12-2020

University Course Evaluations: A Study of the Influence of Faculty, Student, and Course Variables

Tyesha De'Shuan Stewart
Stephen F. Austin State University, stewarttd1@jacks.sfasu.edu

University Course Evaluations: A Study of the Influence of Faculty, Student, and Course Variables

Creative Commons License

UNIVERSITY COURSE EVALUATIONS: A STUDY OF THE INFLUENCE OF FACULTY, STUDENT, AND COURSE VARIABLES

By

Tyesha De'Shuan Stewart, Master of Art

Presented to the Faculty of the Graduate school of

Stephen F. Austin State University

In Partial Fulfillment

Of the Requirements

For the Degree of

Doctor of Philosophy

STEPHEN F. AUSTIN STATE UNIVERSITY

December, 2020

UNIVERSITY COURSE EVALUATIONS: A STUDY OF THE INFLUENCE OF FACULTY, STUDENT, AND COURSE VARIABLES

By

Tyesha De'Shuan Stewart, Master of Art

APPROVED:

_____
Luis Aguerrevere, Ph. D., Dissertation Director

_____
Nina Ellis-Hervey, Ph. D., Committee Member

_____
Elaine Turner, Ph. D., Committee Member

_____
Sarah Savoy, Ph. D., Committee Member

_____
Pauline M. Sampson, Ph. D.
Dean of Research and Graduate Studies

Abstract

Course evaluations impact faculty' annual evaluations and have become somewhat controversial; yet course evaluations in faculty evaluations persists. The purpose of this study is to provide a more in-depth examination of course evaluations by analyzing faculty, student, and course variables. Analyses were performed to address the following research question: "To what degree do faculty gender, faculty race, faculty year of terminal degree, student gender representation in class, student race representation in class, total enrollment, course level, and grading patterns predict the variance in overall course evaluation, standard deviation, and response rate in one university academic unit?" Collected data consisted of course evaluations of Human Services-related courses and information from faculty and students, during one academic year. Results from this study may assist administrators in enhancing the promotion, tenure, and merit review process when considering the impact of faculty and student variables on course evaluations cumulative ratings, standard deviation, and course evaluation response rates.

Acknowledgements

First and foremost, I offer my most sincere gratitude to my dissertation chair(s) Dr. Robbie Steward and Dr. Luis Aguerrevere, for their continuous support throughout the research and writing process of this dissertation. Their expertise and insightful feedback were invaluable and pushed me to think more critically about my study. I would also like to thank my dissertation committee members, Dr. Nina Ellis-Hervey, Dr. Sarah Savoy, and Dr. Elaine Turner for providing their support and insightful commentary throughout my dissertation journey. I'm beyond grateful and blessed for their involvement in this accomplishment.

Last but not least, I wish to praise my family and friends for all of their constant support throughout my graduate experience. I am grateful for my mother, Rosie Holloway, for always providing me with her wisdom, comfort, and love when I struggle on my path. I want to thank my sister, Duana Clark, for keeping me motivated and reminding me to keep exceeding my goals in everything I set my heart out to do. I also want to thank my aunt, Bettye Williams, for continuously uplifting me in prayer and reminding me just how amazing I truly am. I am especially grateful for my angelic grandmother, Bettye Holloway. Thank you for providing me with your warm embrace and loving spirit, you will forever be in my heart.

I love and appreciate each of you for your support. Thankful that God blessed me with a phenomenal support system to ensure that I am physically, mentally, emotionally, and spiritually stable as I continue on my journey to success.

Table of Contents

List of Tables

Chapter I

Introduction

The contentious debate of using university course evaluations as a predominant means to measure teaching effectiveness is a long-standing topic in higher education (Aleamoni & Yimer, 1973; Boring, Ottoboni, & Stark, 2016; Chen, Y., & Hoshower, 2003; Emery, Kramer, & Tian, 2003; Frey, 1978; Kelley, 1972; Kidd & Latif, 2004; Marsh & Bailey, 1993; Marsh & Roche, 1997; McClain, Gulbis, & Hays, 2018; Shevlin, Banyard, Davies, & Griffiths, 2000). Many advocates of the use of this measure view this success-based tool completed by students as a reliable measurement in assessing teaching effectiveness, review for promotion and tenure, and decisions in merit raises (Rebman Jr, Wimmer, & Booker, 2018; Zafar, Ghazal, Parpio, & Amirali, 2017). Those with consistently strong, positive, teaching evaluations have been deemed effective instructors and often validated with increases in salary, teaching awards, and promotion tenure, and merit (Langbein, 2008; Subbaye & Vithal, 2017; Williams & Ceci, 1997). This is especially the case when strong course evaluations are received in addition to documented scholarship and research activities and positively evaluated professional service. However, some scholars question the validity and reliability of this measurement and are opposed to using this tool as a mean of determining faculty members' success in academia (Khong, 2016; Oon, Spencer, & Kam, 2017).

Although the use of a standardized course evaluation continues to be a common practice in the evaluation of university and college faculty teaching, questions are being raised about the meaningfulness of students' ratings of course instruction (Feistauer & Richter, 2017). Do these measures effectively assess competence as instructors or do they measure other unknown processes in the perceptions of students that should not be considered relevant in the annual review process for academicians or validated with promotion and secured employment longevity (Hornstein, 2017). Some argue that teaching evaluations primarily reflect the likability of the faculty. It is purported that students rate more positively those faculty who grade less rigorously, who are more lenient, accessible and friendly; while faculty who have high standards and focus on the content and information delivery are penalized in the process (Heckert, Latier, Ringwald-Burton, & Drazen, 2006; Weinkle, Stratford, & Lee, 2020). If this is the case, course evaluations are not a valid measure that allows faculty to reflect or make pertinent changes to instructional methods or course content to improve students' quality of learning. On the other hand, advocates of the current tradition argue that the measures are accurate reflection of both teaching efficacy and effectiveness (Carlucci, Renna, Izzo, & Schiuma, 2019). The topic has become somewhat controversial in blogs, promotion, tenure, and merit committees, and in the Chronicles of Higher Education (Falkoff, 2018; Linse, 2017; Ray, Babb, & Wooten, 2018), however, the practice of including course evaluations in faculty evaluations persists (Schmelkin, Spencer, & Gellman, 1997).

With course evaluations being a prevalent topic of discussion in higher education, the quality of this measurement is constantly studied to determine the factors associated with students' evaluations of teaching effectiveness. Variables identified as most important in predicting teaching effectiveness consist of content domain items, instructor domain items, and student domain items which include course organization, course materials, faculty interaction, teaching methods, grading methods, assignments/examinations, and requirements/expectations (Park & Dooris, 2020). These items generate the questions that are typically seen on the surveys that are administered to students at the end of every semester. Yet, further research indicates these variables that were once deemed important in this evaluation have been overshadowed by other bias factors (Royal & Stockdale, 2015).

Studies have shown factors such as gender, race, faculty rank, academic discipline, class size, course level, course delivery, and expected grade significantly impact overall course evaluations (Barnes & Barnes, 1993; Capa-Aydin, 2016; Chisadza, Nicholls, & Yitbarek, 2019; Culver, 2010; Kifle & Alauddin, 2016; Mitchell & Martin, 2018; Peterson, Biederman, Andersen, Ditonto, & Roe, 2019; Smith & Hawkins, 2011; Stewart, 2018). These extraneous variables give rise to questions of whether course evaluations are a true representation of teaching effectiveness or merely a reflection of student bias of faculty characteristics (i.e. gender, race, rank) and course variables (i.e. course type, level, delivery type, class size). Faculty are then penalized for these factors of which they have limited to no control. This is a relevant issue given that critical

decision-making outcomes are based on the use of an instrument to measure quality of teaching effectiveness by untrained evaluators; and that the results are associated with the career of academicians' likelihood of promotion, tenure, merit, salary increase, and job security (Stewart, 2018). It is imperative that administrators carefully analyze and monitor factors that influence course evaluations across colleges and departments to ensure results are interpreted accurately and equally when partaking in personnel decisions during annual review of faculty members' dossiers (Linse, 2017).

Nevertheless, literature associated with empirical studies of course evaluation is limited and that most address course characteristics (i.e. course level, class size, course grade), while some address faculty characteristics (i.e. gender, race, rank) and a few address student variables (i.e. race, gender). Conclusions of those who raise questions about the validity and reliability of the course evaluation process is based on anecdotal evidence which suggest some consistency in faculty attitudes whom are opposed to using this tool to determine teaching effectiveness due to the influence of numerous variables and ratings being obtained from untrained evaluators (Abrami, 2001; Nasser & Fresko, 2002; Schmelkin, Spencer, & Gellman, 1997; Theall & Franklin, 2001). Nonetheless, faculty perceptions of the use of course evaluation continue to be evaluated within studies as some are advocates for the intended purpose of this tool while others believe that it is a bias tool that is likely to impede their progress toward promotion, tenure, and merit.

Summary and Statement of the Problem

Faculty, student, and course variables have yet to be examined simultaneously; and unexamined errors in administration decision-making might persist. Therefore, this study will add to the literature by providing a more in-depth examination of all the potential variables associated with the variance of course evaluations at one time. The focus of this study is to examine the following research questions:

1. To what degree do faculty variables (i.e. faculty gender, faculty race, faculty year of terminal degree), student-classroom variables (i.e. student gender representation in class % of female, student race representation in class % of racial majority, total enrollment), and course variables (i.e. course level, grading patterns) predict the variance in overall course evaluations?
2. To what degree do faculty variables (faculty gender, faculty race, faculty year of terminal degree), student-classroom variables (i.e. student gender representation in class % of female, student race representation in class % of racial majority, total enrollment), and course variables (i.e. course level, grading patterns predict the variance in the standard deviation of the course evaluation?
3. To what degree do faculty variables (faculty gender, faculty race, faculty year of terminal degree), student-classroom variables (i.e. student gender representation in class % of female, student race representation in class % of racial majority, total enrollment), and course variables (i.e. course level, grading patterns) predict the variance in response rates?

Due to the challenge of obtaining individual student information per faculty/course, the percentage of gender representation in class and the percentage of student representation in class will be examined for this research design. Each variable will be explored to identify the strongest predictors in the model for overall course evaluation, standard deviation, and response rates. A comprehensive literature review of prior research that examined variables that influence course evaluation will be presented and discussed in Chapter 2.

Chapter II

Literature Review

Though limited research has been conducted to examine factors that influence student rating of teaching and the impact it has on promotion, tenure, and merit, there has been an overreliance on use of course evaluation as the only tool of measuring teaching effectiveness in higher education for years.

In a previous study, Stewart (2018) examined to what degree faculty demographics, faculty academic background, faculty academic status, faculty professional engagement, course grading patterns, and course characteristics predict the variance in overall course evaluation ratings by using archival data from an academic unit housed in a college of education during one academic year. The results of this study showed that faculty year of terminal degree accounted for 60.2% of the variance in overall course evaluation; faculty ethnicity and year of terminal degree accounted for 55% of the variance in standard deviations; and grading patterns predicted 24.5% of the variance in response rate, though it did not reach significance. Based on their findings, faculty that received their degree recently obtained higher course evaluations, faculty that received their degree recently obtained smaller standard deviations, and faculty that assigned fewer grades of D's and Fs's were likely to receive a higher response rate.

The purpose of this study is to replicate the previous study by integrating student variables and further examining how faculty gender, faculty race, faculty year of terminal

degree, student gender representation in class, student race representation in class, total enrollment, level of course, and grading patterns as independent variables predicts or significantly impact the dependent variable, the university faculty cumulative course evaluation, standard deviation of course evaluation, and response rate. The association between each of these variables are hypothesized to contribute to students' ratings of faculty member teaching performance (Anaya & Cole, 2001; Bachen, McLoughlin, & Garcia, 1999; Bennett, 1982; Boring, 2017; Fan et al., 2019; Feldman, 1984; Fernández & Mateo, 1997; Freeman, 1994; Gehrt, Louie & Osland, 2015; Lundberg & Schreiner, 2004; Marsh, 2001; Marsh & Roche, 2000; Martin, 2016; Noel & Smith, 1996; Remedios, Lieberman, & Benton, 2000; Nerger, Viney, & Riedel, 1997; Reid, 2010; Renaud & Murray, 1996; Ryan & Harrison, 1995; Schwitzer, Griffin, Ancis, & Thomas, 1999; Shapiro, 1990; Thornton, Adams, & Sepehri, 2010; Suarez-Balcazar, Orellana-Damacela, Portillo, Rowan, & Andrews-Guillen, 2003). This literature review provides an overview of prior empirical studies and conceptual articles associated with the study of university course evaluation.

