the educator may improve his or her teaching. The intention was normally to facilitate assessment and improvement of teaching for newly hired educators or an educator teaching a class for the first time, or to help an educator trying a new teaching technique. Questions that educators asked their students when administering a student mid-course evaluation include questions such as:

- “What is working in the classroom that helps enhance the student learning?”
- “What specific improvements would help the student learn better?”
- “What could the educator do differently to help this particular class?”

2.5.1.5 Exit evaluation

Exit evaluations were given mostly to graduating seniors before they completed their engineering program. This typically applied to students who graduated from institutions such as doctoral universities and master's colleges and universities. For associate's colleges, this evaluation was given to students who were transferring to a four-year institution. The main purpose of the exit evaluation was to assess what the students have learned and to guide longer-term assessment of the program. It was also used to make adjustments in the engineering program in terms of course offerings or who teaches the courses. This process was not used to evaluate a specific course, although students could

comment on specific courses. Exit evaluations assess the overall quality of the courses taught in the department. Generally, this was used to help the department identify strengths and areas for improvement, and eventually to implement a change to improve the program. However, some educators indicated that they use the results from exit evaluations as formative feedback to improve their teaching as well as student learning. This type of evaluation involved questions such as if the students were satisfied with the program, changes the students would like to see within the program, and if they found themselves ready to enter the workforce. Below are sample questions from exit evaluations:

- “What did you do in your class?”
- “What did you learn?”
- “What do you think about the overall quality of education in your program?”

Students were either given a form to fill out or a link to the survey. However, some participants mentioned administering or accompanying the survey with an oral interview where the students were given an opportunity to discuss and explore their answers in more detail. Approximately 71% ($n=24$) of the 34 participants practiced this method in their department. Nearly 15% ($n=5$) of these participants indicated using this type of evaluation as part of the ABET accreditation program. About 6% ($n=2$) of the participants indicated that the senior exit survey was part of the National Survey of Student Engagement (NSSE), which gathers information nationally about students’ participation in programs and activities at the institution.

2.5.1.6 Alumni evaluation

Approximately 32% ($n=11$) of the total participants administered an alumni evaluation. The alumni evaluation was very similar to the exit evaluation; however, it was usually given to previous students who have successfully completed the engineering program. About 21% ($n=7$) of the 34 total participants indicated using this as a formative assessment, and others ($n=5$) as part of the ABET accreditation process. Just like the exit evaluation, the alumni evaluation was typically not used to evaluate an individual educator. Rather, it was focused on the engineering program as a whole. Yet, some educators used results from

alumni evaluations to help improve their teaching when the students provided comments regarding specific instructors or courses.

2.5.1.7 Practices usage summary

Table 2.4 summarizes the teaching evaluation practices used. The most common method used to evaluate teaching within engineering departments was a student end-of-course evaluation. It was followed by classroom observation with 85% utilization. The student mid-course evaluation was used by 74% of participants, although this study revealed that departments did not require this method.

Table 2.4: Teaching evaluation practices usage summary

Teaching evaluation practices	Usage by number of total participants
Student end-of-course evaluation	34 (100%)
Classroom observation either by peers or non-peers	29 (85%)
Student mid-course evaluation	25 (74%)
Student exit evaluation	24 (71%)
Evaluation of classroom materials	17 (50%)
Alumni evaluation	11 (32%)

Table 2.5 shows the methods usage by institution type using the Carnegie classification of institutions. Doctoral universities, master's colleges and universities, and associate's colleges used similar evaluation approaches. For example, student end-of-course evaluation was used 100% of the time at all three institution types investigated. Similarly, classroom observation either by peer or non-peer and student mid-course evaluation had comparable usage across all three institutions. Master's colleges and universities institutions reported the highest usage of classroom material evaluation with 73%. Doctoral universities tend to utilize student exit evaluation and alumni evaluation more frequently than other institution types.

Table 2.5: Teaching evaluation practices usage summary by institution type

Teaching evaluation methods by institutions	Usage by participants' institution		
	Doctoral Universities	Master's Colleges and Universities	Associate's Colleges
Student end-of-course evaluation	17 (100%)	11 (100%)	6 (100%)
Classroom observation by peers or non-peers	15 (88%)	9 (82%)	5 (83%)
Student mid-course evaluation	12 (71%)	9 (82%)	5 (83%)
Student exit evaluation	15 (88%)	7 (64%)	2 (33%)
Evaluation of classroom materials	7 (41%)	8 (73%)	2 (33%)
Alumni evaluation	7 (41%)	3 (27%)	1 (17%)
Total number of participants	17	11	6

2.5.2 Which practices are used for formative and/or summative purposes?

The exit and alumni evaluations were not usually used to evaluate teaching but to assess the program as a whole. They were both used for formative feedback or for ABET accreditation purposes. Table 2.6 shows the most common methods used to evaluate an educator's teaching, its purpose, and how it is used. In this study, mandatory evaluations correspond to teaching evaluation methods that were required of an educator, while voluntary evaluations were methods that educators used at their own discretion.

Table 2.6: Methods, purposes, and how teaching evaluation is used within engineering departments

Teaching evaluation methods	Purposes	How it is used
Classroom observation by peers and non-peers	Formative and Summative	Mandatory and Voluntary
Evaluation of classroom materials	Formative and Summative	Mandatory and Voluntary
Student end-of-course evaluation	Formative and Summative	Mandatory
Student mid-course evaluation	Formative	Voluntary

Information gathered from classroom observation was commonly used for promotion purposes ($n=11$), annual review ($n=9$), and teaching improvement ($n=17$), regardless of who is the observer: peer or non-peer. Fourteen participants indicated that both peers and

non-peers could be the evaluator. Other participants indicated that only their peers ($n=7$) or non-peers ($n=8$) evaluate them. The review of classroom materials that may accompany classroom observations were reported by around 26% ($n=9$) of the total participants to be used for promotion purposes, 24% ($n=8$) for annual review, 18% ($n=6$) for teaching improvement, and 21% ($n=7$) for ABET accreditation purposes. The student end-of-course evaluation was used for two very different purposes within departments: formative and summative. Nearly 15% ($n=5$) of the participants described using these evaluations both solely and formatively to improve teaching. The same number of participants used the student end-of-course evaluations solely for making summative decisions for educators' promotion, tenure, or annual review. Many participants ($n=24$) used the student end-of-course evaluation information for making both formative and summative decisions. All participants ($n=34$) used the results gathered from the student mid-course evaluation as formative feedback. Figure 2.3 provides the number of teaching evaluation practices usage by purposes, formative or summative.

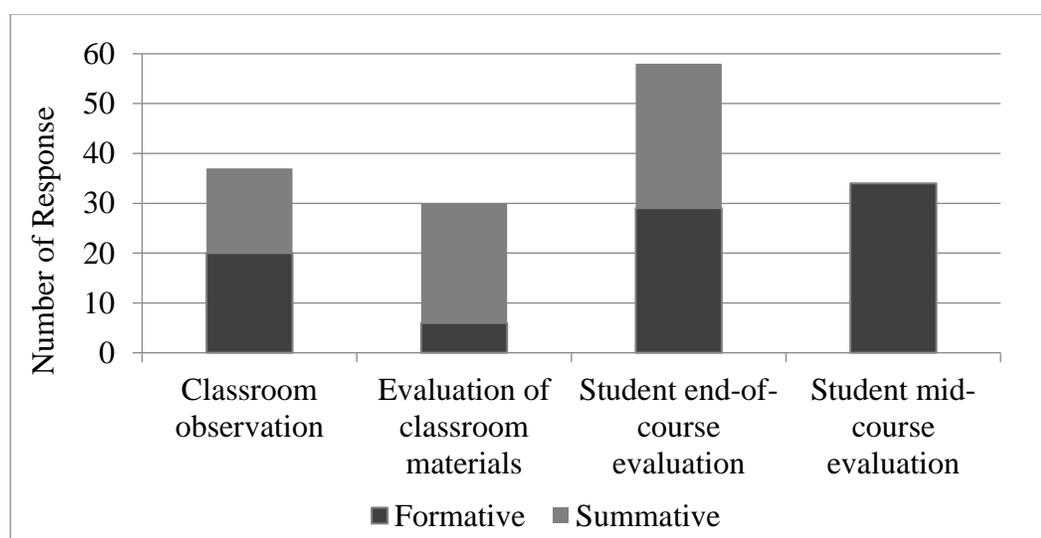


Figure 2.3: Formative and summative teaching evaluation practices

2.5.3 What practices do engineering faculty report as useful?

Participants' perspectives on the effectiveness of teaching evaluation methods were explored with the interview question, "What do you think about the teaching evaluation practices in your department?" The participants were also asked about what practices they

found to be effective or ineffective. Table 2.7 shows the participants' perspectives about the collective teaching evaluation practices in their department, which were broken down into three parts: somewhat effective, ineffective, or neither.

Table 2.7: Effectiveness of teaching evaluation practices based on participants' perspectives

Teaching evaluation practices effectiveness	Number of response
Somewhat Effective	15 (44%)
Ineffective	15 (44%)
Neither	4 (12%)
Total number of participants	34

Participants found two practices, student end-of-course evaluation and classroom observation, more useful for evaluating teaching when the purpose was to gather students' or colleagues' formative feedback to improve teaching, but less useful when it is used for promotion and tenure purposes. For instance, participants reported student end-of-course evaluation to be a useful tool because it gathers student feedback regarding the course and the educator's teaching; students spend more time in the classroom than an observer. However, participants felt it was beneficial to gather additional data, such as data from classroom observations. Examples of student end-of-course evaluation being effective based on two participants' response are shown below:

- a) Participant 150, *"Student course evaluation is a very important instrument. It creates evidence when a faculty member is misbehaving or underperforming."*
- b) Participant 1250, *"I generally find it [SET] pretty helpful and effective. I think that it allows me to reflect on what I do pretty well, and so I don't find it to be too lacking or anything like that. Students are not hesitant to be critical. In terms of at least thinking about how I can improve, that's a good way to do that, I find the student evaluations to be more helpful."*

Based on many participants' perspectives, they believed that student end-of-course evaluations and classroom observations were only useful to some extent and would benefit

from improvements. The improvements or modifications that 44% of the participants ($n=15$) would like to see were aligned with why participants believe that these two teaching evaluation practices were ineffective. Participants critiqued both classroom observation and the student end-of-course evaluation process. Suggestions from participants included using standardized forms or rubrics during a classroom that could provide specificity, as well as remind the evaluator of what needs to be assessed. In addition, regular observations and using a trained observer, who knows what to evaluate and how, will help improve classroom observation. Furthermore, the student evaluation of teaching (SET) process was weak and imperfect in some way. These reasons include, but are not limited to, the type of questions being asked in the SET survey, timing, regularity of the survey, what it measures, consistency of the system, students as evaluators, and the purpose that it served. Table 2.8 provides some participant perspectives on the SET process.