Faculty Variables and Course Evaluation

Faculty Gender

Based on prior research, gender bias has been found to be associated with negative ratings on course evaluations. Students' perspective of their professor characteristics and aspects of teaching effectiveness may also contribute to how they rate

their professor. Basow (1995) examined the influence of professor gender, student gender, and divisional affiliation on course evaluation at a liberal arts college. The researchers used a student evaluation form consisting of questions relating to teacher behavior and questions relating to the course. Based on their results, female professors received lower ratings from male students while receiving higher ratings from female students across divisional affiliations. In contrast, gender bias did not impact the ratings of male professors as they received higher ratings similarly by female and male students across divisional affiliation. Female professors teaching humanities received higher ratings than male professors. Male professors teaching natural science received higher ratings than female professors. Results further suggest expectations of female professors and male professors differ as males received higher ratings on areas of appropriate speech, enthusiasm, thought stimulation, organization, and knowledge while females received higher ratings on areas of sensitivity, respect, fairness, and student freedom (Basow, 1995). Male students tended to rate female professors more negatively in areas of fairness and speech while female students tended to rate them positively in areas of sensitivity and respect. This pattern of same-gender teaching reveals a bias in preference that students may reflect on when completing evaluations of teaching effectiveness. Gender expectations, teaching styles, and areas of discipline in education each contribute to gender roles and impact students' evaluation of their experience in the course. When professors fall short of these expectations, their course evaluations are negatively

affected. These findings support the need for investigation of the impact of gender representation within the class on course evaluations of all faculty, but especially for female faculty.

When female faculty are held to a different standard than their male counterparts, expectations from students may decrease the likelihood of being consider within the promotion, tenure, and merit process especially in areas of academia where female faculty are underrepresented. For example, Morgan et al. (2016) conducted a study in a medical school that focused on difference of students' evaluations of male and female physician faculty within four clinical rotations. Based on their finding, female physicians received lower ratings on students' evaluations in all four clinical rotations including surgery, pediatrics, internal medicine, and obstetrics and gynecology. Ratings of female and male physician in the surgery rotation revealed a larger discrepancy compared to the other clinical rotations. No difference was found between the ratings of male and female students.

Adibifar (2019) examined if students' ratings on course evaluations are influenced by their professor's gender, age, and race. Based on their results, students provided higher ratings to professors who identified as male, white, or younger. It can be assumed that expectations of professors vary based on their gender due to societal norms that influence student perception or expectations of these roles. Some students may perceive male professors to be more competent than female professors and are likely to base their

ratings on this belief without regards to teaching effectiveness. Similarly, these biases negatively impact professors who are identifiable as a racial/ethnic minority. Despite being competent and exceeding expectations in effective teaching, professors of color are rated negatively due to previously held bias by students. Furthermore, professors who are younger are perceived to be more enthusiastic, flexible, and interactive when teaching, while older professors are less prone to use various methods of teaching. The preconceived notions that students have about male and female, racial/ethnic majority and racial/ethnic minority, and young and old professors furthers suggest that students' ratings are influenced by judgment and assumptions rather than the professor's teaching effectiveness. Previous studies have also shown similar results of this relationship, which found that a professor's gender and age to be a significant predictor of student evaluation of teaching effectiveness (Campbell, 2019; Murray et al., 2019).

When female faculty are expected to present themselves in an expected social role, then they may be rated lower than their male colleagues for not exhibiting those particular behaviors. For example, Mitchell & Martin (2018) conducted a study to focus on sources of gender bias and the different criteria women are evaluated than their male counterparts. The researchers hypothesized that female professors are rated differently regarding their competence and personality. Using evidence from content analysis of students' commentary on course evaluations, female professors were perceived to be less experienced in education, have a lower academic rank, and labeled as teachers instead of

professors. Female professors were also evaluated based on their personality, in which they were expected to be warmer, nurturing, and more accessible. Data from their analysis indicated that students' commentary involved addressing these personality attributes when providing feedback to their female professor, while male professors received commentary that address their competence. Some studies went further in ruling-out the possibility of instructor related attributes contributing to students' ratings on evaluations. Despite controlling for instructor characteristics and teaching styles for similar courses, female professors are still rated more critically on evaluations than male professors (MacNell, Driscoll, & Hunt, 2015). Their findings suggest that student evaluations of teaching effectiveness are biased against female professors. This further suggest that male professors are viewed as more competent and qualified, which decreases female professors having an equal opportunity at promotion, tenure, and merit.

Further investigation of gender bias is explored when Wagner, Rieger, and Voorvelt (2016) studied the effect of teacher gender and ethnicity on student course evaluation of teaching effectiveness while controlling for the course and content. The researchers found that gender bias was present while ethnicity was insignificant to ratings on course evaluations. Female professors were found to be evaluated more critically than male professors on course evaluations. Their findings also suggest that female professors are 11 percentage points less likely to obtain the teaching evaluation cut-off for promotion to associate professor when compared to male professors (Wagner et al.,

2016). Gender stereotypes influence students' perception that male faculty are more competent than female faculty, which reduces the likelihood of female faculty becoming tenure or receiving appropriate compensation for their service (Rivera & Tilcsik, 2019). Consequently, female faculty may have to work harder than male faculty to be perceived as more effective and to qualify for the same benefits during annual reviews. This supports the notion that students' ratings on course evaluations are influenced by their perception of gender roles rather than their professor's instructional skills. Therefore, administrator decisions that are based solely on course evaluations puts female faculty at a disadvantage in academia.

Joye and Wilson (2015) studied the effects of the professor gender and perceived age on students' ratings of teaching effectiveness, rapport, and academic performance. In this study, the researchers presented a picture of an instructor who was either male, female, young, or old while playing a gender-ambiguous audio lecture. For example, the students either saw a younger adult man, older adult man, younger adult women, or older adult women while the voices were altered on the lecture audio. Students were instructed to complete a quiz to assess instructor effectiveness, rapport with the instructor, and perceived age and attractiveness of the instructor. Findings led the researcher to conclude that students rated male professors to be more competent than female professors and rated younger female professors more attractive and greater with rapport than professors who were male or older.

Likewise, in previous studies, ratings were found to be significantly higher for young male professors compared to younger female professors, older female professors, and older male professors (Arbuckle & Williams, 2003). Students appear to differ in their expectations and gender preference, which influences student evaluations of teaching that result in an invalid measure of teaching effectiveness. Overall, course evaluations are found to be influenced by students' attitudes about gender instead of a valid representation of teaching effectiveness, which is specifically shown to be a discriminatory tool for female professors (Mengel, Sauermann, & Zölitz, 2019). Acknowledging potential bias is critical when reviewing course evaluations, as various factors that may influence a student rating whether it is gender bias, students' interest of the course, expected grade, or student-faculty interaction that may potentially affect the validity and reliability of an instructors' evaluation (Stewart, 2018).

Faculty Race

The race of a faculty member has also been found to be associated with students' subjective view when rating a faculty member teaching performance based on bias perceptions. Smith (2009) examined the effect of faculty race and gender on student ratings of teaching effectiveness. White male, White female, and male faculty who identified as other received higher ratings on overall course evaluations and overall teaching ability. Black male, Black female, and female faculty who identified as other received lower ratings on overall course evaluations and overall teaching ability.

Similarly, Chisadza, Nicholls, and Yitbarek (2019) conducted a study in South Africa to investigate the affects that race has on course evaluation ratings. Their results indicated that black lectures received lower ratings than white lecturers. These results suggest the likelihood that faculty of color are less likely to thrive or be considered for promotion, tenure, and merit compared to White faculty members due to racial bias contributing to students' ratings on course evaluations.

Further, investigation of race-based biases are explored when Wang and Gonzalez (2020) examined the ratings received by professors who identified as a racial/ethnic minority (i.e. African American, Asian American, Hispanic American) on student evaluation of teaching across multiple universities. Results indicated that professors, who belong to a racial/ethnic minority group, received lower ratings on course evaluations, while White American professors received higher ratings. Though limited attention is given to the effects race has on ratings, these findings warrant further research into race-based bias on groups that are negatively impacted by the influence this how on their career in academia Furthermore, these studies provide more insight of uncontrolled variables that administrators and faculty should consider when reviewing course evaluations for promotion, tenure, and merit and the importance of taking the necessary steps to eliminate these biases .

Faculty Year of Terminal Degree

In higher education, it is commonly inferred that faculty who have a doctoral level degree are more experienced and competent than faculty who have a masters level degree. Though there are few studies that have examined the effects of this variable, studies have shown that faculty years in terminal degree influence ratings of student evaluation of teaching effectiveness (Stewart, 2018). According to Lewis and McKinzie (2019), length of industry experience and years of teaching experience may influence student evaluations of teaching effectiveness. In this study, student evaluation data from business courses was collected from three universities to determine how student evaluations were impacted by faculty academic background. The results of this study found a positive correlation between the length of industry experience and years of teaching experience with students' evaluation of teaching effectiveness. Specifically, a faculty with more experience in the field received higher ratings on course evaluations. While faculty who had longer teaching experience received lower ratings on course evaluations.

Similarly, Andrade and Rocha (2012) showed that instructors with more experience tended to receive better ratings on student evaluations, which supports the notion that receipt of higher scores on evaluations are possibly due to competency and longevity in teaching. It can be assumed that faculty with more experience in an industry will be more positively evaluated because they are able to provide practical or actual

situations applicable to the course, which enhances students' perceptions of teaching effectiveness. However, faculty that have an extensive experience in teaching may begin to decline in other areas such as availability, communication, flexibility, enthusiasm, or work performance.

Rezaei, Haghdoost, Okhovati, Zolala, and Baneshi (2016) found support that academic degree led to better ratings on student evaluations while teaching experience did not reach significance. Professors and instructors received higher ratings on course evaluations than assistant professors. Yet, there are several additional studies in which results differ in which they report that professor and associate professors received higher ratings than instructor and assistant professor, while some findings concluded instructors received higher ratings than those with higher academic degree (Ghafourian Boroujerdnia, Shakurnia, & Elhampour, 2006; McPherson, Jewell, & Kim, 2009). Furthermore, other studies have shown a correlation between teaching experience and teaching effectiveness (Chaudhary & Rathore, 2018; Podolsky, Kini, & Darling-Hammond, 2019). Prior research has shown mixed results, yet their findings indicates that there is a relationship between experience and student evaluations of teaching effectiveness.

Student Variables and Course Evaluation

Student Gender

Previous research has shown numerous evidences of faculty demographic impact on student evaluations of faculty teaching effectiveness.

Yet, there is a limited number of studies that have specifically target student demographic as a variable that influence course evaluations. Santhanam and Hicks (2002) examined whether discipline, course year level, or gender of students had any influence on student ratings on teaching evaluations. The researchers used a student evaluation form consisting of questions relating to professor's teaching and questions relating to the specific unit/course. Based on their results, female students tended to provide higher ratings on course evaluation than male students. Results indicate a possible relationship between variables that could influence this outcome including interaction between student gender and faculty gender, course discipline, and difference in learning behavior across gender (Santhanam & Hicks, 2002). Likewise, prior research has also shown evidence that student characteristics impact their perspective of effective teaching. Female students provide higher ratings to their professors on course evaluation than male students who provided lower ratings on course evaluations (Basow & Silberg, 1987; Hancock, 1992; Summers, 1996; Tatro, 1995; and Thawabieh, 2017).

Female and male students bias and preference may be reflected when completing evaluations of teaching effectiveness, but it does not necessarily determine whether students learn more or less based on their ratings of teacher effectiveness. For example, Young, Rush, and Shaw (2009) conducted a study that focused on specific student and instructor characteristics that influence students' evaluations of teaching. Students rated their instructors on three factors: interpersonal characteristics (i.e. warm and friendly,

respect, humor, tolerance, comfortable atmosphere, adapt to student needs, concern for student learning, enjoyment, enthusiasm, motivation, accessible), pedagogical characteristics (i.e. well prepared, well organized, clear explanations, identify important ideas, subject matter knowledge, use of good examples, communication, self-confident), and course content characteristics (i.e. valuable course, improved understanding, increased interest, worthwhile materials). Based on their finding, female students rated their female instructors higher on pedagogical characteristics and course content characteristics than they rated their male instructors, while male students rated male instructors significantly higher on pedagogical characteristics and course content than their female instructors. It can be inferred that female and male student differ in their preference of teaching styles which supports the difference in ratings on course evaluations across students' gender. These results further suggest that it is imperative that administrators use multiple means to assess course instruction given the potential for gender bias.

Student Race

Due to the anonymous nature of university course evaluation, student identifiers are removed to allow students to be forthright when providing feedback about their overall experience in the course and the quality of teacher effectiveness without information being traced back to their response (Afonso, Cardozo, Mascarenhas, Aranha, & Shah, 2005). For this reason, it is possible that studies that have been conducted are

limited to only examining students' race in relation to their instructional preference and learning outcomes based on their interaction with instructors. Lundberg and Schreiner (2004) examined the quality and quantity of students' interaction with faculty members and its relationship to learning based on seven different racial groups of undergraduate students. The researchers found that student interaction with faculty was a greater predictor of learning for Asian/Pacific Islander students, Mexican American students, and Native American Students. Faculty interaction was a smaller predictor of learning for White students, while student background characteristics (i.e. age, gender, class level major, etc.) was a greater predictor of learning for White students. Yet, White students had a more positive perception of faculty relationship compared to other racial groups. For African American, Hispanic, and Puerto Rican students, working harder after receiving instructor' feedback was a greater predictor of learning. Results indicate that students of color with frequent interactions with faculty contributed significantly to student learning though results varied by race (Lundberg & Schreiner, 2004). Though students of color interacted frequently with faculty, Native Americans and African American students reported less satisfaction with the faculty-student relationship.