Table 2.8: Reasons to improve the SET process based on sample participants' perspectives

Reasons to improve the SET process	Participants' response based on their own perspectives
Student as evaluators	<p><i>PARTICIPANT 1200:</i> <i>"The student evaluation process I feel is highly flawed because it is so affected by the passion the students feel about their grade. I honestly, I really feel like that's huge. Students can't separate their feelings about the class and the message from the grade that they are earning. It's not very objective."</i></p> <p><i>PARTICIPANT 800:</i> <i>"The SET score is informative, when I read my student comments, it is informative for me, and it is nice to know what they are thinking. Students are stressed out during the end of the term when they are given the SET and highly emotional."</i></p>
SET questions	<p><i>PARTICIPANT 1050:</i> <i>"I would say on the student evaluation, I think we need to do better. All that it really boils down to is one or two questions. Was the teacher effective overall, or something like that? I think that's a bad way to evaluate somebody's teaching."</i></p>

Table 2.8: Reasons to improve the SET process based on sample participants' perspectives (Continued)

Reasons to improve the SET process	Participants' response based on their own perspectives
SET questions	<p><i>PARTICIPANT 1300:</i> "I think our student evaluations are not effective. I think they really tend to be more of like I said earlier, do I like this person or do I not like this person? I have a hard time looking at my student evaluations because it feels more personal than it should be, and I think it has to do with the kind of questions that they're asked. I don't think that I generally get a lot of constructive information very often in the evaluations. I would like to see questions that are a lot more content-based."</p>
Regularity	<p><i>PARTICIPANT 1300:</i> "I also wish that there was maybe more of a formal process for students to do it more often than at the end of the term because the classes that I teach, I teach once a year. If somebody says something didn't work for him or her and then I have a whole year before I can change it. I find that when I look at these evaluations, I think if I had only known, I would have had the opportunity to fix it, versus getting it after the final."</p>
Timing	<p><i>PARTICIPANT 900:</i> "The end of the quarter evaluation, it's sort of like it's too late to do anything. Still, it gives you a pretty good sense of what happened, ... That really helps over time. It helps you figure out what's effective."</p>
What it measures	<p><i>PARTICIPANT 750:</i> "I think it's really weak, I don't think it really tells us much about how instructors are actually helping students learn. I don't think there's much sure there are things associated with the actual practices of teaching but I think there are essentially zero assessment students learning how professors are actually getting their student to learn that's the weakness we have."</p>
Consistency	<p><i>PARTICIPANT 550:</i> "I think we definitely take it seriously. I think there are a few places where there are gaps, and there are... even with a really, really well-designed system, which we have, there still is room sometimes for interpretation which makes the system sometimes a little inconsistent. One professor might be praised for one thing and the next person who does a very similar thing might get a different type of feedback. That's a little troubling sometimes."</p>

Participants reported finding other voluntary and informal practices to be highly useful for improving their teaching. These included administering a survey during the course, using classroom assessment techniques, and self-reflection or assessment. For example, surveys given during the course (mid-course evaluation) help the educator know what areas he or she needs to focus on to support struggling students. Additionally, using classroom assessment techniques help the educators improve their teaching because they receive a better response rate and more feedback from their students. Eight participants have completed self-assessments and teaching portfolios which allow them to reflect on their teaching to make necessary adjustments the next time they teach the same course. Five participants found it helpful to have short discussions during meetings, and office hours with their colleagues and students about what is working and not working in the class. The educators could talk and share their strategies that they use or should be using in the classroom with other colleagues, while students can discuss what is working and not working for them.

2.5.4 Role of rewards

Recognizing and rewarding educators' performance with teaching awards may be necessary to retain effective educators and perhaps to make teaching an attractive career choice (Knapper, 2001). For example, five study participants (15%) mentioned that there is an annual teaching excellence award where their peers and students nominate an educator. But, a majority of the participants said that there is no teaching reward being offered. According to participant 1300, *"rewards or an incentive might help change the dynamic in the classroom because we would be incentivized to get good reviews so some people would really change their teaching approach."*

2.6 Discussion

Study participants have expressed interest in identifying alternative ways to evaluate teaching that could be implemented their departments. They are also interested in learning about the practices that have been adopted at other institutions for several reasons. These include identifying what teaching evaluation practices are available and useful, and to know what can be done to improve the current system.

2.6.1 Teaching evaluation practices are used for either formative or summative purposes

The purpose of evaluating teaching was found to be either formative, summative, or both, depending on the department or institution. For instance, the evaluation method was used formatively when the results are used to identify areas of improvement for an individual educator and for professional development. It is summative when the results are utilized for promotion and tenure purposes, and annual reviews. Teaching evaluation practices used for formative purposes are found to be more useful and effective by many participants because they get better response rates, and the feedback they collect allows them to make adjustments and reflect on the course and their teaching to help students learn.

2.6.2 Importance, critiques and possible improvement of SET

“Student evaluations of teaching are now administered in almost all colleges and universities in the United States, and are becoming common in other countries” (McKeachie, 2002, p. 326). Study participants discussed multiple methods used for teaching evaluation, yet it was identified that the student end-of-course evaluation remained the most commonly used. The purpose of SET is to collect data for personnel evaluation and to improve faculty’s teaching. The SET surveys are easy to conduct, inexpensive, and less time consuming than other approaches. However, some participants questioned the usefulness of the student end-of-course evaluation. For example, the questions asked on the survey are so general that they may be irrelevant to a particular class, and even if relevant, are worded so generally that they offer little guidance for improvement (McKeachie, 2002). Some participants found that students often get confused whether they were evaluating the course or the instructor (Table 2.8). There are other factors such as course size and class difficulty that should be considered because every institution, even the departments, and disciplines, are different. McKeachie (2002) suggested using a variety of forms for student evaluation of teaching that will be mostly relevant to the particular class that they are teaching. Nevertheless, a large number of our participants still believe that the student end-of-course evaluation remains an important part of teaching evaluation.

2.6.3 Approximately 85% of the evaluators who observe the classroom are not trained

The majority of the participants used classroom observations along with student end-of-course teaching evaluations. Our study showed that most evaluators, peers, and non-peers who observe the classrooms, did not receive any type of training. An evaluator's lack of training and/or not feeling qualified to evaluate teaching poses another challenge to reliability and validity of the method, as does the fact that they are based usually on only a small portion of faculty member's teaching activities, such as a single classroom observation of teaching (Yon, Burnap, & Kohut, 2016). The researchers suggest that it is essential that any evaluator must receive and continuously undergo training. It may also be beneficial to consider bringing in an expert evaluator who has knowledge about teaching evaluation and assessment for different disciplines and topics. Well-trained evaluators who know the process on how to conduct effective observation and how to use the specific peer observation instrument are ideally suited to evaluate an instructor's teaching effectiveness (Yon, Burnap, & Kohut, 2016).

2.6.4 Alignment with evidence-based teaching practices

A glaring omission from the literature is in the identification of practices and how they align with the evidence-based teaching practices (EBIPs). This omission is surprising because ideally teaching evaluation efforts would be aligned with best practices in course development and implementation, as well as with student learning. Alignment with student learning is very challenging due to several issues; high quality measures of student learning are not consistent across courses, measures of student learning vary widely across courses and offerings of a particular course, and determining any kind of correlational or causal relation between a particular aspect of practice, and its effect on learning are nearly impossible at a large scale. Doing so would require significant attention to practices, learning, improvement of practices, and student work, and the relationship between each of these. However, there is an abundance of research that supports the efficacy of educational practices with student learning, as well as other outcomes (Prince, 2004). Additionally, these practices align with what is reasonably observable with common teaching evaluation practices.

We suggest that these observable core components of a course offering are active learning in the classroom; the alignment of outcomes, activities, and assessments, and the frequency and quality of feedback to students. Active learning has been shown to be effective for student learning through multiple sources (Prince, 2004). Active learning is operationalized in detail in the ICAP (Interactive, Constructive, Active, and Passive) framework (Chi & Wylie, 2014), where it is shown that I, C, A, and P learning environments are decreasingly effective to student learning. The alignment of outcomes, activities and assessments are supported as a best practice in the educational literature. Two examples are constructive alignment and backward design (Biggs, 1996; Wiggins & McTighe, 1998). Both are based on using intentional instructional design to align the outcomes of a course with the types of learning activities and subsequent assessment measures. The focus of these approaches is to ensure that course outcomes are reflected and enacted in every facet of the course. This alignment can be observed through the course syllabus and examples of student assignments and assessments and how they align with outcomes. There is vast literature supporting the importance of the quality and quantity of feedback to students. For example, Juwah et al. (2004) posit that student learning is enhanced when formative feedback measures are used as teachers can often adjust their lesson delivery to meet the needs of students. According to Shute (2008), “the main aim of formative feedback is to increase student knowledge, skills and understanding in some content areas or general skill” (p.156).

Ideally, both formative and summative teaching evaluation practices would be aligned with these best practices. However, our findings showed that there is a lack of alignment between the current evaluation practices and EBIPs previously defined. For instance, our study showed that the majority of the classroom observers and evaluators of classroom materials receive minimal or no training. The classroom observation process usually only consists of a single classroom observation with one observer, and only half of the participants in this study indicated that classroom materials are evaluated. Based on these results, active learning in the classroom is probably not considered or impossible to observe during this process. Classroom observers often focus on the educator’s strengths and weaknesses such as their teaching skills, teaching methods, and content knowledge, and

less on the connection between these skills and student-learning outcomes. Additionally, only eight participants revealed using standard forms or checklists in their department during classroom observation. There is also no protocol for document analysis to help the evaluators know what and how to evaluate the evidence they gather. Without any standard procedures and proper training, this could lead to an educator getting very different scores and feedback depending on who observed him or her. The evidence gathered without using standard-based measures during document analysis and without proper training for observing teaching might be based on one individual's judgment and does not provide evidence that might help an educator improve his or her teaching (Goe, Biggers, & Croft, 2012).

Despite our findings, we would like to provide descriptions of how currently used teaching evaluation practices might be better aligned with EBIPs, see Table 2.9.

Table 2.9: Aligning current teaching evaluation practices with EBIPs

Teaching evaluation practices	Description of how currently used teaching evaluation practices could be aligned with EBIPs
Classroom observations	The observer and the educator being evaluated must discuss the evidence obtained in the observations. The conversation should incorporate evidence of student learning so that educators focus primarily on whether students are learning rather than focusing only on instruction (Goe et al., 2012).
Evaluation of classroom materials	Collecting classroom artifacts, such as assignments and resulting student work, can provide evidence of students' understanding of particular concepts that could lead to instructional design, and strategies to improve student-learning outcomes (Goe et al., 2012).
Student end-of-course evaluations	Designing and using the survey questions to collect specific information about the learning environment and the effectiveness of instructional practices.
Formative assessments (Also known as classroom assessment techniques)	This type of assessment consists of activities designed for helping educators assess and measure students understanding of the course content. The educator can reflect on the feedback gathered from their students. The two types of formative assessment are formal and informal techniques. Formal techniques include in-class activities, quizzes, and online learning modules. For example, if students do poorly on a quiz, then an educator might examine their instructional design or teaching strategies. In-class activities, where students work in pairs or small groups to solve classroom problems, can help students who are struggling. This approach can help educators to revise concept delivery or to reinforce important concepts in a different manner to help students. Informal techniques include written reflections, poll/surveys, and wrappers. For example, some of our participants used poll everywhere (an audience response system application that uses mobile technology to assess students understanding of particular concepts or misconceptions) to check what is working with students and what they find confusing with the lecture. Another participant used clicker questions to get instant feedback with questions to test students' understanding.
Summative assessments	Using these techniques could help educators improve their course, leading to improved student learning. These include collecting exams, projects, and presentations, and teaching portfolios. For instance, teaching portfolios provide examples of student works or evidence of participation, and these could be submitted for promotion purposes.

Although there is substantial interest in improving teaching evaluation practices in engineering programs, it was identified that the existing practices are still somewhat different from the identified best practices in the literature.

2.7 Conclusion

Based on the research questions that guided this research study, the following conclusions can be made:

2.7.1 Student end-of-course evaluation remains the most popular choice for evaluating teaching.

There are a variety of teaching evaluation practices available for implementation. However, this study showed that student end-of-course evaluation is the most widely used because it is fairly easy to administer, inexpensive, and requires less time than the other teaching evaluation practices. A large number of engineering programs also conduct classroom observations as part of teaching evaluation, but the lack of rubric use and little training that might help observers focus on the connection between teaching strategies and student learning is problematic. In addition, classroom materials may be requested and collected from an educator when he or she undergoes an observation, but as in the case of classroom observations using these documents to evaluate or improve teaching requires attention to opportunities for student learning. Student mid-course evaluation was reported by participants to be very useful for gathering feedback from their students to help improve their teaching and to assess student learning. However, this method is only used on a voluntary basis and does not usually go to educator's personnel file. If the purpose of teaching evaluations is primarily to improve teaching, then perhaps all educators should participate in this practice. The choice of methods used depends heavily on the purpose of teaching evaluation, which may vary in terms of department or institution objectives.