Expectations of faculty-student relationship varies across racial backgrounds as many students of color often interact with faculty who differ in race across universities (Noel & Smith, 1996; Schwitzer et al., 1999). Cultural differences may contribute to students' quality interaction with faculty, which may also impact learning outcomes and

ratings on course evaluations (Anaya & Cole, 2001; Suarez-Balcazar et al., 2003). It could be assumed that students' self-reported learning may possibly be associated with course evaluation outcomes; however further research is needed to support this notion. Based on the literature, it can be hypothesized that students of color who are more satisfied with faculty interaction would provide higher ratings on course evaluation. At this time, there is no study that has explored the relationship between students' race and course evaluations based on the use of the following academic search engines: Academic Search Complete, ERIC, Google Scholar, PsychArticles, PsychInfo, Scopus, SpringerLink, which makes these findings a critical contribution to the literature.

Course Variables and Course Evaluation

Total Enrollment

The impact of class size on course evaluations in higher education is a variable that has often been explored considering the variance in range of student enrollment across courses (Elmore & Pohlmann, 1978; Hamilton, 1980; Haslett, 1976; Marsh, Overall, & Kesler, 1979; Wood, Linsky, & Straus, 1974). In one study on the association between total enrollment and course evaluations of teaching effectiveness, Liaw and Goh (2003) examined the course characteristics (i.e., level of subject; type of subject; time of the lecture; day of the lecture; class size) and instructor characteristics (i.e., gender; rank; instructional experience) as bias factors that impacts overall teaching ratings on course evaluations using a regression model. Based on their results, class size had a significant effect on student's rating of teaching effectiveness. Specifically, faculty that taught

courses with smaller enrollment were more likely to receive more positive course evaluations than faculty that taught courses with larger enrollment. The other course and instructor characteristics were not found to be significant and were not considered to be bias factors contributing to ratings on course evaluations, and these results remained consistent after controlling for multicollinearity in this study. According to Liaw and Goh (2003), it is assumed that smaller classroom sizes are predictors of better-quality teaching effectiveness and may slightly reflect the outcome of student learning. This is possibly due to various benefits that students and faculty receive in a more reduced size classroom including better interaction and classroom engagement, improved learning environment, and enhance students' relationship with faculty members resulting in higher ratings on course evaluation of teaching effectiveness.

In a more controlled study, Bedard and Kuhn (2008) conducted an experiment to determine the impact of class size on student evaluations of teaching effectiveness after controlling for instructor and course fixed effects. The variables of importance in this study included class size, average evaluations score, and the course instructor to account for bias and instructor and course heterogeneity (Bedard & Kuhn, 2008). For example, each instructor differs in teaching styles and grading techniques, and it is common to have several instructors assigned to teach the same course within a semester. Based on the results, class size was still shown to have a negative impact on student evaluations of teaching effectiveness. Higher course evaluations could be a result of low-class size instead of a true reflection of teaching effectiveness (Mateo & Fernandez, 1996). If class

size effects are accounted for, this may display a true representation of student evaluations of teaching effectiveness when administrators interpret this data for the annual review.

Chapman and Ludlow (2010) further found support that larger class sizes negatively influence students' learning on course evaluations. After examining the influence of class size, and student and instructor characteristics on students' learning, the results showed that each variable was statistically significant. As a whole, student and instructor characteristics accounted for much of the variance for student learning, while class size independently accounts for student learning. The researchers also noted that other student and instructor characteristics did not explain for the effects of a larger class size. Research has shown that majority of undergraduate classes have larger class sizes and many students in these courses reported lower grades, which may? result in a deterioration in learning outcome and decrease ratings on course evaluations (Chapman & Ludlow, 2010). This is yet another factor that faculty have little to no control which could have an adverse impact on students learning and overall course evaluation based on the difference between student-to-faculty ratio across courses. Administrators are encouraged to take into consideration the pros and cons of adjusting class size and the effects it has on improving student learning aside from the quality of teaching effectiveness. Overall, these studies have emphasized the importance of not using course evaluations as the only means of assessing teaching effectiveness. When administrators interpret course evaluations other measures such as faculty self-assessment or classroom

observations of students' learning should also be considered in the process of promotion, tenure, and merit.

Course Level

While various faculty, student, and course variables have been explored to an extent of influence on course evaluations, literature reviews are found to be limited on the relationship between level of course and student evaluations. All students are un-trained in the evaluation of others but are asked to do so nonetheless and their conclusions are given great weight in decision-making related to a professional's career promotion. Nevertheless, graduate students, though possibly still not trained, would hopefully be more experienced in the process just due to engagement during prior years of undergraduate education. Because of this developmental and experiential change, the predictors of course evaluations may also change. This possibility supports the inclusion of course level in the model. Prior studies have focused more on other course characteristics such as class size, workload, and course difficulty. Bailey, Gupta, and Schrader (2000) examined differences in students' ratings of teaching effectiveness across course level, course content, and individual instructor. Results indicated significant differences across each of these factors, which further showed that students' expectations of the instructor engagement, course demands, and instructor personal characteristics impact their ratings of teaching effectiveness. It is implied that students, who are not intrinsically motivated or interested in the course, were more likely to have a more negative view of the workload, demands, and expectations of the course content.

Also, students' ratings differ between instructors who taught the same course, which may suggest personal bias or dissimilar criteria per professor which further impacts students rating on teaching effectiveness. Based on their finding, students' in lower level courses place higher expectations on their instructors to be more accessible, involved, and interact when teaching the course. Therefore, students' in lower level were more likely to provide lower ratings on evaluations than students' in higher level course when instructors did not meet their expectations. Students' level of maturity and intrinsic motivation is considered as a justification for this difference between course level (Bailey et al., 2000). It can be assumed that student's expectation of course content and faculty characteristics reflect their ratings on course evaluations of teaching effectiveness as they move forward to more advance courses in academia.

Badri, Abdulla, Kamali, and Dodeen (2006) further examined how specific factors influence student ratings on course evaluations of teaching effectiveness. In their study, the variables considered were the department, course level, course timing, class size, expected grade, self-reported GPA, and gender of the student. The results indicated significant influence and bias of ratings on course evaluations by each of the factors. Though class size and self-reported GPA were found to have a significant effect, it was uncertain whether higher ratings on course evaluations was specifically due to smaller class size or provided by students who reported higher GPAs. Yet, the results did show that students that reported lower GPAs and students that reported higher GPAs provided higher ratings on course evaluations than students who reported average GPAs (Badri et

al., 2006). The department or field of study (i.e. math, art, humanities, business) and courses that are offered during the day were more likely to receive higher ratings on course evaluations. Also, female students, students with higher grades, and students enrolled in higher-level courses tended to award higher evaluations. Several outcomes can potentially be predicted when each of the factors are correlated with each other which also impacts the overall course evaluation. It is imperative that faculty and administrators carefully interpret course evaluations of teaching effectiveness upon consideration of salary, promotion, tenure, and merit.

Grading Patterns

When interpreting the relationship between grading patterns and ratings on course evaluations, several theories can be formed. For instance, high course evaluations may be a valid measure of improved learning, enriched instruction, and better grades. Secondly, it could be assumed that course evaluations are based on students' perceptions and motivation for taking the course. Lastly, higher course evaluations could be a result of faculty rewarding higher grades while lower course evaluations could be a result of faculty rewarding lower grades. For example, a prior study found that faculty who assign lower grades were more likely to receive lower response rates. Based on their findings it could be hypothesized that students who received these lower grades were either less likely to complete the course evaluation or give faculty lower ratings, which consequently impacts the validity of their course evaluation (Stewart, 2018). Course evaluations are impacted by both faculty and student characteristics that are often known

to be variables that faculty have little to no control. Each of these factors are suggested to influence students' rating of course evaluation based on the conclusions of previous studies.

In a study of course evaluations, Griffin (2004) examined the impact of grading leniency and grade discrepancy on students' rating of course evaluations. Griffin (2004) found that professors, who were more lenient with grading, received higher course evaluations while professors, who were less lenient with grading, received lower course evaluations. They also found that students provided low ratings on course evaluations when they received lower than the expected grades; yet they did not provide higher ratings on course evaluation when their professors rewarded them higher grades than expected. Similarly, Blackhart, Peruche, DeWall, and Joiner (2006) also concluded that professors are rewarded higher ratings on course evaluations when they are more lenient in grading and provide higher average grade among the class. When compared to other variables (i.e., gender, class difficulty, class workload, class size, students' motivation), grading patterns have shown to be the most significant predictor in influencing overall course evaluations. The relationship between expected grade and deserved grade on course evaluations have supported the theory that bias exists within this measurement tool that reduce its purpose in evaluating teaching effectiveness. Furthermore, these factors then contribute negatively to faculty opportunity for promotion, tenure, and merit.

Prior studies have resulted in mixed findings in which grades have shown little to no significance in contributing to low course evaluations. For example, Remedios and

Lieberman (2008) explored factors that influence course evaluations and the relationship between these factors by assessing grades, workload, course difficulty, student expectations', and students' goals. Their findings indicated that course evaluations were mostly impacted by the professor's quality of teaching based on the students' involvement and level of enjoyment of the course content. Yet, the results also suggested that grades, workload, course difficulty, and student's goals had a slight impact on students' rating on course evaluations. Students' expectations were found not to have a direct effect on course evaluations in this model. Students who were more likely to learn more during the course were more likely to receive higher grades and provide higher ratings on course evaluations; while students who struggled with the workload and found the course difficult were more likely to lower grades and provide lower ratings on course evaluations. Remedios and Lieberman (2008) explored their research from a valid perspective in which students rate course evaluations based on how well their professor taught the course content, and from a bias perspective in which students rated course evaluations solely on the grades they received for the course. This supports the assumption that grades in conjunction with course workload and expectations can impact students' rating on course evaluations of teaching quality.

Correspondingly, Boysen (2008) studied students' reasons for giving professors low evaluations and how often they engage in this vengeful act. Boysen (2008) found that only a few students reported that they provided professors low evaluations to retaliate for the low grades they received in a course. Results indicated that low evaluation were

predominantly influenced by poor teaching style, yet other factors also contributed to students' justification of providing low course evaluations. From the students' perspective, low course evaluations were also given due to their professors being rude, unprepared, unfair in grading, poor communication, unknowledgeable, unavailable, poor attitude, hard grading, and not teaching test materials. Other rationales consisted of workload too heavy, not learning enough, objectives unclear, class being pointless, and class not matching syllabus. Findings also indicated that professors received low evaluations from students who received higher grades and students that received lower grades. Several studies have concluded a positive correlation between course evaluations and grades which suggest that higher grades results in higher student evaluations. This may also contribute to an increase in grade inflation and influence professors to reward better grades in return for higher course evaluations. Grading leniency is often a factor that is discussed in the controversial argument of course evaluations being an unreliable measuring tool of teaching evaluations. It important that administrators address these potential bias factors when reviewing annual evaluations during the promotion, tenure, and merit process. Based on findings from the literature, it may be hypothesized that grading patterns will have a positive yet small contribution to overall course evaluations.

Dependent Variables

Course Evaluation

Data collected from university course evaluations provides an overall rating of students' response of the course and teaching effectiveness. The cumulative course

evaluation provides the average response obtained from students within the course that participated in the evaluation. Faculty that receive a lower score may indicate negative ratings from students while higher scores may suggest positive ratings from students. Given the findings from previous studies, it can be assumed that multiple variables (i.e. faculty gender, faculty race, faculty year of terminal degree, faculty rank, course level, course delivery, course class size, grading patterns, etc.) may contribute to the outcome of overall course evaluations (Beran & Rokosh, 2009; Beran & Violato, 2005; Brockx, Spooren, & Mortelmans, 2011; Isely & Singh, 2005; Johnson, Narayanan, & Sawaya, 2013; MacNell et al., 2015; Mohan, 2011; Prince, Felder, & Brent, 2007; Reid, 2010; Stewart, 2018; Young & Duncan, 2014). For example, one study revealed that faculty year of terminal degree predicted the variance in overall course evaluation which indicated that faculty that received their degree recently received higher scores on course evaluations (Stewart, 2018). However, some studies differ and found that course evaluations are a true measurement of course content and teacher effectiveness and not skewed by other variables.

Park and Dooris (2019) conducted a study to determine which variables (i.e. instructor characteristics, student characteristics, course characteristics, and course and instruction items on the evaluation) may significantly predict high teaching and overall course evaluations by using a decision tree analysis. Based on their research, they found that the items on the course evaluation which address faculty and student interaction, course delivery, course organization and planning, course assignment and examinations,

and grading significantly predict high overall evaluations and teaching effectiveness. Specifically, the items on the evaluation that predict the overall rating of teaching included helping the student better understand course material and providing an intellectually stimulating class. The items that predicted the overall rating on course evaluation included helping the student better understand course material and assignments that help students learn the material (Park & Dooris, 2019). Though student and instructor characteristics were found not to be significant predictors, results from this study indicate that specific items on course evaluation were more important in predicting high teaching and course evaluations for female and male instructors. For example, the items regarding a well-organized course and providing a stimulating class impacts overall course evaluations for female instructors while the item about helping students understand the course material is an important item for male instructor (Park & Dooris, 2019). This suggest that items on the course evaluation may be interpreted differently by students based on their gender and the instructor gender; yet further studies should be conducted to compare these findings. Nonetheless, the results from this study found items on the course evaluations to be valid predictors of teaching effectiveness and researchers did not detect any biases related to instructor, student, or course characteristics.