2.7.2 Teaching evaluation methods are used for making formative and/or summative decisions.

One important finding our study highlighted was the fact that the methods used to evaluate teaching were either used for formative or summative evaluation. All but one of the teaching evaluation methods discussed in this study were used for both formative and summative purposes. This study demonstrated that student mid-course evaluations are used in all the institutions but only for formative evaluation. In addition, there was an almost equal split between formative and summative use for student-end-of course evaluation while the evaluation of classroom materials was mostly used for summative assessment. The choice of method used was dependent on the purpose of teaching evaluation. For example, when the purpose is to assess the quality and magnitude of student learning or for the faculty member to use data to improve teaching, then the methods were used formatively (e.g. student mid-term evaluations). In cases where the evaluation was ultimately used in the institution's rewards system such as promotion and tenure or continued employment for non-tenure line faculty, then the method was used summatively.

2.7.3 Formative evaluations and assessments were reported to be useful by educators.

Student end-of-course evaluation and classroom observation are found to be useful practices for teaching evaluations. They are highly effective and beneficial to educators when they are used as a means of improvement instead of as the primary tool for making decisions about promotion and tenure, or rehiring. The participants reported that enactment of both these practices should still be improved in order to be fully effective. Formative evaluation practices that are used informally and voluntarily such as student mid-course evaluation, self-reflection or self-assessment, were found to be beneficial to educators because they are frequent and lead educators to reflect on their teaching. Since teaching is the primary or a critical part of a faculty member's work, an educator's commitment to using formative assessment to improve teaching might be what is summatively assessed to make promotion or rehiring decision.

This work answers the call for research focused on teaching related issues in engineering education. As the discipline continues to grow and more formalized programs are established, it is necessary to link teaching strategies reported as effective to enhance student learning to methods of assessing quality teaching in engineering. In this study, we sought to explore the range of practices used to conduct teaching evaluation in engineering departments across the country. However, the number of participants were not evenly distributed across the three types of higher education institutions. This remains an area in which further investigation could be conducted. This would facilitate comparison of practices within and across institutions with varying value and focus on quality teaching.

2.8 Acknowledgment

This material is based upon work supported by the National Science Foundation under Grant No. DUE-1347817. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

2.9 References

- Baxter, P., & Jack, S. (2008). Qualitative Case Study Methodology : Study Design and Implementation for Novice Researchers *Qualitative Case Study Methodology : Study Design and Implementation*, 13(4), 544–559.
- Berk, R. A. (2005). Survey of 12 Strategies to Measure Teaching Effectiveness. *International Journal of Teaching and Learning in Higher Education*, 17(1), 48–62.
- Biggs J. (1996). Enhancing Teaching Through Constructive Alignment. *Higher Education*, 32, 347–364.
- Brandenburg, D. C., Braskamp, L. A., & Ory, J. C. (1984). *Evaluating Teaching Effectiveness: A Practical Guide*.
- Broder, J. M., & Taylor, W. J. (2015). Teaching Evaluation in Agricultural Economics and Related Departments. *American Journal of Agricultural Economics*, 76(1), 153–162
- Campbell, J., & Bozeman, W. (2007). The value of student ratings: Perceptions of students, teachers, and administrators. *Community College Journal of Research and Practice*, 32(1), 13–24. <http://doi.org/10.1080/10668920600864137>

- Chen, Y., & Hoshower, L. B. (2003). Student Evaluation of Teaching Effectiveness: An Assessment of Student Perception and Motivation. *Assessment & Evaluation in Higher Education*, 28(1), 71–88.
- Chi, M. T. H., & Wylie, R. (2014). The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational Psychologist*, 49(4), 219–243. <http://doi.org/10.1080/00461520.2014.965823>
- Creswell, J. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. SAGE (4th ed.). <http://doi.org/10.1007/s13398-014-0173-7.2>
- Crotty, M. (1998). *The Foundations of Social Research: Meaning and Perspective in the Research Process*.
- Dedoose Version 7.1.3. (2015). Web Application for Managing, Analyzing, and Presenting Qualitative and Mixed-Method Research Data.
- Dressel, P. (1960). Teaching , Learning , and Evaluation. *Improving College and University Teaching*, 8(1), 11–15.
- Eddy, S. L., Converse, M., & Wenderoth, M. P. (2015). PORTAAL: A Classroom Observation Tool Assessing Evidence-Based Teaching Practices for Active Learning in Large Science, Technology, Engineering, and Mathematics Classes. *CBE - Life Science Education*, 14, 1–16. <http://doi.org/10.1187/cbe-14-06-0095>
- Ferguson, R. (2007). Constructivism and Social Constructivism. In *Theoretical Frameworks for Research in Chemistry/Science Education* (pp. 28–49).
- Fink, L. D., Ambrose, S., & Wheeler, D. (2005). Becoming a professional engineering educator: A new role for a new era. *Journal of Engineering Education*, 94(January), 185–194. <http://doi.org/10.1002/j.2168-9830.2005.tb00837.x>
- Goe, L., Biggers, K., Croft, A. (2012). *Linking Teacher Evaluation to Professional Development: Focusing on Improving Teaching and Learning*.
- Goldschmid, M. L. (1978). The evaluation and improvement of teaching in higher education. *Department of Psychology and Higher Education, Swiss Federal Institute of Technology, Lausanne, Switzerland*, 7, 221–245.
- Henderson, C., Turpen, C., Dancy, M., & Chapman, T. (2014). Assessment of teaching effectiveness: Lack of alignment between instructors, institutions, and research recommendations. *Physical Review Special Topics - Physics Education Research*, 10(1), 1–20. <http://doi.org/10.1103/PhysRevSTPER.10.010106>
- Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education*, 39(4), 372–400.

- Iqbal, I. (2013). Academics' resistance to summative peer review of teaching: questionable rewards and the importance of student evaluations. *Teaching in Higher Education*, 18(5), 1–13. <http://doi.org/10.1080/13562517.2013.764863>
- Johnson, T. D., & Ryan, K. E. (2000). A comprehensive approach to the evaluation of college teaching. *New Directions for Teaching and Learning*, (83), 109–123. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=tfh&AN=9178975&lang=es&site=ehost-live>
- Juwah, C., Macfarlane-Dick, D., Matthew, B., Nicol, D., Ross, D., & Smith, B. (2004). Enhancing student learning through effective formative feedback. *Higher Education*, (68), 1–41. Retrieved from http://www.heacademy.ac.uk/resources/detail/resource_database/id353_effective_formative_feedback_juwah_et al
- Kite, M. E. (2012). *Effective Evaluation of Teaching: A Guide for Faculty and Administrators*.
- Knapper, C. (2001). Broadening Our Approach to Teaching Evaluation. *New Directions for Teaching and Learning*, 2001(88), 3–9. <http://doi.org/10.1002/tl.32>
- McKeachie, W. . (2002). *McKeachie's Teaching Tips: Strategies, Research, and Theory for College and University Teachers*.
- Mohanty, G., Gretes, J., Flowers, C., Algozzine, B., & Spooner, F. (2005). Multi-method evaluation of instruction in engineering classes. *Journal of Personnel Evaluation in Education*, 18(2), 139–151. <http://doi.org/10.1007/s11092-006-9006-3>
- Morgan, P. (2008). The Course Improvement Flowchart : A description of a tool and process for the evaluation of university teaching. *Journal of University Teaching & Learning Practice*, 5(2), 1–14. Retrieved from <http://ro.uow.edu.au/jutlp/vol5/iss2/2>
- National Research Council. (2003). *Evaluating and improving undergraduate teaching in science, technology, engineering, and mathematics*. National Research Council. <http://doi.org/10.1002/bmb.2003.494031059998>
- Nygaard, C., & Belluigi, D. Z. (2011). A proposed methodology for contextualised evaluation in higher education. *Assessment & Evaluation in Higher Education*, 36(6), 657–671. <http://doi.org/10.1080/02602931003650037>
- Ohanlon, J., & Mortensen, L. (1980). Making Teacher Evaluation Work. *Journal of Higher Education*, 51(6), 664–672. <http://doi.org/10.2307/2943969>
- Ory, J. C. (2000). Teaching Evaluation: Past, Present, and Future, (83), 13–18.
- Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education- Washington*, 93(July), 223–232. <http://doi.org/10.1002/j.2168-9830.2004.tb00809.x>

- Reeves, T. C. (2006). How do you know they are learning?: the importance of alignment in higher education. *Int. J. Learning Technology J. Learning Technology*, 2(4), 294–309. <http://doi.org/10.1504/IJLT.2006.011336>
- Shute, V. aleri. J. (2008). Focus on Formative Feedback. *Review of Educational Research*, 78(1), 153–189. <http://doi.org/10.3102/0034654307313795>
- Wankat, & Oreovicz. (2009). Teaching engineering. *IEEE Engineering Management Review*, 37(1), 1–9. <http://doi.org/10.1109/EMR.2008.2012072>
- Wiggins, G., & McTighe, J. (1998). What is backward design? In *Understanding by Design* (pp. 7–19). <http://doi.org/10.1016/j.cie.2006.03.005>
- Yon, M., Burnap, C., & Kohut, G. (2016). Evidence of Effective Teaching-Perceptions of Peer Reviewers, *50*(3), 104–110.

Chapter 3 – Investigating Current Approaches to Assessing Teaching Evaluation in Engineering Departments

Keisha A. Villanueva, Shane A. Brown, Nicole P. Pitterson, and Ann Sitomer

2016 IEEE Frontiers in Education Conference (FIE)

Camino Real El Paso

101 South El Paso Street

El Paso, TX, USA

Date: Accepted July 2016

3.1 Abstract

The nature of teaching evaluation is, more often than not, determined primarily by the results of surveys based on students' evaluation and, sometimes, peer evaluation ratings. However, previous work has highlighted that in most cases student evaluation surveys are the main and/or only method used to evaluate teaching. This led to the critique of using student evaluation surveys as insufficient as a stand-alone method of evaluating teaching. These criticisms are based on the fact that instruments designed for students to evaluate teaching commonly consist of the same set of generic closed-ended questions and lack the ability to appropriately judge the multi-dimensional nature of teaching and learning. Additionally, the focus of teaching evaluation is theorized to differ based on the type of institution (which are doctoral universities, baccalaureate and associate colleges). With this in mind, this study seeks to explore current approaches to evaluate teaching in engineering departments. Based on the call for multiple approaches to evaluating teaching in higher education this research seeks to answer the following questions "what are the current approaches to assessing teaching evaluation and how are these approaches similar or different based on institution type?" As part of a larger study, this paper will describe the teaching evaluation practices within three types of institutions aforementioned. The goal of this study is to provide stakeholders with a repository of teaching evaluation practices. The authors also posit that having a broad spectrum of teaching evaluation approaches is highly valuable in improving teaching and student learning in the engineering community.

Keywords – evaluation of teaching, formative and summative assessments, higher education

3.2 Introduction

Teaching evaluation is used in higher education to offer valuable feedback to instructors/professors on the quality of their teaching as well as ways in which they can improve their performance so as to ensure optimal student learning (Ory, 2000). The importance of conducting formative and summative assessments of teaching lies in the following three core benefits teaching evaluation provide: 1) feedback aimed at improving teaching for increased student learning, content covered in courses and overall classroom design, 2) measure of teaching effectiveness that can be used in the promotion and tenure process and, 3) information for future revision to courses and programs so as to effectively meet the need of the students who enroll in these courses/programs (Kite, 2012; National Research Council, 2003; Ory, 2000).

Research focused on the evaluation of teaching suggests student evaluation surveys are the most commonly used methods and in some cases the only method (Chen & Hoshower, 2003). Student evaluation surveys are reported as having high levels of validity since students are best positioned to judge particular aspects of teaching and the classroom in general. However, student evaluation surveys have been critiqued as an invalid and unreliable method of evaluating teaching when used as the only assessment method (Felder & Brent, 2004). This critique has fueled the call for multiple sources of evidence in assessing teaching effectiveness that has become stronger over the last decade (National Research Council, 2003). This is based on the fact that student evaluation surveys usually consist of closed-ended, generic type questions which, while intent on making the process fairly easy for students to complete, on their own are not enough to capture all facets of teaching effectiveness. The lack of open-ended or explanatory approaches, such as interviews or focus groups, limit the possibility of collecting student explanations of their perceptions of instructional strategies chosen to enhance the learning process (Hoyt & Pallett, 1999; Kite, 2012; National Research Council, 2003).

Employing additional methods to evaluate teaching effectiveness create the opportunity to truly assess all the interacting pieces of teaching and learning. This is especially necessary since “learning and teaching are multi-dimensional constructs, therefore an evaluation of

teaching should involve a multi-dimensional approach” (Kite, 2012, p. 30). Consequently, there is a need for additional methods of evaluation so that the breadth of teaching effectiveness can be assessed. Additionally, there are some aspects of teaching and learning students are incapable of assessing such as instructor commitment to teaching, decisions made concerning course design and class structure, and the instructor’s rationale for instructional decisions made in the classroom (Hoyt & Pallett, 1999; Saroyan & Amundsen, 2010).

This study examines teaching evaluation processes in engineering departments and whether these processes are consistent across the different institution types that diverge in their focus on teaching and learning. Using the Carnegie classification of institutions, this study was conducted to investigate the various teaching evaluation processes at doctoral universities, baccalaureate and associate colleges. From our study, stakeholders will become aware of the different evaluation techniques that can be used in teacher evaluation at various types of institutions. The authors posit the teaching evaluation system in engineering departments can benefit from having a wealth of approaches to assess the various aspects of teaching. In addition, this work will open up the conversation about how teaching evaluation can be reformed to improve the quality and magnitude of student learning.

3.3 Perspectives from literature

Teaching evaluation is described as the process whereby the quality of teaching is assessed (Kite, 2012). This measure of assessment can be conducted using formative and/or summative approaches. In higher education, formative assessments of teaching are focused on providing instructors/professors with information that can help them to improve their teaching. Formative assessment provides feedback on content delivery, design of classroom, course materials and, in some cases, teacher preparedness. Summative assessment, on the other hand, is used in the promotion and tenure process whereby information collected by the evaluation method is used to measure overall performance of the instructor/professor (Hoyt & Pallett, 1999; National Research Council, 2003; Wachtel,

1998). At associate-degree granting colleges that rely on a large number of adjunct faculty, these summative evaluations are also used in decisions to rehire (Campbell & Bozeman, 2007). However, the primary focus of teaching evaluation should be the educator's ability to create the conditions necessary to optimize student learning. It is on this premise that some researchers seek to determine what is "good" teaching. For example, the work of Chickering and Gamson (1987) on good practice in teaching and learning emphasize the following seven principles:

1. encourages contact between students and faculty,
2. develops reciprocity and cooperation among students,
3. encourages active learning,
4. gives prompt feedback,
5. emphasizes time on task,
6. communicates high expectations,
7. respects diverse talents and ways of learning.

Based on these seven principles previous work has categorized "good teaching" into three distinct facts:

1. creation of a student-centered learning environment
2. appropriate assessment of student learning and,
3. sufficient preparation for teaching and learning.

These three categories therefore require a multi-method approach to teaching evaluation. "The key to effective teaching evaluation is to collect data from multiple sources (triangulation) making sure that all education-related activities are rated by the people best qualified to rate them" (Felder & Brent, 2004, p. 38).

3.3.1. Research on the purpose of teaching evaluation

From previous research on the evaluation of postsecondary teaching, evaluation of teaching is conducted in disparate silos, mostly by researchers in higher education or educational leadership (Campbell & Bozeman, 2007; Iqbal, 2013; Nygaard & Belluigi, 2011; Saroyan & Amundsen, 2001; Smith, 2012). In addition, some scholars with an expertise in evaluation in general (Nygaard & Belluigi, 2011) and researchers working within

disciplinary practices such as social work (Kealey, 2010) or engineering (Mohanty et al., 2005) assess teaching and teacher effectiveness. Although scholars researching postsecondary evaluation of teaching all build upon previous work, our literature review did not uncover a clear research agenda that is currently being pursued to help us explain the relationship between evaluation of teaching and improvements in postsecondary teaching. However, this is not surprising given the variety of research interests of those engaged in this inquiry.

As mentioned above, these scholars highlight that evaluation of teaching often serves two purposes, data for improving teaching or as a summative evaluation. However, conflating these two purposes is rarely interrogated, with a few notable exceptions (Mohanty et al., 2005; Nygaard & Belluigi, 2011). The literature does highlight three primary categories of strategies for evaluating postsecondary teaching: student evaluation of teaching (SETs), peer evaluation of teaching (POT), and self-evaluation (Morgan, 2008; Saroyan & Amundsen, 2001). Within each of these categories, there are a variety of quantitative and qualitative tools. SETs may consist of numerically scored questionnaires used either formatively or summatively, open-ended responses to items on a questionnaire, or be gathered using student focus groups (Morgan, 2008). Research on the peer evaluation of teaching underscores a number of issues, for example, culture or personal bias (Yon et al., 2016), lack of training in the observation and evaluation of teaching (Smith, 2012), the need to observe more than one class meeting or use multiple peer evaluators and to consider other aspects of the course, such as activities and assessments (Iqbal, 2013). Self-evaluation also takes several forms, from completing checklists (Morgan, 2008) to creating portfolios that significant reflection on one's teaching (Felder & Brent, 2004; Saroyan & Amundsen, 2001). Several studies compared two of these strategies, but few compared the various tools within particular strategies. Further, our review uncovered no research that compares self-evaluation strategies. Saroyan and Amundsen (Saroyan & Amundsen, 2001) is a frequently cited review of research that both summarizes past findings and makes recommendation for best practices. Significant from our perspective is the idea that both SET and POT are observations of behavior and cannot reveal educators' underlying

rationale for instructional decisions (Grossman, 1991). In addition, these authors, as well as Kealey (2010) and Nygaard and Belluigi (2011) acknowledge that teaching is a complex and contextualized activity and that development is rarely linear. To this end it is necessary to employ multiple approaches to evaluate teaching that goes beyond students' perception. The purpose of this paper is therefore to examine the various approaches to teaching evaluation used in the engineering departments/units of different institutions. The research questions we seek to answer are:

- a. What are the current approaches used in engineering departments/units to assessing teaching evaluation?
- b. How are these approaches similar or different based on institution type?

In the engineering context it is especially necessary to explore the various methods of teaching evaluation being used in different institutions. This is an important area to investigate primarily because of the lack of studies on this topic. In addition, the recommendation made by the National Research Council's book on *Evaluating and Improving Undergraduate Teaching in Science, Technology, Engineering and Mathematics* (National Research Council, 2003) speaks to "the collection of different kinds of evidence that can be used to determine whether faculty and departments are indeed promoting student learning" (p. 116).

3.4 Methodology

3.4.1 Study Design

This study guided by a descriptive case study with multiple embedded units approach which is situated within the constructivist research paradigm. "A descriptive case study enables the researcher to describe an intervention or phenomenon and the real-life context in which it occurs" (Baxter & Jack, 2008, p. 549). The units of analysis are the various institutions from which our data were collected. The data were collected using an exploratory sequential mixed-method design to answer the research questions. This approach involved collecting both quantitative and qualitative data (Creswell, 2014). In

this paper we will report on the qualitative data collected. The preliminary analysis of these data was used to guide the collection of quantitative data.

3.4.2 Participants (Institutional Overview)

Educators with varying responsibility and ranks such as course instructors, department heads, and program coordinators in engineering programs at schools across the U.S., were contacted and asked to participate through email in a semi-structured interview. Snowball sampling was used to identify participants. One author provided contact information for individuals at some institutions who were invited to be interviewed and asked to identify further colleagues they believed would be interested in this research. A total of 34 participants were interviewed once in this study. Almost half of the participants were department chairs ($n=15$). Table 3.1 shows the number of participants and their role in their department and/or institution. The participants who were interviewed in this study came from 31 different schools. There were few participants who were identified to be from the same school, however they were not from the same engineering department.

Table 3.1: Participants' Department Role

Participants' department role	<i>n</i>	Percent
Department Chair	15	44%
Associate Department Chair	1	3%
Full Professor	1	3%
Associate Professor	6	18%
Assistant Professor	3	9%
Engineering Instructor	6	18%
Program Coordinator	2	6%
Total	34	

3.4.3 Data Collection

The data collected and used for this stage of our project were gathered using semi-structured interviews. An interview protocol was developed to query each participant's department practices in conducting teaching evaluations. However, many of the participants also reported their *own* practices in conducting teaching evaluations. The interview protocol was developed based on a literature search of teaching evaluation

procedures. The duration of each interview ranged between 15-60 minutes. During the interview, copies of any documents related to teaching evaluation such as student end of course questionnaires, letters, memos, and questionnaires used for student midterm evaluation were requested from the participants at the end of each interview. These materials were requested and collected to help with the comparison and analysis for this study. The documents collected from some participants were kept confidential and any identifiers such as name, school name, and department name are edited and replaced by codes. However, not all participants were able to share any or all of the documents and it was not possible to perform a comparative document analysis. The interviews were audiotaped and notes were also taken.

3.4.4 Data Analysis

Each interview was transcribed either via professional transcription company or by a graduate researcher. The text documents were uploaded into Dedoose a web application for managing, analyzing, qualitative and mixed methods research data. An excel spreadsheet was also used to tabulate and keep track of the data obtained from Dedoose. The data were analyzed in two phases. In phase one, all evaluation techniques were grouped based on similarity in their execution; in phase two, the evaluation techniques were categorized based on institution type to identify similarities and differences. However, an unequal number of participants were interviewed from each institution type. For example, half ($n=17$) of the total participants came from research-intensive universities, 11 of the participants were from teaching-focused universities and the rest came from the community colleges ($n=6$). Therefore, this comparison may be biased due to the number of sample sizes from each institution type. However, the authors' intention was to see if there was any pattern that might emerge by grouping the participants within the three institution types.

3.5 Results

3.5.1 What are the current approaches to assessing teaching evaluation?

We found a total of five approaches used in the evaluation of teaching. These approaches are summarized in Table 3.2 and a brief discussion of how each assessment is conducted.

Table 3.2: Summary of methods used to evaluate teaching

Methods to evaluate teaching	How assessment is conducted
Student end-of-course evaluation	Usually administered through the use of closed-ended survey items, students are instructed to rate a given prompt.
Student mid-course evaluation	Considered an informal method of assessment, teachers gauge students learning midway through the course by gathering information about students' perception of the instructor's teaching and the progression of the course. Instructors may use the feedback gathered to adjust the delivery during the second half of the course.
Student exit evaluation	Graduating students are surveyed using predetermined questions to gather information about the effectiveness of courses and educational programs.
Classroom observation by peers	The purpose of this method can be twofold; it can be used to improve teaching, as well as inform promotion decisions. A peer, usually a senior department member, visits the classroom of the faculty member being evaluated. The peer may also evaluate course artifacts; and self-assessment documents such as a teaching portfolio/dossier. Collective feedback is given to the faculty with constructive criticism.
Classroom observation by administrators	This method is usually used to gather information about in-class practices associated with teaching and learning. Classroom observations are usually conducted by personnel who are knowledgeable about pedagogical strategies or the content being taught. This observation is sometimes conducted by an external evaluator such as someone who works with the institutions' Center for Teaching and Learning.