As course evaluations continue to be used by institutions for summative and formative purposes, more research should be conducted to provide further information regarding the various factors that should be consider when interpreting this tool for teaching effectiveness and promotion, tenure, and merit.

Standard Deviation

While the cumulative course evaluation may indicate a high score, it is not sufficient to support the distribution of students' response. Standard deviation influences the direction of the cumulative course evaluation which further suggest it's importance when faculty interpret the response of their students (Franklin, 2001). Therefore, the standard deviation is a significant dependent variable to review when examining course evaluations to gather more information regarding the variance around the mean (Stewart, 2018). For example, if a small standard deviation is present this suggest that students rated the course/faculty the same on the evaluation. However, a large standard deviation suggests that the students' response of the course/faculty differed on the evaluation and indicate that administrators and faculty should consider reviewing plausible reasons for the variance in ratings. When a faculty member receives a high mean on their course evaluations, a small standard deviation is preferred, which indicates that majority of their students found the course and faculty teaching effectiveness satisfactory. For faculty who receive a low mean on their course evaluation, a large standard deviation is preferred which indicates that some students found the course and faculty teaching effectiveness satisfactory while other students found it unsatisfactory. Therefore, careful investigation is warranted by faculty, administrators, and study reviewers.

The results of the examination of this variable represent an appropriate illustration of the distribution of responses and would seem imperative to review given the critical decision being made based on the ratings faculty receive on course evaluations. Yet, this

variable has typically not been considered among common factors that are measured in studies when predicting variables that contribute to the overall course evaluations (Fjortoft, 2015). A prior study found that faculty who received their degree most recently were more likely to have smaller standard deviations which indicated consistency in students' response on course evaluations (Stewart, 2018). The results from this study suggests instructor characteristics and potential bias that may skew the distribution of students' ratings on course evaluations. It further indicates the importance of interpreting the standard deviation and cumulative course evaluation simultaneously. Because comparable studies are not available, it is essential that further research is conducted to examine other variables that may significantly influence the variance in standard deviations of course evaluations.

Response Rate

When examining overall course evaluations, it is also important to interpret the response rate considering the impact of the number of student participants has on the validity of the evaluation. The method in which course evaluations are administered has changed from paper-format to electronic distribution which has been found to influence students' motivation and likelihood of completing the evaluation. Previous studies have shown that this switch from paper to electronic format has resulted in a reduction in response rates for course evaluations (Adams & Umbach, 2012; Anderson, Cain, & Bird, 2005; Avery, Bryant, Mathios, Kang, & Bell, 2006; Guder & Malliaris, 2013). As a result of the low response rates, the feedback provided when examining the overall course

evaluation is less likely to be a true representation of the course, result in bias, and will be ineffective in providing information for promotion, tenure, and merit. Research studies have been conducted to identify factors that are associated with increasing response rates on university course evaluations (Anderson, Brown, & Spaeth, 2006; Chapman & Joines, 2017; Dommeyer, Baum, Hanna, & Chapman, 2004; Johnson, 2003; Norris & Conn, 2005; Nulty, 2008). For example, Young, Joines, Standish, and Gallagher (2019) examined whether response rate would increase if faculty offered students in-class time to complete course evaluations compared to a group of faculty that did not offer this option to their students. Based on their results, they found that response rates increased when faculty provided in-class time for students to complete the evaluation. Faculty that discussed the significance of course evaluations, provided several reminders, and dedicated time during the term to allow students to partake in the evaluation were more likely to increase student motivation to participate and reduced the chances of them forgetting to complete the evaluation (Young et al., 2019). These findings further suggest that students are more likely to be motivated in completing course evaluations once they realize the value of their feedback and become aware that their responses are being taken into consideration in improving course content and teaching effectiveness.

Goodman, Anson, and Belcheir (2015) further examined various techniques faculty use to increase response rate with consideration of course and instructor characteristics. The results of this study found that using incentives were the most effective technique in increasing response rates on course evaluations. Specifically, extra

points are a common incentive faculty use to increase student completion of course evaluations. Other techniques such as explanation of the importance of course evaluation, providing reminders, sending students e-mails to complete the evaluation, and providing in-class time to complete the evaluation were also found to be effective in increasing response rates (Goodman et al., 2015). In addition, the number of techniques faculty use further increase response rates by two to five percent across undergraduate and graduate courses. In regard to course characteristics (i.e. class size, class level, instruction mode) and instructor characteristics (i.e. tenured/tenure track, adjunct, gender), results indicated that small classes, face-to-face courses, graduate courses, and courses taught by tenure/tenure-track faculty had higher response rates. However, the approach or strategy faculty use to increase response rates has a greater impact than course and faculty characteristics. Nevertheless, some faculty members choose not to provide incentives possibly due to concerns regarding grade inflation or the validity of the ratings (Jaquett, VanMaaren, & Williams, 2016; Love & Kotchen, 2010). This study further addresses some of these concerns by noting that non-point and point-based incentives are equally effective. The class-wide approach in providing incentives were also found to be as equally effective as an individual-based approach in monitoring student completion of course evaluations.

Further investigation of tools utilized to increase response rates is explored when Crews and Curtis (2011) examine faculty perspective on improving student motivations to complete course evaluations. Their findings align with previous studies in which

faculty reported using multiple approaches including: explaining the importance of course evaluation, including and reviewing the information about course evaluation in the syllabi, providing reminders, sending emails with the evaluation link, giving incentives such as extra points, and allowing class time for students to complete the evaluation (Crews & Curtis, 2011). Researchers indicated that though some faculty members were apprehensive about providing incentives other strategies are taken into consideration and utilized across disciplines in higher education. When evaluating the response rate, it is still important to consider potential bias and other possible factors including instructor, student, and course characteristics that may contribute to the overall course evaluation.

Summary and Research Questions

Currently there are no studies that have examined faculty, student, and course variables, simultaneously. Therefore, this study will add in a meaningful and significant way to the literature by exploring all the potential variables associated with the variance of course evaluations at one time. The specific purpose of this research is to determine:

1. To what degree do faculty variables (faculty gender, faculty race, faculty year of terminal degree), student-classroom variables (i.e. student gender representation in class % of female, student race representation in class % of racial majority, total enrollment), and course variables (i.e. course level, grading patterns) predict the variance in overall course evaluations?
2. To what degree do faculty variables (faculty gender, faculty race, faculty year of terminal degree), student-classroom variables (i.e. student gender representation

in class % of female, student race representation in class % of racial majority, total enrollment), and course variables (i.e. course level, grading patterns) predict the variance in the standard deviation of the course evaluation?

3. To what degree do faculty variables (faculty gender, faculty race, faculty year of terminal degree), student-classroom variables (i.e. student gender representation in class % of female, student race representation in class % of racial majority, total enrollment), and course variables (i.e. course level, grading patterns) predict the variance in response rates?

Based on the literature review and results of prior studies, it is hypothesized that:

1. In the model, faculty, student, and course variables will significantly predict the variance in overall course evaluations.
2. In the model, student variables will be the sole predictors of the variance in the standard deviation of the course evaluations.
3. In the model, course variables will significantly predict the variance in response rates.

Chapter III

Methods

Procedure

This study was conducted from archival data from an academic unit housed in a college of education. IRB approval and permission to access the course evaluation data was obtained before gathering faculty information from this academic unit. The university website was used to collect instructors' vita using a secure technology device that has internet access. Individuals' identity was protected by coding and through maintaining collected campus-wide data, instead of identifying results by program or department. To ensure confidentiality of the data, faculty and students names and any other identifiable information was removed.

Given that archival data was used, informed consent was not needed. No compensation was involved within this study. The risks were minimal given that most of the faculty and student demographic and content included in this study were archival data. Once copies of teaching evaluations, annual evaluations, and faculty vitae were printed, materials were secured in a locked cabinet in the principal investigator's office, which was always locked and secured.

The results of this study will provide a positive contribution to the literature and will increase understanding of the meaning of course evaluations in the review process for annual evaluations and promotion, tenure, and merit. The information is valuable to

students who aspire to become faculty, tenure-track faculty, and administrators who establish policy related to employment, promotion, tenure, and merit.

The data analyzed was based on the archival data of a total of fifty-three (n = 53) faculty members. Within this sample of faculty members, the following information was gathered: 40 (75%) females and 13 (25%) males; and 46 (87%) faculty members, who identified as racial majority (i.e., White/Caucasian) and 7 (13%) faculty members who identified as racial/ethnic minority (i.e., Black/African-American, American Indian or Alaskan Native, Asian, and Hispanic). Each instructor's course load assignment varied and course evaluations were collected during the 2018-2019 academic year from a total of two hundred two (n = 202) courses department wide. Within this sample of courses, the following information was gathered: 105 (52%) undergraduate course level, 85 (42%) master course level, and 12 (6%) doctorate course level.

University Setting

The state university is a comprehensive, regional institution which enrolls approximately 13,000 students and offers more than 120 areas of study and concentrations within six academic colleges including business, education, fine arts, forestry and agriculture, sciences and mathematics, and liberal and applied arts. The student-to-faculty ratio is 18:1 and an average class size of 27.

Description of College. The college of education consists of five departments including: Elementary Education, Human Science, Human Services, Kinesiology & Health Science, and Secondary Education & Educational Leadership Department. During

the school year in which the data was obtained, the college of education had a student body of approximately 3, 701 in which 3,011 were females (81.36%) and 690 were males (18.64%). Of these students, 2,300 identified as White (62.15%), 629 as Black (17.00%), 551 as Hispanic (14.89%), and 221 as other (5.97%). Students enrolled in this college included 2,709 undergraduates (73%) and 992 graduates (27%). Approximately 1,077 (29%) students were housed in the Elementary Education Department, 883 (24%) students in the Kinesiology & Health Science Department, 658 (18%) students in the School of Human Sciences, 681 (19%) students in the Human Services Department, and 387 (10%) in the Secondary Education & Educational Leadership Department. The demographic of the faculty within the college included 19 full professors (8.88%), 32 associate professors (14.95%), 10 teaching assistants (4.67%), 86 adjunct faculty (40.19%), 6 instructors (2.80%), 7 lecturers (3.27%), and 19 as other (8.88%). Of these faculty members, 124 were non-tenure track (57.94%), 52 were tenured (24.30%) and 38 were on-track (17.76%). About 103 of faculty had no terminal degree (48.13%), and 111 faculty had completed a terminal degree in their field of study (51.87%).

Description of Department. The department of this study had a student body of approximately 681, with 511 being female (90.60%) and 64 being male (9.40%). Of these students, 463 identified as White (67.99%), 88 as Black (12.92%), 86 as Hispanic (12.63%), and 44 as other (6.46%). Students enrolled in this department included 681 undergraduates (66%) and 357 graduates (34%). The demographic of the faculty for the department included five full professors (10.64%), two teaching assistants (4.26%), 14

adjunct faculty (29.79%), eight assistant professors (17.02%), five associate professors (10.64%), one instructor (2.13%), two lecturer (4.26%), and ten as other (21.28%). Of these faculty members, 28 were non-tenure-track (59.57%), 11 were tenured (23.40%), and 8 were on-track (17.02%). About 24 of the faculty had no terminal degree (51.06%), and 23 faculty had completed a terminal degree in their field of study (48.94%).

Dependent Variables

The course evaluation contains questions about both course content and instructor effectiveness for the use of program improvement and for instructor evaluation. The course evaluation used in this study resembles a traditional measure. The college of education course evaluation rating is on a 5-item scale ranging from 5 = *very good*, 4 = *good*, 3 = *average*, 2 = *poor*, to 1 = *very poor* (see Appendix for college of education course evaluation). A column is also provided for students to select no opinion and an option to provide additional comments. The dependent variables being influenced in faculty course evaluations include: *cumulative course evaluation* – the overall average evaluation score of the course, or the average score of students' responses; *standard deviation of the course evaluation* – mean of faculty cumulative course evaluation; and *response rate* – the percentage of students in a course responding to the evaluations, or the number of students who completed the course evaluation divided by the number of people in the course.

Independent Variables

Variables manipulating the variance in faculty course evaluations include: *faculty gender* – the sex of faculty members (female vs. male); *faculty race* – the ethnicity of faculty members (racial/ethnic minority vs. racial/ethnic majority); *faculty year of terminal degree* – the year of their terminal degree; *student gender representation per course* – percentage of students identified as female; *student race representation per course* – percentage of students identified as racial majority; *total enrollment* – total number of students enrolled per course; *course level* – undergraduate vs. masters vs. doctoral; and *grading patterns* – the number of high grades (As and Bs) vs. the number of low grades (Cs, Ds, and Fs) assigned during the academic year course evaluations were collected.