One key finding on the use of these methods is the fact that some of the methods summarized in Table 3.2 are used formally (at the departmental or unit level) as well as informally (at the faculty-to-faculty level). In addition, we also found that these methods

are sometimes used for formative assessment, to provide feedback to improve teaching and for summative assessment, to make decisions about faculty promotion and tenure. Table 3.3 illustrates this information. From the table it can be seen that student end-of-course evaluation of teaching is the most utilized assessment approach with a frequency usage of 100%. Classroom observation is the second highest with a frequency count of 85% while peer evaluation has a frequency usage of 79%. Student mid-course evaluation also has a high usage of 74% while student exit evaluation was reported at 62%. In some interviews, participants mentioned additional methods such as teaching portfolios, alumni evaluation and in-class informal techniques to gauge students' perception. However, the use of these methods to evaluate teaching was not explicitly discussed.

Table 3.3: Teaching Evaluation Practices Utilization Summary

Teaching Evaluation Practices	Utilization by Participants	
	<i>n</i>	Percent
Student end-of-course evaluation	34	100%
Student mid-course evaluation	25	74%
Student exit evaluation	21	62%
Classroom observation by peers	27	79%
Classroom observation by administrators	29	85%
Total number of participants in this study	34	

3.5.2 How are these approaches similar or different based on institution type?

Our study consisted of 34 institutions of which 17 (approximately 50%) were research-intensive universities, 11 (approximately 32%) were teaching focused institutions while six (approximately 18%) were community colleges. Figure 3.1 summarizes this distribution.

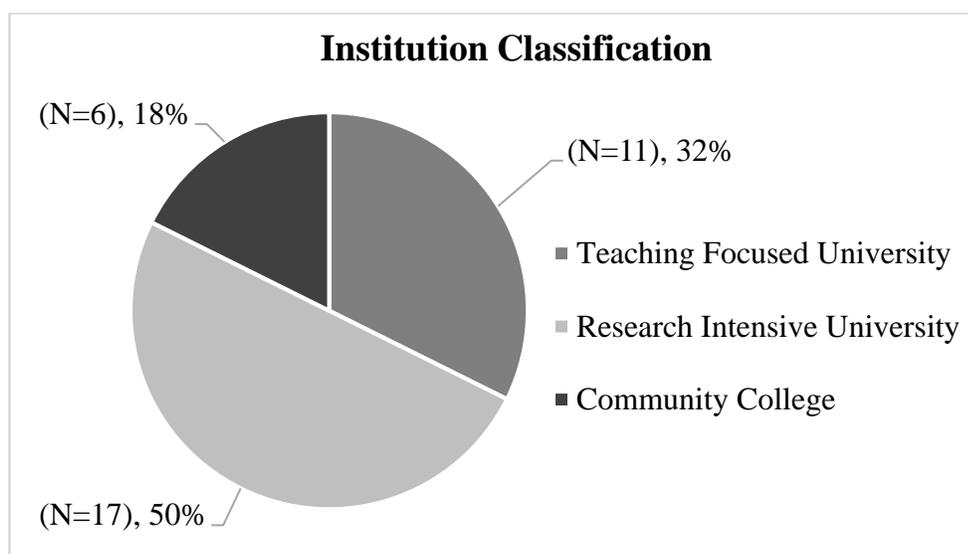


Figure 3.1: Distribution of institution type used in our study

Across the institution types investigated in this study, we found similar assessment approaches. For example, student end-of-course evaluation was used 100% of the time at all three institution types investigated. Similarly, classroom observation, peer and student exit evaluation had comparable usage across all three institution types. The use of student mid-course evaluation was the only major difference; community colleges had the highest usage of this form of teaching evaluation at 83%, teaching focused universities had an 82% usage and research intensive universities a 59% usage. The utilization of teaching evaluation approaches across institution types are illustrated in Table 3.4.

Table 3.4: Teaching Evaluation Practices Utilization Classified by Institution-type

Teaching evaluation practices by institutions	Utilization by Participants		
	Teaching University	Research University	Community College
Student end-of-course evaluation	100%, (<i>n</i> =11)	100%, (<i>n</i> =17)	100%, (<i>n</i> =6)
Student mid-course evaluation	82%, (<i>n</i> =9)	59%, (<i>n</i> =10)	83%, (<i>n</i> =5)
Student exit evaluation	64%, (<i>n</i> =7)	59%, (<i>n</i> =10)	50%, (<i>n</i> =3)
Classroom observation by peers	91%, (<i>n</i> =10)	71%, (<i>n</i> =12)	83%, (<i>n</i> =5)
Classroom observation by administrators	82%, (<i>n</i> =9)	88%, (<i>n</i> =15)	83%, (<i>n</i> =5)
Total number of participants by institution, (<i>n</i>)	11	17	6

3.6 Discussion

The methods used to evaluate teaching unearthed by our study did not appear to have significant differences in how they were utilized at the three institution types. For example, we found that student evaluation of teaching similar frequency of usage in all institution types. Our findings also indicate that teaching focused institutions and community colleges had more frequent usage of all the methods used to evaluate teaching with the exception of classroom observations and student exit evaluation. One important finding our study highlighted was the fact that the methods used to evaluate teaching were either used for formative or summative evaluation. The choice of method used was dependent on the purpose of teaching evaluation. For example, when the purpose is to assess the quality and magnitude of student learning or for the faculty member to use data to improve teaching, then the methods were used formatively (e.g. student mid-course evaluations). In cases where the evaluation was ultimately used in the institution's rewards system such as promotion and tenure or continued employment for non-tenure line faculty, then the method was used summatively.

Additionally, formative and summative approaches and how they are used are subjected to the policies and practices of schools/departments. However, to increase the validity and reliability of the teaching evaluation process it is necessary to combine multiple methods. This finding is supported by the NRC (2003) report on *Evaluation and Improving Undergraduate Teaching in Science, Technology, Engineering and Mathematics* which states "assessment that is based on a single teaching activity e.g. classroom presentation or depends on information from a single source such as student evaluation ratings is less reliable, useful and valid than an assessment of an instructor's strengths and weaknesses that is based on multiple sources" (p. 51). Though our participants discussed multiple methods used for teaching evaluation, student evaluation surveys remain one of the main methods used in summative assessment of teaching evaluation. While there are benefits to having students report their perception of teaching and decisions made by the professor/instructor about their learning using multiple methods of assessment create a more complete picture of the many facets of teaching. Additionally, we were unable to

ascertain that when multiple methods are used in teaching evaluation, are these are weighed in the same manner? However, our participants have expressed interest in identifying alternative ways to evaluate teaching that could be used in their departments. Although there is substantial interest in improving teaching evaluation practices in engineering programs, it was identified that the existing practices are still somewhat different from the identified best practices in the literature. One such practice is an argument made by several researchers (Felder & Brent, 2004; Saroyan & Amundsen, 2001; Wachtel, 1998) to utilize multiple forms of evaluation to teaching that also attend to the many facets of teaching. For example, in Felder and Brent's model (2004), students would evaluate practices related to the classroom such as instruction, assessment methods, advising and mentoring provided by the instructors. Peers would be responsible for assessing instructional pedagogy and the appropriateness of their use, course documents and feasibility of stated learning outcomes. The instructor/professor being evaluated would reflect and rate their own philosophy of teaching, how well they complete their role as an instructor and how are they designing the students learning experience to meet their own learning outcomes. Documentation of these different areas of evaluation would then help to develop a teaching portfolio, which could be used as a summative artifact for promotional decisions. The advantage of using this model is that multiple perspectives are included and the process of teaching evaluation becomes structured with all stakeholders having some level of investment.

3.7 Conclusions

Based on the research questions that guided this student the following important conclusions can be made:

3.7.1 Student evaluation of teaching surveys are still the most prominent teaching evaluation approach

We found that despite the call for increasing variety of methods for evaluating an educator's teaching, student evaluation of teaching is still widely used because it is fairly easy and less demanding than other approaches. The second most used approach is classroom observations conducted by administrators and peers. We found that these two

approaches to teaching evaluation are mandated by most of the institutions explored. While we acknowledge that to implement other teaching evaluation approaches could be time consuming and additional service commitment for faculty, we still argue that the use of multiple approaches is beneficial. Being able to triangulate various approaches allows for a complete assessment of all the interacting factors associated with teaching. However, the methods or practices chosen to use should be dependent on the purpose of the evaluation. The issue of time constraints could be addressed by combining a set of practices/techniques discussed in the literature and result sections. In addition, new practices such as student focus group interviews using present and past students of a particular course can be added to student evaluation and peer observation approaches as a third data point.

3.7.2 Level of usage of teaching evaluation approaches were similar among teaching focused institutions and community colleges

A key finding was the similarity in usage of teaching evaluation methods at teaching focused and community colleges. However, this finding is hardly surprising as emphasis on “good teaching” varies based on institution type. In cases where teaching evaluation determines contract renewal and job promotion, such as at the community college level, the importance of assessing all aspects of teaching is of utmost importance. Similarly, at teaching-focused institutions, student learning and achievement is a primary goal of instruction. This being the case it is necessary to evaluate the teaching processes that encompass class preparation, information dissemination and student assessment. These different emphases on the teaching and teaching evaluation dictate a multi-faceted approach. In the next phase of this work, a Qualtrics survey will be administered to participants to collect additional information based on findings from our preliminary analysis. However, one key limitation of this study was the disproportionate distribution of institution type. For example, half of the institutions studied were research-intensive while only six were community colleges. A recommendation for future study is the exploration of more community colleges and other teaching focused institutions. Another area of exploration could be to examine based on institution type what practices are mandated or voluntarily done at the discretion of departments or units

3.8 Acknowledgment

This material is based upon work supported by the National Science Foundation under Grant No. DUE-1347817. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

3.9 References

- Baxter, P., & Jack, S. (2008). Qualitative Case Study Methodology : Study Design and Implementation for Novice Researchers *Qualitative Case Study Methodology : Study Design and Implementation*, 13(4), 544–559.
- Campbell, J., & Bozeman, W. (2007). The value of student ratings: Perceptions of students, teachers, and administrators. *Community College Journal of Research and Practice*, 32(1), 13–24. <http://doi.org/10.1080/10668920600864137>
- Chen, Y., & Hoshower, L. B. (2003). Student Evaluation of Teaching Effectiveness: An Assessment of Student Perception and Motivation. *Assessment & Evaluation in Higher Education*, 28(1), 71–88.
- Chickering, A. W., & Gamson, Z. F. (1987). Seven principles for good practice in Undergraduate Education. *American Association of Higher Education Bulletin*, 39, 3–7. [http://doi.org/10.1016/0307-4412\(89\)90094-0](http://doi.org/10.1016/0307-4412(89)90094-0)
- Creswell, J. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. SAGE (4th ed.). <http://doi.org/10.1007/s13398-014-0173-7.2>
- Felder, R. M., & Brent, R. (2004). How to evaluate teaching. *Chemical Engineering Education*, 38(3), 200–202. <http://doi.org/10.1177/002248717102200217>
- Grossman, P. L. (1991). Overcoming the apprenticeship of observation in teacher education coursework. *Teaching and Teacher Education*. [http://doi.org/10.1016/0742-051X\(91\)90004-9](http://doi.org/10.1016/0742-051X(91)90004-9)
- Hoyt, D. P., & Pallett, W. H. (1999). Appraising Teaching Effectiveness - Idea Paper # 36, (36), 1–14.
- Iqbal, I. (2013). Academics’ resistance to summative peer review of teaching: questionable rewards and the importance of student evaluations. *Teaching in Higher Education*, 18(5), 1–13. <http://doi.org/10.1080/13562517.2013.764863>

- Kealey, E. (2010). Assessment and Evaluation in Social Work Education: Formative and Summative Approaches. *Journal of Teaching in Social Work, 30*(April), 64–74. <http://doi.org/10.1080/08841230903479557>
- Kite, M. E. (2012). *Effective Evaluation of Teaching: A Guide for Faculty and Administrators*.
- Mohanty, G., Gretes, J., Flowers, C., Algozzine, B., & Spooner, F. (2005). Multi-method evaluation of instruction in engineering classes. *Journal of Personnel Evaluation in Education, 18*(2), 139–151. <http://doi.org/10.1007/s11092-006-9006-3>
- Morgan, P. (2008). The Course Improvement Flowchart : A description of a tool and process for the evaluation of university teaching. *Journal of University Teaching & Learning Practice, 5*(2), 1–14. Retrieved from <http://ro.uow.edu.au/jutlp/vol5/iss2/2>
- National Research Council. (2003). *Evaluating and improving undergraduate teaching in science, technology, engineering, and mathematics*. National Research Council. <http://doi.org/10.1002/bmb.2003.494031059998>
- Nygaard, C., & Belluigi, D. Z. (2011). A proposed methodology for contextualised evaluation in higher education. *Assessment & Evaluation in Higher Education, 36*(6), 657–671. <http://doi.org/10.1080/02602931003650037>
- Ory, J. C. (2000). Teaching Evaluation: Past, Present, and Future, (83), 13–18.
- Saroyan, A., & Amundsen, C. (2001). Evaluating University Teaching: Time to Take Stock. *Assessment & Evaluation in Higher Education, 26*(4), 341–353.
- Smith, H. (2012). Points of Departure: The unintended Consequences of Grading Teaching. *Teaching in Higher Education, 17*(6), 747–754.
- Wachtel, H. K. (1998). Student evaluation of college teaching effectiveness: A brief review. *Assessment & Evaluation in Higher Education, 23*(2), 191–212. Retrieved from <http://www.tandfonline.com/doi/abs/10.1080/0260293980230207>
- Yon, M., Burnap, C., & Kohut, G. (2016). Evidence of Effective Teaching-Perceptions of Peer Reviewers, *50*(3), 104–110.

Chapter 4 – Conclusion

This research investigated current approaches used to conduct teaching evaluation in engineering programs. It also sought to examine what teaching evaluation approaches were perceived by participants' to be effective. This study was guided by an exploratory sequential mixed-method design. Data were collected using semi-structured interviews and an online survey. In the subsequent sections, the two previous chapters are summarized independently followed by the overall conclusions of this thesis.

Chapter 2 - Teaching evaluation practices in engineering programs: Current approaches and usefulness.

Findings indicated that six approaches are used to evaluate teaching in engineering programs. These approaches are student end-of-course evaluation, classroom observation by peers or non-peers, evaluation of classroom materials, student mid-course evaluation, exit evaluation, and alumni evaluation. These approaches are used either for formative or summative evaluations. Formative evaluations are informal assessments of an educator's work used primarily to collect feedback from students or peers to help the educator improve his or her teaching. Summative evaluations are used for formal personnel decisions such as promotion and tenure, annual review or continued employment. Student end-of-course evaluation and classroom observation are the two most common approaches used for evaluating teaching in engineering programs; these approaches are both used formatively and summatively. Based on the participants' perspectives, these approaches are more beneficial when their purpose is to collect formative feedback from students or colleagues to improve faculty's teaching. Classroom materials are reviewed and evaluated such as samples of student work. Evaluation of classroom materials is also used for formative and summative evaluations. About three-quarters of the participants administered student mid-course evaluations to assess their own teaching and measure student learning. This approach is used voluntarily by educators to gather evidence for formative purposes only. The main purpose of conducting exit and alumni evaluations was to evaluate engineering programs; this evidence helps programs identify strengths and areas for improvement. However, educators might use the result from these evaluations as formative feedback

when the students comment on a specific course or certain aspect of teaching within the program. Evaluation of teaching approaches that are found to be effective and useful based on participant's perspectives were also explored. Teaching evaluation approaches such as SET and classroom observations were found to be more useful by educators when they are used for making formative evaluations than summative evaluations of teaching. Lastly, this chapter suggests descriptions of how currently used teaching evaluation practices might be aligned with evidence-based instructional practices. For instance, classroom materials such as student work results can be collected to provide evidence of students' learning.

Chapter 3 - Investigating current approaches to assessing teaching evaluation in engineering departments.

In this manuscript, the findings of this research based on the preliminary analysis of the data are described. This chapter investigates the teaching evaluation approaches at three institution types, using the data collected only from the interviews. The results showed that current approaches used to evaluate teaching in engineering programs can be used for formatively or summatively, a result that also emerged from the analysis of the interview and survey data. Student end-of-course evaluation is still widely used across three institutions while classroom observation is the second most utilized approach. The range and level of usage of teaching evaluation practices by institutions were examined to identify similarities and differences.

4.1 Challenges and Recommendations

There are many challenges on the current teaching evaluation practices used in engineering programs (see Table 4.1). The following recommendations, Table 4.1, are provided to improve the current state of teaching evaluation practices utilized in this discipline to help enhance both the educators teaching and student learning.

Table 4.1: Teaching evaluation practices challenges and recommendations

Challenges	Recommendations
Questions asked on the student end-of-course evaluation survey	Using variety of forms
Evaluators are not trained	Receive and undergo continuous training, and bringing an expert evaluator
Using standardized forms	Developing and using standardized-based forms
Single classroom observation	Regular observation
one evaluator for classroom observation and document analysis	Multiple evaluator

The questions asked on the student end-of-course evaluation survey is one of the challenges identified in this study. The questions on the survey were worded so poorly and generally that they either provide little guidance for improvement because they are often irrelevant to a particular course. Therefore, it would be beneficial to develop and have a variety of forms of survey that educators can choose from when evaluating a certain class that would be most relevant to the particular class that they are teaching. The majority of the participants indicated that the classroom observer often didn't receive any training or if so, the training is very minimal. It is important that observers receive and undergo continuous training to gain knowledge on how to conduct proper observations and analysis of the documents collected. It also essential that evaluators of classroom materials and observers used standardized forms. The standard forms help guide and remind the evaluator and observer on what to evaluate and how to evaluate classroom documents. It is also essential that educators are evaluated regularly and by multiple observers per course to assess all aspects of teaching and learning.

The literature contains many examples of successful standards and practices for effective teaching evaluation that have been proven to improve teaching and student learning. The findings from this research highlights how important it is that educators are evaluated on teaching practices that align with evidence-based instructional practices to determine their effectiveness. Teaching evaluation should recognize characteristics of effective instruction and include these in the evaluation. However, it was identified that many faculties from engineering programs have received minimal or no formal training in techniques or

strategies to evaluate their own or their colleagues teaching effectiveness, which was also found across STEM disciplines by the National Research Council (2003). They may not also be familiar with the literature of teaching evaluation practices that could lead to direct improvements. Based on these findings, teaching evaluation in engineering programs do not align to those evidence-based instructional practices. For that reason, it means that an educator's teaching is neither improving nor are students learning.

4.2 Future Research

In chapter 3, the teaching evaluation approaches used in engineering were investigated and categorized by types of institutions. It was determined that it is necessary to explore the same sample size from each type of institutions. Therefore, further investigation could be conducted to make appropriate comparisons of practices within and across institutions.

During interview, we didn't investigate what evaluators consider effective or not when they evaluate other educator's teaching. Further inquiry on this topic might be an important area of research to identify and understand what aspects of teaching are evaluated that help improve teaching and student learning.

Bibliography

- Baxter, P., & Jack, S. (2008). Qualitative Case Study Methodology : Study Design and Implementation for Novice Researchers Qualitative Case Study Methodology : Study Design and Implementation, *13*(4), 544–559.
- Berk, R. A. (2005). Survey of 12 Strategies to Measure Teaching Effectiveness. *International Journal of Teaching and Learning in Higher Education*, *17*(1), 48–62.
- Biggs J. (1996). Enhancing Teaching Through Constructive Alignment. *Higher Education*, *32*, 347–364.
- Brandenburg, D. C., Braskamp, L. A., & Ory, J. C. (1984). *Evaluating Teaching Effectiveness: A Practical Guide*.
- Broder, J. M., & Taylor, W. J. (2015). Teaching Evaluation in Agricultural Economics and Related Departments. *American Journal of Agricultural Economics*, *76*(1), 153–162
- Brown, S. A., Villanueva, K. A., Pitterson, N. P., & Sitomer, A. (2016). Investigating current approaches to assessing teaching evaluation in engineering departments. In *IEEE Frontiers in Education Conference* (pp. 1–6).
- Brown, S. A., Villanueva, K. A., Pitterson, N. P., Hurwitz, D. S., & Sitomer, A. (2016). *Teaching evaluation practices in engineering programs: Current approaches and usefulness*. Manuscript under review.
- Campbell, J., & Bozeman, W. (2007). The value of student ratings: Perceptions of students, teachers, and administrators. *Community College Journal of Research and Practice*, *32*(1), 13–24. <http://doi.org/10.1080/10668920600864137>
- Chen, Y., & Hoshower, L. B. (2003). Student Evaluation of Teaching Effectiveness: An Assessment of Student Perception and Motivation. *Assessment & Evaluation in Higher Education*, *28*(1), 71–88.
- Chi, M. T. H., & Wylie, R. (2014). The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational Psychologist*, *49*(4), 219–243. <http://doi.org/10.1080/00461520.2014.965823>
- Chickering, A. W., & Gamson, Z. F. (1987). Seven principles for good practice in Undergraduate Education. *American Association of Higher Education Bulletin*, *39*, 3–7. [http://doi.org/10.1016/0307-4412\(89\)90094-0](http://doi.org/10.1016/0307-4412(89)90094-0)
- Creswell, J. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. SAGE (4th ed.). <http://doi.org/10.1007/s13398-014-0173-7.2>
- Crotty, M. (1998). *The Foundations of Social Research: Meaning and Perspective in the Research Process*.