Analysis

Using the archival data, in a larger regression model with cumulative overall course evaluation, standard deviation of the course evaluation, and response rate as the dependent (criterion) variable, and faculty gender, faculty race, faculty year of terminal degree, student gender representation in class, student race representation in class, total enrollment, course level, and grading patterns will be included as independent variables in a regression equation. The variables incorporated in the model were the following: Instructor-related variables included faculty members' personal demographics (i.e., gender, race/ethnicity status) and professional related information including faculty

members' academic background (i.e., year of terminal degree). Coding was assigned for each of the independent variables as follows:

Faculty gender: coded: 1 = female, 2 = male

Faculty race: coded: 1 = racial/ethnic majority, 2 = racial/ethnic minority

Student-related variables will include students' personal demographics per class (i.e., student gender representation in class: percentage of students identified as female, student race representation in class: percentage of students identified as racial majority).

Course-related variables included course total enrollment, grading patterns (i.e., high grades vs. low grades), and course characteristics (i.e., course level). Coding will be assigned for each of the independent variables as follows:

Course level: coded: 1 = undergraduate (100-400 level), 2 = master (500 level), 3 = doctorate (600 level)

Independent studies, practicum, internship, theses, and dissertations were not included in the study.

Chapter IV

Results

Preliminary Analysis

First, Pearson product-moment correlation data analyses were conducted to examine the relationships between each of the independent and dependent variables to determine the strength of each pair of variables and assess for correlations that met the criteria of multicollinearity. Multicollinearity occurs when there are high correlations between independent variables to be included in a regression analysis, resulting in misinterpretation of the *p*-values to identify independent variables that are statistically significant in the regression model (Kim, 2019; Vatcheva, Lee, McCormick, & Rahbar, 2016). The criterion for determining multicollinearity includes a cutoff value of $r = .80$ or variance inflation factor (VIF) greater than 5 (Thompson, Kim, Aloe, & Becker, 2017). The correlations that met the criterion for multicollinearity were total enrollment and high grades. High grades are included because this variable is supported by the literature; therefore, total enrollment is excluded from the main analyses. Then, enter regression analyses were performed to examine the degree to which the independent variables (i.e., faculty gender, faculty race, faculty year of terminal degree, student gender representation in class, student race representation in class, total enrollment, course level, grading patterns) predict the three dependent variables: (i.e., overall course evaluation, standard deviation, and response rates.

Overall Course Evaluation Analysis

Pearson Product Correlations were conducted for the total sample between overall course evaluation and the variables: gender (r = -.251, p < .001), year of terminal degree (r = .178, p = .011), low grades (r = -.406, p < .001), standard deviation (r = -.861, p <.001), and response rate (r = .210, p = .003). Findings indicate that faculty with higher course evaluations tend to be female; completed terminal degree recently; assign fewer low grades; have a smaller standard deviation, and receive higher response rates (see Table 1 for the Pearson Product correlation matrix indicating the relationships between all pair of variables).

Table 1

Pearson Product Correlation Matrix

Variables	1	2	3	4	5	6	7	8	9	10	11	12
1. Gender [a]	-	.201**	-.109	-.075	-.003	.074	.000	.058	.096	-.251**	.209**	-.023
2. Ethnicity [b]	.201**	-	.210**	-.097	.017	.096	.187**	.090	.074	-.044	.086	-.001
3. Year Terminal Degree	-.109	.210**	-	-.196**	-.070	-.180*	.270**	-.115	-.254**	.178*	-.140*	.213**
4. Percentage Female	-.075	-.097	-.196**	-	.433**	.433**	-.512**	.410**	.207**	-.004	.002	-.145*
5. Percentage Racial Majority	-.003	.017	-.070	.433**	-	.137	-.150*	.184**	-.135	.057	-.072	-.079
6. Total Enrollment	.074	.096	-.180*	.433**	.137	-	-.499**	.968**	.427**	-.111	.191**	-.084
7. Course Level [c]	.000	.187**	.270**	-.512**	-.150*	-.499**	-	-.410**	-.500**	.098	-.111	.375**
8. High Grades [d]	.058	.090	-.115	.410**	.184**	.968**	-.410**	-	.192**	-.001	.092	.012
9. Low Grades [e]	.096	.074	-.254**	.207**	-.135	.427**	-.500**	.192**	-	-.406**	.390**	-.372**
10. Overall Course Evaluation	-.251**	-.044	.178*	-.004	.057	-.111	.098	-.001	-.406**	-	-.861**	.210**
11. Standard Deviation	.209**	.086	-.140*	.002	-.072	.191**	-.111	.092	.390**	-.861**	-	-.194**
12. Response Rate	-.023	-.001	.213**	-.145*	-.079	-.084	.375**	.012	-.372**	.210**	-.194**	-
N	202	202	202	201	201	202	202	202	202	202	202	202

Note. [a] Female = 1, male = 2. [b] Racial majority = 1, racial minority = 2. [c] Undergraduate (100-400 level) = 1, master (500 level) = 2, doctorate (600 level = 3).[d] Number of As and Bs assigned. [e] Number of Cs, Ds, and Fs assigned.

* $p < .05$. ** $p < .01$.

A enter regression was conducted to determine the degree that faculty variables (faculty gender, faculty race, faculty year of terminal degree), student-classroom variables (i.e., student gender representation in class % of female, student race representation in class % of racial majority, total enrollment), and course variables (i.e., course level, grading patterns) predict the variance in overall course evaluations. Results of the enter regression analysis indicated that faculty gender and low grades significantly predicted the variance in overall course evaluations, F (8, 192) = 7.573, $p < .000$, $R^2 = .240$ (see Table 2). The regression indicated that faculty gender and low grades significantly predicted the variance in overall course evaluations and accounted for 24% of the variance. Faculty that are female and faculty that assigned fewer low grades obtained higher course evaluations.

Table 2

Regression Analysis for Overall Course Evaluation

Variable	B	SE B	β
Gender	-.220	.067	-.217
Ethnicity	.061	.089	.048
Year Terminal Degree	.002	.003	.046
Percentage Female	.056	.283	.016
Percentage Racial Majority	-.102	.176	-.042
Course Level	-.108	.064	-.149
High Grades	.002	.003	.040
Low Grades	-.080	.013	-.474

Note. $R^2 = .24$

Standard Deviation Analysis

Pearson Product correlations were conducted for the total sample between standard deviation of course evaluations and the variables: gender ($r = .209$, $p = .003$),

year terminal degree (r = -.140, p = .048), total enrollment (r = .191, p = .006), low grades (r = .390, p < .001), overall course evaluation (r = -.861, p < .001), and response rate (r = -.194, p = .006). Findings indicate that faculty with greater standard deviations in course evaluations tend to be male; completed their terminal degree earlier; have a greater number of student enrollment; assign more low grades; have lower overall course evaluation; and lower response rates (see Table 1 for the Pearson Product correlation matrix indicating the relationships between all pair of variables).

A enter regression was conducted to determine the degree that faculty variables (faculty gender, faculty race, faculty year of terminal degree), student-classroom variables (i.e., student gender representation in class % of female, student race representation in class % of racial majority, total enrollment), and course variables (i.e., course level, grading patterns) predict the variance in the standard deviation of the course evaluation. Results of the enter regression analysis indicated that faculty gender and low grades predicted the variance in standard deviations of course evaluations, F (8, 192) = 5.899, p < .000, R^2 = .197 (see Table 3). The regression indicated that faculty gender and low grades predicted the variance in standard deviations of course evaluations and accounted for 19.7% of the variance. Faculty that are male and faculty that assigned more low grades obtained smaller standard deviations.

Table 3

Regression Analysis for Standard Deviation

Variable	B	SE B	β
Gender	.106	.045	.159
Ethnicity	-.001	.060	-.002
Year Terminal Degree	-.001	.002	-.037
Percentage Female	-.122	.192	-.055
Percentage Racial Majority	.019	.120	.012
Course Level	.055	.044	.116
High Grades	.002	.002	.066
Low Grades	.047	.009	.425

Note. $R^2 = .20$

Response Rate Analysis

Pearson Product Correlations were conducted for the total sample between response rates and the variables: year terminal degree ($r = .213$, $p = .002$), student gender representation ($r = -.145$, $p = .040$), course level ($r = .375$, $p < .001$), low grades ($r = -.372$, $p < .001$), overall course evaluation ($r = .210$, $p = .003$), and standard deviation ($r = -.194$, $p = .006$). Findings suggest that faculty with higher response rates tend to received their degree recently; have fewer percentage of students who identified as female; teach higher level courses; assign fewer lower grades; receive higher course evaluations, and have a smaller standard deviation (see Table 1 for the Pearson Product correlation matrix indicating the relationships between all pair of variables).

A enter regression was conducted to determine the degree that faculty variables (faculty gender, faculty race, faculty year of terminal degree), student-classroom variables (i.e., student gender representation in class % of female, student race

representation in class % of racial majority, total enrollment), and course variables (i.e., course level, grading patterns) predict the variance in response rates. Results of the enter regression analysis indicated that course level, high grades, and low grades predicted the variance in response rates, F (8, 192) = 7.550, $p < .000$, $R^2 = .239$ (see Table 4). The regression indicated that course level, high grades, and low grades predicted the variance in response rates, and accounted for 23.9% of the variance in response rates. Faculty that teach higher level courses, assigned higher grades, and assigned fewer low grades obtained higher response rates.

Table 4

Regression Analysis for Response Rate

Variable	B	SE B	β
Gender	.004	.018	.014
Ethnicity	-.028	.024	-.081
Year Terminal Degree	.001	.001	.088
Percentage Female	.047	.076	.052
Percentage Racial Majority	-.076	.047	-.116
Course Level	.065	.017	.333
High Grades	.003	.001	.211
Low Grades	-.011	.004	-.251

Note. $R^2 = .24$

Chapter V

Discussion

The current study was conducted to further interpret the meaning of course evaluations and determine which variables are associated with students' evaluations of teaching effectiveness. Based on previous literature addressing student course evaluations, there are currently no studies that have examined faculty, student, and course variables simultaneously. This study provides a more in-depth examination of all the potential variables associated with course evaluations variance. The following research question was addressed in this study: To what degree do faculty variables (faculty gender, faculty race, faculty year of terminal degree), student-classroom variables (i.e., student gender representation in class % of female, student race representation in class % of racial majority, total enrollment), and course variables (i.e., course level, grading patterns) predict the variance in overall course evaluations, the standard deviation of the course evaluation, and response rates?

It was hypothesized that in the model: faculty, student, and course variables will significantly predict the variance in overall course evaluations; student variables will predict the variance in standard deviation, and course variables will predict the variance in response rates. According to the results, faculty gender and low grades predict the variance in overall course evaluation and standard deviation. In contrast, course level, high grades, and low grades predict the variance in response rate. The regression results

indicate that faculty that are female and faculty that assigned fewer C's, D's, and F's obtained higher course evaluations. Faculty who are male and faculty who assigned more C's, D's, and F's obtained smaller standard deviations. Faculty that teach master and doctorate level courses, faculty that assigned more A's and B's, and faculty that assigned fewer C's, D's, and F's obtained higher response rates.

This study concluded that faculty gender, course levels, and grading patterns significantly contributed to the amount of variance in overall course evaluation, standard deviation, and response rate. There was no statistically significant relationship found between the dependent variables and the following variables: faculty race, faculty year of terminal degree, student gender representation per course, and student race representation per course. While these independent variables were not found to be significant, it is important to note that these variables were included as some were found to be predictors that impact student evaluation of teaching effectiveness and faculty members' promotion, tenure, and merit in previous studies (Chisadza, Nicholls, & Yitbarek, 2019; Lewis & McKinzie, 2019; Lundberg & Schreiner, 2004; Santhanam & Hicks, 2002; Stewart, 2018; Wang & Gonzalez, 2020). Moreover, it is important to consider factors that have yet to be addressed within the literature and examine how each of these variables contributed to university course evaluations within this academic unit.

Faculty Gender and University Course Evaluation

Prior research has shown faculty gender to influence students' ratings on course evaluations. Students' perception of teaching effectiveness differs for female and male

faculty members, which has also resulted in negative ratings on course evaluations. Based on societal and cultural norms, there are gender role expectations that are expected to be demonstrated by women and men. It is implied that female faculty exhibit traits that are perceived as feminine, such as nurturing, warm, friendly, empathetic, and understanding (MacNell, Driscoll, & Hunt, 2015). In contrast, male faculty are expected to exhibit traits that are perceived as masculine such as dominant, critical, objective, intelligent, and professional (Tran & Do, 2020). Researchers have found that male faculty and female faculty have received a different rating on areas of teaching across divisional affiliations (i.e., humanities, science, art, math) as a result of these gender schemes (Basow, 1995; Mitchell & Martin, 2018). Studies have also shown female faculty to be evaluated more critically and receive lower ratings than male faculty on course evaluations though, in some studies, instructor characteristics were controlled (Adibifar, 2019; MacNell, Driscoll, & Hunt, 2015; Morgan et al., 2016). In this specific study, course evaluation and standard deviation were found to be associated with faculty gender. Female faculty received higher course evaluations as a reflection of their teaching effectiveness, and male faculty obtained smaller standard deviations, which indicates that their students tended to have similar perceptions of the course/faculty within this academic unit.

Findings suggest that expectations of gender roles may influence students' perception of teaching effectiveness when rating their professor and could therefore decrease the likelihood of faculty members having an equal opportunity at promotion, tenure, and merit when they fall below these expectations. Information from the results of

this study further supports the importance of considering the influence faculty gender has on the validity and reliability of university course evaluations and when making decisions regarding promotion, tenure, and merit.