- Dedoose Version 7.1.3. (2015). Web Application for Managing, Analyzing, and Presenting Qualitative and Mixed-Method Research Data.
- Dressel, P. (1960). Teaching , Learning , and Evaluation. *Improving College and University Teaching*, 8(1), 11–15.
- Eddy, S. L., Converse, M., & Wenderoth, M. P. (2015). PORTAAL: A Classroom Observation Tool Assessing Evidence-Based Teaching Practices for Active Learning in Large Science, Technology, Engineering, and Mathematics Classes. *CBE - Life Science Education*, 14, 1–16. <http://doi.org/10.1187/cbe-14-06-0095>
- Felder, R. M., & Brent, R. (2004). How to evaluate teaching. *Chemical Engineering Education*, 38(3), 200–202. <http://doi.org/10.1177/002248717102200217>
- Ferguson, R. (2007). Constructivism and Social Constructivism. In *Theoretical Frameworks for Research in Chemistry/Science Education* (pp. 28–49).
- Fink, L. D., Ambrose, S., & Wheeler, D. (2005). Becoming a professional engineering educator: A new role for a new era. *Journal of Engineering Education*, 94(January), 185–194. <http://doi.org/10.1002/j.2168-9830.2005.tb00837.x>
- Goe, L., Biggers, K., Croft, A. (2012). *Linking Teacher Evaluation to Professional Development: Focusing on Improving Teaching and Learning*.
- Goldschmid, M. L. (1978). The evaluation and improvement of teaching in higher education. *Department of Psychology and Higher Education, Swiss Federal Institute of Technology, Lausanne, Switzerland*, 7, 221–245.
- Grossman, P. L. (1991). Overcoming the apprenticeship of observation in teacher education coursework. *Teaching and Teacher Education*. [http://doi.org/10.1016/0742-051X\(91\)90004-9](http://doi.org/10.1016/0742-051X(91)90004-9)
- Henderson, C., Turpen, C., Dancy, M., & Chapman, T. (2014). Assessment of teaching effectiveness: Lack of alignment between instructors, institutions, and research recommendations. *Physical Review Special Topics - Physics Education Research*, 10(1), 1–20. <http://doi.org/10.1103/PhysRevSTPER.10.010106>
- Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education*, 39(4), 372–400.
- Hoyt, D. P., & Pallett, W. H. (1999). Appraising Teaching Effectiveness - Idea Paper # 36, (36), 1–14.
- Iqbal, I. (2013). Academics' resistance to summative peer review of teaching: questionable rewards and the importance of student evaluations. *Teaching in Higher Education*, 18(5), 1–13. <http://doi.org/10.1080/13562517.2013.764863>

- Johnson, T. D., & Ryan, K. E. (2000). A comprehensive approach to the evaluation of college teaching. *New Directions for Teaching and Learning*, (83), 109–123. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=tfh&AN=9178975&lang=es&site=ehost-live>
- Juwah, C., Macfarlane-Dick, D., Matthew, B., Nicol, D., Ross, D., & Smith, B. (2004). Enhancing student learning through effective formative feedback. *Higher Education*, (68), 1–41. Retrieved from http://www.heacademy.ac.uk/resources/detail/resource_database/id353_effective_formative_feedback_juwah_et al
- Kealey, E. (2010). Assessment and Evaluation in Social Work Education: Formative and Summative Approaches. *Journal of Teaching in Social Work*, 30(April), 64–74. <http://doi.org/10.1080/08841230903479557>
- Kite, M. E. (2012). *Effective Evaluation of Teaching: A Guide for Faculty and Administrators*.
- Knapper, C. (2001). Broadening Our Approach to Teaching Evaluation. *New Directions for Teaching and Learning*, 2001(88), 3–9. <http://doi.org/10.1002/tl.32>
- McKeachie, W. . (2002). *McKeachie's Teaching Tips: Strategies, Research, and Theory for College and University Teachers*.
- Mohanty, G., Gretes, J., Flowers, C., Algozzine, B., & Spooner, F. (2005). Multi-method evaluation of instruction in engineering classes. *Journal of Personnel Evaluation in Education*, 18(2), 139–151. <http://doi.org/10.1007/s11092-006-9006-3>
- Morgan, P. (2008). The Course Improvement Flowchart : A description of a tool and process for the evaluation of university teaching. *Journal of University Teaching & Learning Practice*, 5(2), 1–14. Retrieved from <http://ro.uow.edu.au/jutlp/vol5/iss2/2>
- National Research Council. (2003). *Evaluating and improving undergraduate teaching in science, technology, engineering, and mathematics*. National Research Council. <http://doi.org/10.1002/bmb.2003.494031059998>
- Nygaard, C., & Belluigi, D. Z. (2011). A proposed methodology for contextualised evaluation in higher education. *Assessment & Evaluation in Higher Education*, 36(6), 657–671. <http://doi.org/10.1080/02602931003650037>
- Ohanlon, J., & Mortensen, L. (1980). Making Teacher Evaluation Work. *Journal of Higher Education*, 51(6), 664–672. <http://doi.org/10.2307/2943969>
- Ory, J. C. (2000). Teaching Evaluation: Past, Present, and Future, (83), 13–18.
- Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education- Washington*, 93(July), 223–232. <http://doi.org/10.1002/j.2168-9830.2004.tb00809.x>

- Reeves, T. C. (2006). How do you know they are learning?: the importance of alignment in higher education. *Int. J. Learning Technology J. Learning Technology*, 2(4), 294–309. <http://doi.org/10.1504/IJLT.2006.011336>
- Saroyan, A., & Amundsen, C. (2001). Evaluating University Teaching: Time to Take Stock. *Assessment & Evaluation in Higher Education*, 26(4), 341–353.
- Shute, V. aleri. J. (2008). Focus on Formative Feedback. *Review of Educational Research*, 78(1), 153–189. <http://doi.org/10.3102/0034654307313795>
- Smith, H. (2012). Points of Departure: The unintended Consequences of Grading Teaching. *Teaching in Higher Education*, 17(6), 747–754.
- Wachtel, H. K. (1998). Student evaluation of college teaching effectiveness: A brief review. *Assessment & Evaluation in Higher Education*, 23(2), 191–212. Retrieved from <http://www.tandfonline.com/doi/abs/10.1080/0260293980230207>
- Wankat, & Oreovicz. (2009). Teaching engineering. *IEEE Engineering Management Review*, 37(1), 1–9. <http://doi.org/10.1109/EMR.2008.2012072>
- Wiggins, G., & McTighe, J. (1998). What is backward design? In *Understanding by Design* (pp. 7–19). <http://doi.org/10.1016/j.cie.2006.03.005>
- Yon, M., Burnap, C., & Kohut, G. (2016). Evidence of Effective Teaching-Perceptions of Peer Reviewers, 50(3), 104–110.

Appendix A - Interview Protocol

Teaching Evaluation and Assessment Study Protocol

Name of Interviewee:

Name of Interviewer:

Date/Time:

Department & School Name:

Time Duration of the Interview: approximately 30-40 minutes

Introduction

The purpose of this research study is to identify current practices used to evaluate teaching in engineering departments across the country to understand and assess teaching with qualitative interviews.

The following steps will be conducted:

- Gather documentation of current practice in the College of Engineering to develop a summary document that describes existing process.
- Identify best practices in the COE departments and schools throughout the country by interviewing faculty members, engineering department heads or any individual in the department who is knowledgeable of the teaching evaluation and assessment practices.

- There will only be one interview (via phone or video conference or in person) that will take approximately 30-40 minutes.
- The answers provided by the interviewee will be anonymous and the interviewee will not be identified.
- Documents gathered will be confidential
- With their permission, it would be beneficial if the interview was audio recorded for depth analysis, and so you won't miss or forget anything during the conversation
- Any article or documentation resulted from this work will be provided to the interviewee department/school

Note: Dr. Shane Brown will share the documents via email.

- The interviewee can skip any question they wish, and can stop the interview anytime.

Does the instructor has any questions or concerns before starting the interview?

Interviewee Background

1. What is the name of your university?
2. What is the name of your department?
3. Could you please tell me a little bit about yourself and your role in your department?

Teaching Evaluation Questions

1. What procedures have you used or are currently using to evaluate teaching in your department?
2. Could you tell me the different methods to evaluate teaching that is being utilized in your department?

Method Name	Formal	Informal	Not Used	Survey or Interview
A. Student Evaluation of Teaching (SET)				
B. Peer Evaluation				
C. Student Midterm Evaluation				
D. Senior Exit Evaluation				
E. Instructor Self-Evaluation				
F. Other				
G. Other				
H. Other				

3. Is the teaching evaluation process/procedures the same for the every department?

(For example: Civil Engineering Department and Mechanical Engineering Department have different process of teaching evaluation.

Peer Review Questions

(Peer Review - is the evaluation of performance by other people in the same field or discipline.)

1. What is the main purpose of peer review?
2. On average, how many peer reviews does an instructor receive per year?
(Assistant Professor, Associate Prof. and Full Prof.)
 - Is there a different set of requirements for each professor title (Assistant Professor, Associate Professor, and Full Professor)?
3. Is the peer review used for promotion purposes?
4. If used for promotion purposes, how many peer reviews does a faculty member in the department have on average before applying for promotion?
5. What does a single peer review consist of?
 - a. Observation of single or more lectures?
 - b. Review of syllabus, course notes, assignments, or exams?
6. Who administers the instructor's peer review?
(For example: Other faculty member from the same department or other department, member of a certain committee or hired specialist?)
7. Does the evaluator/observer get any form of training?
(For example: seminar such as Fall Retreat)
8. Do the observers/evaluators have a checklist or standard form that they use when they conduct the peer review?
9. If they have a checklist or standard form, could you describe what are the observers exactly looking for during evaluation?
 - Is there a rubric used to guide evaluate the instructor?
 - Do the evaluators just write comments on that standard form regarding on what they have observed such as strengths or weakness of the instructors teaching?
10. What feedback is documented in a faculty member's permanent file from a peer review?
 - a. Signed letter?
 - b. Other
11. What documentation is provided to the faculty member?

Student Evaluation of Teaching (SET)

1. What is the purpose of the student evaluation of teaching (SET)?
2. Is the survey done online or paper-based?
 - If they switch from paper-based to online-based or vice versa:
Did the department response rate increase or decrease?
3. Who determines what questions are on the student evaluation of teaching?
4. Are they standard questions? (The same in every department, maybe school)
5. What questions are asked on the SET? Examples?
6. Can students write comments?
7. Can instructors add their own questions to get more feedback?
(Can they write their own questions or do they choose from a list of questions?)
8. When do the students take the survey?
(This is usually done 2 or 3 weeks before the end of the term)
9. Can students still take the survey after grades have been posted?
10. Are the evaluations published for students and for public to see? Yes or No?
Explain.

Student Midterm, Senior Exit and Instructor Self-Evaluation

Main Question:

Other than the formal methods, are there any informal methods to evaluate teaching in the department such as student midterm evaluation, senior exit evaluation and instructors' self-evaluation?

Follow Up Questions:

1. Midterm Student Evaluation

(The instructor, sometimes by the Center for Teaching and Learning, in the middle of the term/semester usually gives this survey. Instructors can use the feedback from this method to make adjustments on the course or to get students' opinion.)

- a. Could you describe the procedure/process for the midterm evaluation?
- b. Is this method done on a faculty-to-faculty basis? (Not Mandatory, only if instructors want to?)
- c. What is the purpose/goal of the midterm survey?

2. Senior Exit Evaluation

(This method is usually given to graduating seniors. The feedback or results from this method is used to improve the program or engineering courses to benefit future students.)

- a. What is the purpose/goal of the senior exit evaluation?
- b. Could you describe the procedure/process for this method?
- c. Does the department send a survey to all graduating seniors or the department conducts an interview with the graduating seniors?

For Survey: What questions are asked on the survey?

For Interview: Describe the procedure/process of the senior interview?
(One on one interview? Open forum?)

3. Instructor Self-Evaluation

(Evaluation of oneself or performance at a job or learning task considered in relation to an objective standard.)

- a. What is the purpose of this method?
- b. Helpful or not?

Follow Up Questions / Opinion

- ❖ Could you provide us any materials that are used to evaluate and assess teaching in the department/school?
(For example: survey copy, checklist used for evaluating a faculty member, etc.)
- ❖ What do you think about the teaching evaluation and assessment in your department/school? (Effective or ineffective?). Why?
- ❖ Are evaluations/assessments mandatory?
- ❖ If the teaching evaluation and assessment is not effective, did anyone question the system to make changes and improve the system? What are the limitations for implementing a better evaluation system for teaching?
- ❖ Center for Teaching and Learning (CTL)
 - Is there a center for teaching and learning
 - What is the CTL role?
- ❖ What changes would you like to see regarding the teaching assessment and evaluation in your department?
- ❖ Does the department gives any incentives or any form of rewards to instructors who gets good review/ratings or is there any consequences when an instructor gets a bad review from their evaluations?
- ❖ Is there any other method to assess and evaluate teaching that you prefer other than the one we have talked about?
- ❖ Is the process/procedure to evaluate or assess teaching the same for newer faculty member and tenured faculty member?