Course Level and University Course Evaluation

Exploring the relationship between university course evaluations and course level has shown to be associated with response rates. These findings indicate that faculty that taught higher-level courses tended to have higher response rates on course evaluations. Results indicate that faculty who teach master and doctorate level courses are more likely to have students who participate and complete course evaluations than faculty who teach undergraduate courses. It could be assumed that graduate students are more likely to have more experience in completing course evaluations and providing constructive feedback than undergraduate students due to prior years of academic experience. Though research is limited in examining the relationship between course level and university course evaluations, studies have shown that students' maturity, expectation of the course, and motivation for taking the course could be associated with their ratings on course evaluations (Bailey et al., 2000). This study further contributes to this research by identifying the importance of considering the impact of course levels when evaluating teaching effectiveness.

Grading Patterns and University Course Evaluation

Grading patterns have also been found to be associated with response rates and contribute to the likelihood of students completing a course evaluation. Faculty that

assigned more C's, D's, and F's obtained smaller standard deviations. This result suggested that students tended to have similar perceptions of the course/faculty and provided consistent ratings of either being satisfied or unsatisfied with their grades. Faculty who assign more A's and B's and assign fewer C's, D's, and F's are more likely to have higher response rates. Given the results of this study, the data suggest that students who receive higher grades are more likely to participate in providing course evaluations than students who receive lower grades. Prior studies have shown that response rates are impacted by students' satisfaction with the course and the quality of teaching and may slightly be impacted by their grades (Remedios & Lieberman (2008). While students who receive lower grades are less likely to participate in completing course evaluations, previous studies have shown that students may provide lower ratings on course evaluation to retaliate due to their low grade (Boysen, 2008; Stewart, 2018). Grading leniency is another factor to consider, resulting in higher course evaluations and response rate and subsequently affecting the validity of faculty course evaluation (Blackhart, Peruche, DeWall, & Joiner, 2006; Griffin, 2004). Nonetheless, grading patterns influence course evaluations and it is recommended that faculty and administrators consider the impact grading patterns may have on course evaluations and the annual review process when assessing teaching effectiveness at the end of each term.

Implications

With university course evaluations being a persistent topic of discussion in measuring teaching effectiveness, this study's findings suggest implications that would be

relevant for academicians and administrators in higher education. Students are pertinent in providing feedback regarding their learning experience in the classroom and the quality of the instructors teaching effectiveness (Coates, 2005). Given that students are the active participants and observers, their perception of the course helps guide approaches to improve the overall quality of education. Considering this, outcomes from students' ratings on course evaluations impact the process of promotion, tenure, and merit (Alshare, Wenger, & Miller, 2007).

Within higher education, academicians are tasked with providing an enriching learning experience while also engaging in activities that involve grants/scholarships, research, publication, and presentations to further expand their professional development to advance in their career. While faculty are responsible for providing documentation of their involvement and contributions to their department in their dossier for the annual review process, their students are given the responsibility and opportunity to provide further insight of their teaching which also has a significant impact on promotion, tenure, and merit. As research has shown, student evaluations of teaching effectiveness are a common tool many institutions have adopted as a measure to assess quality of the course, instruction, and teacher performance. Given that teaching is the foundation of academicians' career, it could be inferred that a lot of weight is placed on this area in the review process. This could possibly result in lower academic expectations, reduced course requirements/workload, and grade inflation as a result of a fear response and/or pressure placed on faculty to secure employment and advance in their career.

Administrators are recommended to consider the negative impact of this expectation, which could result in an inaccurate reflection of teaching effectiveness.

Despite the ongoing debate regarding the validity and reliability of course evaluations, this measure provides insightful feedback and guidance to improve the educational experience for current and future students (Rowan et al., 2017). This data is pertinent as it provides faculty an opportunity to reflect and improve on areas of growth and restructure the course content, materials, expectations, and requirements. This information could also be shared among colleagues and could possibly benefit faculty who teach similar courses or work within the same academic unit/department/major. It is important to note that course evaluations should not be used as comparative data but instead to evaluate teaching effectiveness based on each faculty member personal experience, patterns, and areas of growth as each will differ from another. Also, using course evaluation as the only means of measuring teaching effectiveness is not recommended. Administrators are encouraged to consider reviewing peer evaluations, teacher portfolios, student achievements, and other data related to faculty engagement and achievements for the annual performance review process. The data collected from this analysis further contributes to a clearer understanding of potential bias that could arise when using course evaluations. Furthermore, it suggests how data from this tool could be considered invalid, unreliable, and unfair if not interpreted accurately and used as an only measure in assessing teaching effectiveness, review for promotion and tenure, and decisions in merit raises.

The results from this study provide new insight into the relationship between the dependent variables (i.e., faculty, student, and course variables) and independent variables (i.e., course evaluations, standard deviation, and response rates). It further builds on the existing literature of variables associated with predicting teaching effectiveness within university course evaluations. This study's findings have shown that faculty gender and low grades can be accounted for 24% of the variance in overall course evaluation and 19.7% of the variance in standard deviation. Findings suggest that gender bias may potentially impact the validity and reliability of course evaluations, which should be considered in employment decisions. Students were found to differ in ratings they provided their female and male professors, possibly due to beliefs they may hold regarding societal gender roles. Their expectations of the behavior and attitudes that their professor should exhibit based on their gender can influence faculty chances of having lower course evaluations than their colleague of the opposite sex (Kogan, Schoenfeld-Tacher, & Hellyer, 2010). For example, female faculty may receive higher or lower ratings when perceived as more understanding, accessible, lenient, supportive, and warm. In contrast, male faculty may receive higher or lower ratings when perceived as professional, competent, stern, and direct. These are variables in which faculty members' have little to no control and could impact their chances of equal opportunity to advance in their academic careers.

Course level and grading patterns are also variables that should be considered during the review of student evaluations of teaching effectiveness as it could affect the

cumulative score of course evaluations. The findings from this study have shown that course level, high grades, and low grades can be accounted for 23.9% of the variance in response rate. These results indicate that students in higher-level courses are more likely to participate in providing feedback than those in lower-level classes, which may be due to their expectations and satisfaction of the course and instructor (Baeten, Kyndt, Struyven, & Dochy, (2010). Findings further indicate that students who receive lower grades are also less likely to complete course evaluations than students who receive higher grades due to dissatisfaction with their grades. To increase response rates, it is recommended that faculty consider discussing the impact and purpose that course evaluations serve for faculty job security, improvement in course content, and overall enhancement of learning. Faculty members may consider using the section about course evaluation on the course syllabus to provide a recent example of how student evaluations have helped improve their course and/or teaching. Moreover, faculty may consider using class-time to complete evaluations, provide consistent reminders, and implementing incentives (Laguilles, Williams, & Saunders, 2011). Administrators may also consider a reasonable percentage of response rates that should be obtained as a valid measure of university course evaluations when reviewing faculty members' dossiers' progress in teaching performance (Stewart, 2018).

Examining the standard deviation of each course evaluation during the review is also suggested, given that the results of this variable provide further evidence of the distribution of response per course. Information from the standard deviation is important

in determining whether the faculty was rated similarly or differently by students regarding their teaching performance and overall experience within the course. This may prompt a need to determine the reason for the difference in ratings and further suggest whether the overall ratings received on course evaluations are a true representation of faculty members' teaching performance. Thus, these results should be considered by faculty and administrators when reviewing course evaluations to improve teaching effectiveness and when used in the decision-making of promotion and tenure and merit raises.

Limitations

A primary limitation of this study is that data were only collected in one department housed in the college of education during one academic year. Though data collection was a representation of this particular university demographic, there was a lack of diversity within the sample size across gender and race of the participants in the department, which may not truly reflect the geographic scope or general population. Student demographics were also limited due to the challenge of obtaining individual student information per faculty/course to protect students' anonymity. Therefore, student data were restricted to the percentage of gender representation and the percentage of race representation in class. Consequently, the generalizability of the research findings is limited due to the small sample size and characteristics of the participants in this study.

The generalizability of the results may also be limited due to the research design. When assessing for correlations and multicollinearity between independent variables, the

variable that is found to correlate with others is removed and label as not significant to the model. Furthermore, independent variables that were explored for this study may not be pertinent to include across other institutions and university populations. It is important to note that data collection from course evaluations differs based on location, university demographic, and academic units/colleges (i.e., business, education, fine arts, science, and mathematics, liberal and applied arts, forestry, and agriculture) and could have a significant impact on variables associated with students' evaluation of teaching effectiveness (Bianchini, Lissoni, & Pezzoni, 2013; Mittal, Gera, & Batra, 2015).

The structure and content of the course evaluation forms used in this analysis is yet another limitation that is likely to vary across universities and academic units. In this study, contents of the evaluation included specific questions related to the following categories/domains: course, instructor, and student items. It is possible that the overall results could have been impacted by subjective responses due to student directed questions regarding their level of interest, expectations, and satisfaction of the course. Though individualized feedback from each student's perspective is welcome and important to obtain, it may overshadow the purpose of the course evaluation. In order to effectively use this instrument to enrich the learning environment, greater value should be placed on course-content and instructor directed questions to improve teaching effectiveness. To improve the meaningfulness of course evaluation, it may be beneficial to include student items as open-ended questions or allow students the opportunity to include written comments instead of including these questions as a Likert-style item for

ratings. This may allow faculty and administrators to review and attend to specific feedback each student provided regarding their experience within that course without it being a reliant factor of teacher effectiveness. Administrators should also consider creating an individualized course evaluation per academic unit. Taking this approach could be more useful as faculty would receive specific type of feedback from students that would be more useful in improving course content, material, instruction, and the overall quality of the learning environment Also, in order to reiterate and explain the purpose and importance of completing course evaluations, these forms may include the prominent uses of teaching evaluation in the promotion, tenure, and merit process.

The course evaluation form used in this study also assess students' evaluation of the course on a five-point rating scale, in which very poor is equivalent to one point and very good is equivalent to five points. This numerical range may have possibly decreased the degree of variability given the weight placed on the mean. Student ratings on course evaluation can often result in skewed distributions which impacts the mean due to the variance of positive and negative ratings. Of note, other Likert scale values could conclude in different results of teaching effectiveness given that these numbers on the rating scales are ordinal and not quantitative. Therefore, the difference in size between each rating is inconsistent and may not hold value or be seen as statistically viable. However, presenting the data differently can be more meaningful when analyzing student ratings per course. Instead, administrators and faculty may benefit from interpreting the distribution of the ratings by reporting the median and mode values while being caution

of the effects the mean rating score has on the overall evaluation. Analyzing the distribution further may provide a better representation of students' perception of the course and instructor. Also, more value should be placed on student commentary that is provided, especially those that are common and representative of student view of their experience within that course.

Future Research and Recommendations

While previous research has focused on analyzing variables separately, the results from this study demonstrate that studying the variables simultaneously provides further understanding of the relationship between each variable and how each predictor contributes to the variance in university course evaluations. Future research in replicating this analysis is needed across various disciplines to determine the relationship between these variables and course evaluations in other universities. Researchers should consider examining variables identified in evaluation surveys (i.e., teaching methods, course materials, course organization) as most important in predicting teaching effectiveness along with faculty, student, and course variables simultaneously. Additional measures of teaching effectiveness and other methods that are used in the process of reviewing faculty's dossier should also be examined to determine how much weight is given to each assessment when determining promotion, tenure, and merit raises.

Faculty and administrators should consider administering mid-semester course evaluations to identify and address areas of concern prior to the end of semester course evaluations. This may offer an opportunity for faculty to have an open dialogue with

students regarding their perspective and demonstrate that their feedback is vital in improving their educational experience. Academicians may also consider developing a course evaluation that is specific in targeting the discipline and context of the course instead of a general course evaluation used across various departments. This could offer a more meaningful measure in gathering data that is applicable and practical in improving the course and teaching effectiveness to enrich student learning within their area of study.

Conclusion

This study contributes to the field by addressing the meaning of course evaluations and the variables that are associated with students' ratings of teaching effectiveness. University course evaluations offer valuable information that can aid in enhancing the learning environment. To take advantage of the benefits of using course evaluation in higher education, faculty and administrators will first have to accurately interpret the findings to determine what information is essential and rule out meaningless variables. Replication in this area is needed to improve university course evaluation as a valid and reliable measure in determining the quality of teaching effectiveness. Furthermore, improving university course evaluations demonstrates continuous efforts in maintaining a tool that will provide fair and equal evaluations for all faculty across disciplines to be considered for promotion, tenure, and merit raises and thrive in their career in higher education.

References

Abrami, P. C. (2001). Improving Judgments About Teaching Effectiveness Using Teacher Rating Forms. *New Directions for Institutional Research, 2001*(109), 59-87. doi:10.1002/ir.4

Adams, M. J. D., & Umbach, P. D. (2012). Nonresponse and Online Student Evaluations of Teaching: Understanding the Influence of Salience, Fatigue, and Academic Environments. *Research in Higher Education, 53*(5), 576-591. doi:10.1007/s11162-011-9240-5

Adibifar, K. (2019). Student evaluations of professors: Does a professors' gender, race, or age influence student ratings?. *International Journal of Active Learning*, *4*(2), 88-98.