Closing Statement

- Thank you for your time and participating in an interview.

(Please contact Keisha or Dr. Shane Brown if they have questions/concerns or just want an update regarding the study via email.)

Appendix B – Survey Protocol

EXPLANATION OF RESEARCH

Project Title:	Investigating the Current Approaches to Assessing Teaching Evaluation Practices in Engineering Departments
Principal Investigator:	Shane Brown
Student Researcher:	Keisha Villanueva
Co-Investigator(s):	David Hurwitz, Ann Sitomer, Nicole Pitterson

Research Purpose: You are being asked to take part in a research study. The purpose of this research study is to identify current practices in the evaluation of teaching in engineering departments across the country.

Survey Purpose: You are being asked to participate in a survey to gather additional data to supplement the data gathered from the semi-structured interview you participated in earlier. The results from this survey will be used to clarify some of the teaching evaluation practices that emerged during the interviews.

Activities: During this study, you will be asked to participate in a survey

Risks: There is a chance that we could accidentally disclose information that identifies you.

Time: Your participation in this survey will last for approximately 5-15 minutes.

Confidentiality: The information you provide in the survey will be kept confidential to the extent allowed by law. When we write about the study you will not be identified.

Voluntary: Participation in this study is voluntary. You may skip or not answer any question at any time.

Study contacts: If you have any questions about this research project, please contact: Shane Brown, Associate Professor, Civil and Construction Engineering. If you have questions about your rights or welfare as a participant, please contact the Oregon State University Institutional Review Board (IRB) Office, at (541) 737-8008 or by email at IRB@oregonstate.edu

Investigating the Current State of Teaching Evaluation Practices in Engineering Departments Survey Protocol

Introduction

The purpose of this research is to identify current practices in the evaluation of teaching in engineering departments across the country. The purpose of this survey is to gather additional data to supplement the data gathered from the semi-structured interview you participated in earlier. The results from this survey will be used to clarify some of the teaching evaluation practices that emerged during the interviews.

Context

1. What is the name of your college or university?
2. What is the name of your department or school?
3. On average, how many undergraduate students complete your engineering program annually?
4. On average, how many graduate students complete your engineering program annually (if applicable)?
5. What degrees are offered? (Check all that apply)
 - i. AA/AS
 - ii. BS
 - iii. MS
 - iv. MEng
 - v. PhD
 - vi. Other (explain)

In this survey, we will ask you about the teaching evaluation practices that you have utilized or currently using in your department. This survey has three parts: teaching evaluation practices gathered from peers and/or administrators, from students, and the last part ask for other practices not mentioned during the interview. The questions will be in a multiple choices, and some text format.

Teaching Evaluation Practices

Peers or Administrators

The following questions will ask about the teaching evaluation practices that are gathered from either your peers or administrators. These practices involve classroom observation by either your peers or administrators, and review of materials such as syllabus, exams, and assignments

1. Classroom Observation

The collaborative process by which the observer sits in on one or more classroom sessions, records the instructor's teaching practices and student actions, and then meets with the faculty member to discuss the observations.

- a) Does your school or department use classroom observation?
 - i. Yes
 - ii. No
 - iii. Other, Explain
 - iv. I don't know
- b) How is the information on the classroom observation used? (Check all that apply)
 - i. Promotion
 - ii. Annual Review
 - iii. Teaching Improvement
 - iv. Other, Explain
 - v. I don't know
- c) What does a classroom observation consist of?
 - i. Observation of a single lecture
 - ii. Observation of more than one lecture
 - iii. Other, Explain
 - iv. I don't know
- d) Who administers the classroom observation in your school or department? (Check all that apply).
 - i. Another faculty member from the same department
 - ii. Another faculty member from different department
 - iii. Department Chair
 - iv. Dean
 - v. Center for Teaching and Learning (CTL) personnel
 - vi. Other, Explain
 - vii. I don't know
- e) Are all faculty members required to get a classroom observation in each course they teach regardless of rank or position? (e.g., employment status, or rank such as full, assistant or associate professor, adjunct instructor)
 - i. Yes
 - ii. No (if no, please explain the differences)
 - iii. Other, Explain

- iv. I don't know
- f) On average, how many classroom observation does a faculty member receive per year?
 - i. Less than 1
 - ii. 1
 - iii. 2
 - iv. 3 or more
 - v. Depends on the faculty member's employment status or rank
 - vi. I don't know
- g) Is the classroom observation feedback documented in a faculty member's personnel file?
 - a. Yes
 - b. No
 - c. Other, Explain
 - d. I don't know
- h) Does the observer use a standard form or checklist for observation?
 - a. Yes
 - b. No
 - c. Other, Explain
 - d. I don't know
- i) Does the observer get any form of training?
 - a. Yes
 - b. No
 - c. Other, Explain
 - d. I don't know
- j) What type of training does the observer gets? (check all that apply)
 - a. Workshop
 - b. Ongoing Seminar
 - c. Other, Explain
 - d. I don't know

2. Evaluation of Classroom Materials

This process involves collecting and reviewing student work (e.g., exams, quizzes, projects), course materials (e.g., syllabus, assignments, rubrics).

- a) Does your school or department use classroom materials evaluation?
 - i. Yes
 - ii. No
 - iii. Other, Explain
 - iv. I don't know
- b) How is the information gathered from reviewing classroom materials used? (Check all that apply)
 - i. Promotion
 - ii. Annual Review
 - iii. Teaching Improvement
 - iv. Other, Explain

- v. I don't know
- c) What type of materials are collected and reviewed? (Check all that apply)
 - i. Exams
 - ii. Quizzes
 - iii. Projects
 - iv. Syllabus
 - v. Assignments
 - vi. Other, Explain
 - vii. I don't know
- d) Who evaluates the materials in your school or department? (Check all that apply).
 - i. Another faculty member from the same department
 - ii. Another faculty member from different department
 - iii. Department Chair
 - iv. Dean
 - v. Center for Teaching and Learning (CTL) personnel
 - vi. Other, Explain
 - vii. I don't know
- e) Are all faculty members required to submit their classroom materials in each course they teach regardless of rank or position?
(e.g., employment status, or rank such as full, assistant or associate professor, adjunct instructor)
 - v. Yes
 - vi. No (if no, please explain the differences)
 - vii. Other, Explain
 - viii. I don't know
- f) On average, how often do they evaluate the faculty member's classroom materials per year?
 - vii. Less than 1
 - viii. 1
 - ix. 2
 - x. 3 or more
 - xi. Depends on the faculty member's employment status or rank
 - xii. I don't know
- g) Is the classroom materials documented in a faculty member's personnel file?
 - a. Yes
 - b. No
 - c. Other, Explain
 - d. I don't know
- h) Does the evaluator use a standard form or checklist for evaluating classroom materials?
 - a. Yes
 - b. No
 - c. Other, Explain
 - d. I don't know

- i) Does the evaluator get any form of training?
 - a. Yes
 - b. No
 - c. Other, Explain
 - d. I don't know
- j) What type of training does the evaluator gets? (check all that apply)
 - a. Workshop
 - b. Ongoing Seminar
 - c. Other, Explain
 - d. I don't know

Students

The following questions will ask about the teaching evaluation practices that are gathered from your students.

3. Student End-of-Course Evaluation

By student end-of-course evaluation, we mean a survey given at the end of the term or semester.

- a) Does your school or department use student end-of-course evaluation?
 - i. Yes
 - ii. No
 - iii. Other, Explain
 - iv. I don't know
- b) How is the information on the student end-of-course evaluation used? (Check all that apply)
 - i. Promotion
 - ii. Annual Review
 - iii. Teaching Improvement
 - iv. Other, Explain
 - v. I don't know
- c) Are all faculty members required to do student end-of-course evaluation in each course they teach regardless of rank or position? (e.g., employment status, or rank such as full, assistant or associate professor)
 - i. Yes
 - ii. No (if no, please explain the differences)
 - iii. Other, Explain
 - iv. I don't know
- d) How is the student end-of-course evaluation administered?
 - v. Online
 - vi. Paper-based
 - vii. Other, Explain
 - viii. I don't know

- e) Are the questions on the student end-of-course evaluation standard?
If yes, indicate if the method is standard by across the institution, within College of Engineering (COE) or within the department only (e.g., Civil, Mechanical, Electrical etc.)
 - a. Yes, (specify if it's either by Institution, COE or Department only)
 - b. No
 - c. I don't know
- f) Can instructors add their own questions in the student end-of-course evaluation?
 - a. Yes
 - b. Depends on the faculty member's employment status or rank
 - c. No
 - d. I don't know
- g) Are instructors provided a set of questions to choose from to add on the student end-of-course evaluation?
 - a. Yes
 - b. No
 - c. I don't know

4. Student Mid-Course Evaluation

The mid-course student evaluation is used to assess students' perceptions of teaching strategies and/or how students are learning while the class is in progress.

- a) Does your school or department use student mid-course evaluation?
 - i. Yes
 - ii. No
 - iii. Other, Explain
 - iv. I don't know
- b) How is the information on the student mid-course evaluation used? (Check all that apply)
 - i. Promotion
 - ii. Annual Review
 - iii. Formative Feedback
 - iv. Other, Explain
 - v. I don't know
- c) Are the questions on the student mid-course evaluation standard? If yes, indicate if the method is standard across the Institution, within College of Engineering (COE) or within department only (e.g., Civil, Mechanical, Electrical etc.)
 - i. Yes, (specify if it's by Institution, COE or Department only)
 - ii. No
 - iii. I don't know

d) Briefly describe the practice for the student mid-course evaluation.

5. Exit Evaluation

By exit evaluation, we mean a survey given to students who will be completing the program. This practice is also done by conducting an individual or group interview with the students. This allows the students to provide feedback about their experiences in the engineering program.

- a) Does your school or department use exit evaluation?
 - i. Yes
 - ii. No
 - iii. Other, Explain
 - iv. I don't know
- b) How is the information on the exit evaluation used? (check all that apply)
 - i. Promotion
 - ii. Formative Feedback
 - iii. Other, Explain
 - iv. I don't know
- c) Briefly describe the practice for the exit evaluation.

6. Alumni Evaluation

This evaluation is given to alumni of the program about their experiences in the engineering program.

- a) Does your school or department use alumni evaluation?
 - i. Yes
 - ii. No
 - iii. Other, Explain
 - iv. I don't know
- b) How is the information on the alumni evaluation used? (Check all that apply)
 - i. Promotion
 - ii. Formative Feedback
 - iii. Other, Explain
 - iv. I don't know
- c) Briefly describe the practice for the alumni evaluation.

Other Practices

The next questions will ask about the Center for Teaching and Learning (CTL) and if there are any teaching evaluation practices that we did not mention.

7. Center for Teaching and Learning (CTL) or equivalent

They promote teaching excellence and provide supports and other services to assist faculty members with their teaching.

- a) Does your college or university have a Center for Teaching and Learning, or equivalent, that provide support and services for faculty members related to teaching evaluation and/or teaching improvement?
 - i. Yes
 - ii. No
 - iii. Other, Explain
 - iv. I don't know

- b) What services does your Center for Teaching and Learning provides? (check all that apply)
 - i. Classroom observation for teaching evaluation
 - ii. Review of syllabus
 - iii. Review of exams
 - iv. Review of assignments
 - v. Workshops
 - vi. Ongoing seminars or professional learning communities
 - vii. Other, Explain
 - viii. I don't know

- c) Does your school or department use any other teaching evaluation practices that are not mention?
 - i. Yes
 - ii. No

- d) List all the other teaching evaluation name and a brief description of the practices for each.

End of the Survey