Aleamoni, L. M., & Yimer, M. (1973). An investigation of the relationship between colleague rating, student rating, research productivity, and academic rank in rating instructional effectiveness. *Journal of Educational Psychology, 64*(3), 274-277. doi:10.1037/h0034584

Alshare, K. A., Wenger, J., & Miller, D. (2007). The Role of Teaching, Scholarly Activities, and Service on Tenure, Promotion, and Merit Pay Decisions: Deans' Perspectives. *Academy of Educational Leadership Journal*, *11*(1), 53.

Afonso, N. M., Cardozo, L. J., Mascarenhas, O. A., Aranha, A. N., & Shah, C. (2005). Are anonymous evaluations a better assessment of faculty teaching performance? A comparative analysis of open and anonymous evaluation processes. *Fam Med*, *37*(1), 43-7.

Anaya, G., & Cole, D. G. (2001). Latina/o student achievement: Exploring the influence of student-faculty interactions on college grades. *Journal of college student development*, *42*(1), 3-14.

Anderson, J., Brown, G., & Spaeth, S. (2006). Online student evaluations and response rates reconsidered. Innovate: *Journal of Online Education, 2*(6), 1–7.

Anderson, H. M., Cain, J., & Bird, E. (2005). Online Student Course Evaluations: Review of Literature and a Pilot Study. *American Journal of Pharmaceutical Education, 69*(1), 5. doi:10.5688/aj690105

Andrade, E. D. C., & Rocha, B. D. P. (2012). Factors affecting the student evaluation of teaching scores: Evidence from panel data estimation. *Estudos Econômicos (São Paulo)*, *42*(1), 129-150.

Arbuckle, J., & Williams, B. D. (2003). Students' Perceptions of Expressiveness: Age and Gender Effects on Teacher Evaluations. *Sex Roles, 49*(9), 507-516. doi:10.1023/A:1025832707002

Avery, R. J., Bryant, W. K., Mathios, A., Kang, H., & Bell, D. (2006). Electronic course evaluations: Does an online delivery system influence student evaluations? *The Journal of Economic Education, 37*(1), 21-37. doi:10.3200/JECE.37.1.21-37

Bachen, C. M., McLoughlin, M. M., & Garcia, S. S. (1999). Assessing the role of gender in college students' evaluations of faculty. *Communication Education, 48*(3), 193-210. doi:10.1080/03634529909379169

Badri, M. A., Abdulla, M., Kamali, M. A., & Dodeen, H. (2006). Identifying potential biasing variables in student evaluation of teaching in a newly accredited business program in the UAE. *International Journal of Educational Management, 20*(1), 43-59. doi:10.1108/09513540610639585

Baeten, M., Kyndt, E., Struyven, K., & Dochy, F. (2010). Using student-centered learning environments to stimulate deep approaches to learning: Factors encouraging or discouraging their effectiveness. *Educational Research Review*, *5*(3), 243-260.

Bailey, C. D., Gupta, S., & Schrader, R. W. (2000). Do students' judgment models of instructor effectiveness differ by course level, course content, or individual instructor? *Journal of Accounting Education, 18*(1), 15-34. doi:10.1016/S0748-5751(00)00006-3

Barnes, L. L., & Barnes, M. W. (1993). Academic discipline and generalizability of student evaluations of instruction. *Research in Higher Education*, *34*(2), 135-149.

Basow, S. A. (1995). Student Evaluations of College Professors: When Gender Matters. *Journal of Educational Psychology, 87*(4), 656-665. doi:10.1037/0022-0663.87.4.656

Basow, S. A., & Silberg, N. T. (1987). Student evaluations of college professors: Are female and male professors rated differently? *Journal of educational psychology*, *79*(3), 308.

Bedard, K., & Kuhn, P. (2008). Where class size really matters: Class size and student ratings of instructor effectiveness. *Economics of Education Review, 27*(3), 253-265. doi:10.1016/j.econedurev.2006.08.007

Bennett, S. K. (1982). Student perceptions of and expectations for male and female instructors: Evidence relating to the question of gender bias in teaching evaluation. *Journal of Educational Psychology, 74*(2), 170-179. doi:10.1037/0022-0663.74.2.170

Beran, T. N., & Rokosh, J. L. (2009). Instructors' perspectives on the utility of student ratings of instruction. *Instructional Science, 37*(2), 171-184. doi:10.1007/s11251-007-9045-2

Beran, T., & Violato, C. (2005). Ratings of university teacher instruction: How much do student and course characteristics really matter? *Assessment & Evaluation in Higher Education*, *30*(6), 593-601. doi:10.1080/02602930500260688

Bianchini, S., Lissoni, F., & Pezzoni, M. (2013). Instructor characteristics and students' evaluation of teaching effectiveness: Evidence from an Italian engineering school. *European Journal of Engineering Education, 38*(1), 38-57. doi:10.1080/03043797.2012.742868

Blackhart, G. C., Peruche, B. M., DeWall, C. N., & Joiner, T. E. (2006). Factors influencing teaching evaluations in higher education. *Teaching of Psychology*, *33*(1), 37–39.

Boring, A. (2017). Gender biases in student evaluations of teaching. *Journal of Public Economics*, *145*, 27-41.

Boring, A., Ottoboni, K., & Stark, P. (2016). Student evaluations of teaching (mostly) do not measure teaching effectiveness. *Science Open Research*. 1–11. doi:10.14293/S2199-1006.1.SOREDU.AETBZC.v1

Boysen, G. A. (2008). Revenge and Student Evaluations of Teaching. *Teaching of Psychology, 35*(3), 218-222. doi:10.1080/00986280802181533

Brockx, B., Spooren, P., & Mortelmans, D. (2011). Taking the grading leniency story to the edge. The influence of student, teacher, and course characteristics on student evaluations of teaching in higher education. *Educational Assessment, Evaluation and Accountability*, *23*(4), 289-306. doi:10.1007/s11092-011-9126-2

Campbell, H. E. (2019). JPAE at 25: Looking back and moving forward on teaching evaluations. *Journal of public affairs education: J-PAE., 25*(1), 23-29. doi:10.1080/15236803.2018.1558823

Capa-Aydin, Y. (2016). Student evaluation of instruction: comparison between in-class and online methods. *Assessment & Evaluation in Higher Education*, *41*(1), 112-126.

Carlucci, D., Renna, P., Izzo, C., & Schiuma, G. (2019). Assessing teaching performance in higher education: a framework for continuous improvement. *Management Decision, 57*(2), 461-479. doi:10.1108/md-04-2018-0488

Chapman, D. D., & Joines, J. A. (2017). Strategies for Increasing Response Rates for Online End-of-Course Evaluations. *International Journal of Teaching and Learning in Higher Education, 29*(1), 47-60.

Chapman, L., & Ludlow, L. (2010). Can Downsizing College Class Sizes Augment Student Outcomes?: An Investigation of the Effects of Class Size on Student Learning. *The Journal of General Education, 59*(2), 105-123. doi:10.1353/jge.2010.0012

Chaudhary, G., & Rathore, M. S. (2018). A critical study of teacher educators teaching effectiveness in relation to their teaching subjects and teaching experience. *International Journal for Advance Research and Development*, *3*(5), 120-123.

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. doi:10.1080/02602930301683

Chisadza, C., Nicholls, N., & Yitbarek, E. (2019). Race and gender biases in student evaluations of teachers. *Economics Letters*, *179*, 66-71.

Coates, H. (2005). The value of student engagement for higher education quality assurance. *Quality in Higher Education*, *11*(1), 25-36.

Crews, T. B., & Curtis, D. F. (2011). Online Course Evaluations: Faculty Perspective and Strategies for Improved Response Rates. *Assessment & Evaluation in Higher Education, 36*(7), 865-878. doi:10.1080/02602938.2010.493970

Culver, S. (2010). Course Grades, Quality of Student Engagement, and Students' Evaluation of Instructor. *International Journal of Teaching and Learning in Higher Education*, *22*(3), 331-336.

Dommeyer, C. J., Baum, P., Hanna, R. W., & Chapman, K. S. (2004). Gathering faculty teaching evaluations by in-class and online surveys: Their effects on response rates and evaluations. *Assessment & Evaluation in Higher Education, 29*(5), 611-623. doi:10.1080/02602930410001689171

Elmore, P. B., & Pohlmann, J. T. (1978). Effect of teacher, student, and class characteristics on the evaluation of college instructors. *Journal of Educational Psychology, 70*(2), 187-192. doi:10.1037/0022-0663.70.2.187

Emery, C. R., Kramer, T. R., & Tian, R. G. (2003). Return to academic standards: A critique of student evaluations of teaching effectiveness. *Quality Assurance in Education, 11*(1), 37-46. doi:10.1108/09684880310462074

Falkoff, M. (2018). Why we must stop relying on student ratings of teaching. *The Chronicle of Higher Education*, *25*.

Fan, Y., Shepherd, L. J., Slavich, E., Waters, D., Stone, M., Abel, R., & Johnston, E. L. (2019). Gender and cultural bias in student evaluations: Why representation matters. *PloS one*, *14*(2), e0209749. doi:10.1371/journal.pone.0209749

Feistauer, D., & Richter, T. (2017). How reliable are students' evaluations of teaching quality? A variance components approach. *Assessment & Evaluation in Higher Education*, *42*(8), 1263-1279.

Feldman, K. A. (1984). Class Size and College Students' Evaluations of Teachers and Courses: A Closer Look. *Research in Higher Education, 21*(1), 45-116. doi:10.1007/BF00975035

Fernández, J., & Mateo, M. A. (1997). Student and faculty gender in ratings of university teaching quality. *Sex Roles, 37*(11), 997-1003. doi:10.1007/BF02936351

Fjortoft, N. (2015). A Reflection of Faculty and Course Evaluations. *American journal of pharmaceutical education, 79*(9), 129. doi:10.5688/ajpe799129

Franklin, J. (2001). Interpreting the Numbers: Using a Narrative to Help Others Read Student Evaluations of Your Teaching Accurately. *New Directions for Teaching and Learning, 2001*(87), 85-100. doi:10.1002/tl.10001

Freeman, H. R. (1994). Student Evaluations of College Instructors: Effects of Type of Course Taught, Instructor Gender and Gender Role, and Student Gender. *Journal of Educational Psychology, 86*(4), 627-630. doi:10.1037/0022-0663.86.4.627cha

Frey, P. W. (1978). A Two-Dimensional Analysis of Student Ratings of Instruction. *Research in Higher Education, 9*(1), 69-91. doi:10.1007/BF00979187

Gehrt, K., Louie, T. A., & Osland, A. (2015). Student and professor similarity: Exploring the effects of gender and relative age. *Journal of Education for Business*, *90*(1), 1-9.

Ghafourian Boroujerdnia, M., Shakurnia, A. H., & Elhampour, H. (2006). The opinions of academic members of Ahvaz University of Medical Sciences about the effective factors on their evaluation score variations. *Strides in Development of Medical Education*, *3*(1), 19-25.

Goodman, J., Anson, R., & Belcheir, M. (2015). The effect of incentives and other instructor-driven strategies to increase online student evaluation response rates. *Assessment & Evaluation in Higher Education, 40*(7), 958-970. doi:10.1080/02602938.2014.960364

Griffin, B. W. (2004). Grading leniency, grade discrepancy, and student ratings of instruction. *Contemporary Educational Psychology, 29*(4), 410-425. doi:10.1016/j.cedpsych.2003.11.001

Guder, F., & Malliaris, M. (2013). Online course evaluations response rates. *American Journal of Business Education (Online)*, *6*(3), 333.

Hamilton, L. C. (1980). Grades, Class Size, and Faculty Status Predict Teaching Evaluations. *Teaching Sociology, 8*(1), 47-62. doi:10.2307/1317047

Hancock, G. R. (1992). Student and Teacher Gender in Ratings of University Faculty: Results from Five Colleges of Study. *Journal of Personnel Evaluation in Education*, *6*(3), 235-48.

Haslett, B. J. (1976). Student Knowledgeability, Student Sex, Class Size, and Class Level: Their Interactions and Influences on Student Ratings of Instruction. *Research in Higher Education, 5*(1), 39-65. doi:10.1007/BF00991959

Heckert, T. M., Latier, A., Ringwald-Burton, A., & Drazen, C. (2006). Relations among student effort, perceived class difficulty appropriateness, and student evaluations of teaching: is it possible to" buy" better evaluations through lenient grading? *College Student Journal*, *40*(3), 588-597.

Hornstein, H. A. (2017). Student evaluations of teaching are an inadequate assessment tool for evaluating faculty performance. *Cogent Education*, *4*(1), 1304016.

Isely, P., & Singh, H. (2005). Do higher grades lead to favorable student evaluations? *The Journal of Economic Education*, *36*(1), 29-42. doi:10.3200/JECE.36.1.29-42

Jaquett, C. M., VanMaaren, V. G., & Williams, R. L. (2016). The effect of extra-credit incentives on student submission of end-of-course evaluations. *Scholarship of Teaching and Learning in Psychology*, *2*(1), 49-61. doi:10.1037/stl0000052

Johnson, T. D. (2003). Online Student Ratings: Will Students Respond? *New Directions for Teaching and Learning, 2003*(96), 49-59. doi:10.1002/tl.122

Johnson, M. D., Narayanan, A., & Sawaya, W. J. (2013). Effects of course and instructor characteristics on student evaluation of teaching across a college of engineering. *Journal of Engineering Education*, *102*(2), 289-318. doi:10.1002/jee.20013

Joye, S., & Wilson, J. H. (2015). Professor age and gender affect student perceptions and grades. *Journal of the Scholarship of Teaching and Learning*, *15*(4), 126-138.

Kelley, A. C. (1972). Uses and Abuses of Course Evaluations as Measures of Educational Output. *The Journal of Economic Education*, *4*(1), 13-18. doi:10.1080/00220485.1972.10845358

Khong, T. L. (2016). The Validity and Reliability of the Student Evaluation of Teaching: A case in a Private Higher Educational Institution in Malaysia. *International Journal for Innovation Education and Research*, *2*(9), 57-63.

Kidd, R. S., & Latif, D. A. (2004). Student Evaluations: Are They Valid Measures of Course Effectiveness? *American Journal of Pharmaceutical Education, 68*(3), 61. doi:10.5688/aj680361

Kifle, T., & Alauddin, M. (2016). What determines students' perceptions in course evaluation rating in higher education? An econometric exploration. *Economic Analysis and Policy*, *52*(52), 123–130. https://doi.org/10.1016/j.eap.2016.09.004

Kim, J. H. (2019). Multicollinearity and misleading statistical results. *Korean journal of anesthesiology*, *72*(6), 558.

Kogan, L. R., Schoenfeld-Tacher, R., & Hellyer, P. W. (2010). Student evaluations of teaching: Perceptions of faculty based on gender, position, and rank. *Teaching in Higher Education*, *15*(6), 623-636.

Laguilles, J. S., Williams, E. A., & Saunders, D. B. (2011). Can lottery incentives boost web survey response rates? Findings from four experiments. *Research in Higher Education*, *52*(5), 537-553.

Langbein, L. (2008). Management by results: Student evaluation of faculty teaching and the mis-measurement of performance. *Economics of Education Review*, *27*(4), 417-428.

Lewis, V. J., & McKinzie, K. (2019). Impact of industry and teaching experience, course level, and department on student evaluations. *Quarterly Review of Business Disciplines*, *5*(4), 335-356.

Liaw, S., & Goh, K. (2003). Evidence and control of biases in student evaluations of teaching. *International Journal of Educational Management, 17*(1), 37-43. doi:10.1108/09513540310456383

Linse, A. R. (2017). Interpreting and using student ratings data: Guidance for faculty serving as administrators and on evaluation committees. *Studies in Educational Evaluation*, *54*, 94-106.

Love, D. A., & Kotchen, M. J. (2010). Grades, course evaluations, and academic incentives. *Eastern Economic Journal*, *36*(2), 151. doi:10.1057/eej.2009.6

Lundberg, C. A., & Schreiner, L. A. (2004). Quality and Frequency of Faculty-Student Interaction as Predictors of Learning: An Analysis by Student Race/Ethnicity. *Journal of College Student Development, 45*(5), 549-565. doi:10.1353/csd.2004.0061

MacNell, L., Driscoll, A., & Hunt, A. N. (2015). What's in a name: Exposing gender bias in student ratings of teaching. *Innovative Higher Education*, *40*(4), 291-303. doi:10.1007/s10755-014-9313-4

Marsh, H. W. (2001). Distinguishing between Good (Useful) and Bad Workloads on Students' Evaluations of Teaching. *American Educational Research Journal, 38*(1), 183-212. doi:10.3102/00028312038001183

Marsh, H. W., & Bailey, M. (1993). Multidimensional Students' Evaluations of Teaching Effectiveness: A Profile Analysis. *The Journal of Higher Education, 64*(1), 1-18. doi:10.1080/00221546.1993.11778406

Marsh, H. W., Overall, J. U., & Kesler, S. P. (1979). Class Size, Students' Evaluations, and Instructional Effectiveness. *American Educational Research Journal, 16*(1), 57-70. doi:10.2307/1162403

Marsh, H. W., & Roche, L. A. (1997). Making Students' Evaluations of Teaching Effectiveness Effective: The Critical Issues of Validity, Bias, and Utility. *American Psychologist, 52*(11), 1187-1197. doi:10.1037/0003-066X.52.11.1187

Marsh, H. W., & Roche, L. A. (2000). Effects of grading leniency and low workload on students' evaluations of teaching: Popular myth, bias, validity, or innocent bystanders? *Journal of Educational Psychology*, *92*(1), 202.

Martin, L. L. (2016). Gender, Teaching Evaluations, and Professional Success in Political Science. *PS: Political Science & Politics, 49*(2), 313-319. doi:10.1017/S1049096516000275

Mateo, M. A., & Fernandez, J. (1996). Incidence of class size on the evaluation of university teaching quality. *Educational and Psychological Measurement*, *56*(5), 771-778.

McClain, L., Gulbis, A., & Hays, D. (2018). Honesty on student evaluations of teaching: Effectiveness, purpose, and timing matter. *Assessment & Evaluation in Higher Education, 43*(3), 369-385. doi:10.1080/02602938.2017.1350828

McPherson, M. A., Jewell, R. T., & Kim, M. (2009). What determines student evaluation scores? A random effects analysis of undergraduate economics classes. *Eastern economic journal*, *35*(1), 37-51.

Mengel, F., Sauermann, J., & Zölitz, U. (2019). Gender bias in teaching evaluations. *Journal of the European Economic Association*, *17*(2), 535-566.

Mitchell, K. M., & Martin, J. (2018). Gender bias in student evaluations. *PS: Political Science & Politics*, *51*(3), 648-652.

Mittal, S., Gera, R., & Batra, D. K. (2015). Evaluating the validity of Student Evaluation of Teaching Effectiveness (SET) in India. *Education + Training, 57*(6), 623-638. doi:10.1108/ET-06-2013-0072

Mohan, N. (2011). On the use of non-tenure track faculty and the potential effect on classroom content and student evaluation of teaching. *Journal of Financial Education*, *37*(1/2), 29-42.

Morgan, H. K., Purkiss, J. A., Porter, A. C., Lypson, M. L., Santen, S. A., Christner, J. G., . . . Hammoud, M. M. (2016). Student Evaluation of Faculty Physicians: Gender Differences in Teaching Evaluations. *Journal of Women's Health, 25*(5), 453-456. doi:10.1089/jwh.2015.5475

Murray, D., Boothby, C., Zhao, H., Minik, V., Bérubé, N., Larivière, V., & Sugimoto, C. R. (2019). Exploring the personal and professional factors associated with student evaluations of teachers. *PloS one, 15*(6), e0233515. doi:10.1371/journal.pone.0233515

Nasser, F., & Fresko, B. (2002). Faculty Views of Student Evaluation of College Teaching. *Assessment & Evaluation in Higher Education, 27*(2), 187-198. doi:10.1080/02602930220128751

Nerger, J. L., Viney, W., & Riedel, R. G. I. (1997). Student ratings of teaching effectiveness: Use and misuse. *The Midwest Quarterly, 38*(2), 218.

Noel, R. C., & Smith, S. E. (1996). Self-Disclosure of College Students to Faculty: The Influence of Ethnicity. *Journal of College Student Development*, *37*(1), 88-94.

Norris, J., & Conn, C. (2005). Investigating Strategies for Increasing Student Response Rates to Online-Delivered Course Evaluations. *Quarterly Review of Distance Education, 6*(1), 13.

Nulty, D. D. (2008). The adequacy of response rates to online and paper surveys: What can be done? *Assessment & Evaluation in Higher Education, 33*(3), 301-314. doi:10.1080/02602930701293231

Oon, P. T., Spencer, B., & Kam, C. C. S. (2017). Psychometric quality of a student evaluation of teaching survey in higher education. *Assessment & Evaluation in Higher Education*, *42*(5), 788-800.

Park, E., & Dooris, J. (2019). Predicting student evaluations of teaching using decision tree analysis. *Assessment & Evaluation in Higher Education, 45*(5), 1-18. doi:10.1080/02602938.2019.1697798

Peterson, D. A., Biederman, L. A., Andersen, D., Ditonto, T. M., & Roe, K. (2019). Mitigating gender bias in student evaluations of teaching. *PloS one*, *14*(5), e0216241. doi:10.1371/journal.pone.0216241

Podolsky, A., Kini, T., & Darling-Hammond, L. (2019). Does teaching experience increase teacher effectiveness? A review of US research. *Journal of Professional Capital and Community, 4*(4), 286-308. doi:10.1108/jpcc-12-2018-0032

Prince, M. J., Felder, R. M., & Brent, R. (2007). Does faculty research improve undergraduate teaching? An analysis of existing and potential synergies. *Journal of Engineering Education*, *96*(4), 283-294. doi:10.1002/j.2168-9830.2007.tb00939.x

Ray, B., Babb, J., & Wooten, C. A. (2018). Rethinking SETs: Retuning Student Evaluations of Teaching for Student Agency. *Composition Studies*, *46*(1), 34-194.

Rebman Jr, C. M., Wimmer, H., & Booker, Q. E. (2018). A Nationwide Exploratory Study on Faculty Opinions on Student Preparation, Performance, and Evaluations. In *Proceedings of the EDSIG Conference ISSN* (Vol. 2473, p. 3857).

Reid, L. D. (2010). The role of perceived race and gender in the evaluation of college teaching on RateMyProfessors. Com. *Journal of Diversity in higher Education*, *3*(3), 137.

Remedios, R., & Lieberman, D. A. (2008). I liked your course because you taught me well: The influence of grades, workload, expectations and goals on students' evaluations of teaching. *British Educational Research Journal, 34*(1), 91-115. doi:10.1080/01411920701492043

Remedios, R., Lieberman, D. A., & Benton, T. G. (2000). The effects of grades on course enjoyment: Did you get the grade you wanted?. *British Journal of Educational Psychology*, *70*(3), 353-368.

Renaud, R. D., & Murray, H. G. (1996). Aging, personality, and teaching effectiveness in academic psychologists. *Research in Higher Education*, *37*(3), 223-240.

Rezaei, M., Haghdoost, A. A., Okhovati, M., Zolala, F., & Baneshi, M. R. (2016). Longitudinal relationship between academic staffs' evaluation score by students and their characteristics: Does the choice of correlation structure matter?. *Journal of Biostatistics and Epidemiology*, *2*(1), 40-46.

Rivera, L. A., & Tilcsik, A. (2019). Scaling down inequality: Rating scales, gender bias, and the architecture of evaluation. *American Sociological Review*, *84*(2), 248-274.

Rowan, S., Newness, E. J., Tetradis, S., Prasad, J. L., Ko, C., & Sanchez, A. (2017). Should Student Evaluation of Teaching Play a Significant Role in the Formal Assessment of Dental Faculty? Two Viewpoints: Viewpoint 1: Formal Faculty

Assessment Should Include Student Evaluation of Teaching and Viewpoint 2: Student Evaluation of Teaching Should Not Be Part of Formal Faculty Assessment. *Journal of Dental Education, 81*(11), 1362.

Royal, K. D., & Stockdale, M. R. (2015). Are Teacher Course Evaluations Biased against Faculty That Teach Quantitative Methods Courses? *International Journal of Higher Education*, *4*(1), 217-224.

Ryan, J. M., & Harrison, P. D. (1995). The Relationship between Individual Instructional Characteristics and the Overall Assessment of Teaching Effectiveness across Different Instructional Contexts. *Research in Higher Education, 36*(5), 577-594. doi:10.1007/BF02208832

Santhanam, E., & Hicks, O. (2002). Disciplinary, Gender and Course Year Influences on Student Perceptions of Teaching: Explorations and implications. *Teaching in Higher Education, 7*(1), 17-31. doi:10.1080/13562510120100364

Schmelkin, L. P., Spencer, K. J., & Gellman, E. S. (1997). Faculty Perspectives on Course and Teacher Evaluations. *Research in Higher Education, 38*(5), 575-592. doi:10.1023/A:1024996413417

Schwitzer, A. M., Griffin, O. T., Ancis, J. R., & Thomas, C. R. (1999). Social Adjustment Experiences of African American College Students. *Journal of Counseling & Development, 77*(2), 189-197. doi:10.1002/j.1556-6676.1999.tb02439.x

Shapiro, E. G. (1990). Effect of Instructor and Class Characteristics on Students' Class Evaluations. *Research in Higher Education, 31*(2), 135-148. doi:10.1007/BF00992258

Shevlin, M., Banyard, P., Davies, M., & Griffiths, M. (2000). The Validity of Student Evaluation of Teaching in Higher Education: Love me, love my lectures? *Assessment & Evaluation in Higher Education, 25*(4), 397-405. doi:10.1080/713611436

Smith, B. P. (2009). Student ratings of teaching effectiveness for faculty groups based on race and gender. *Education, 129*(4), 615-624.

Smith, B. P., & Hawkins, B. (2011). Examining student evaluations of Black college faculty: Does race matter? *The Journal of Negro Education*, *80*(2), 149-162.

Stewart, T. (2018). Examining the Meaning of Course Evaluation.

Suarez-Balcazar, Y., Orellana-Damacela, L., Portillo, N., Rowan, J. M., & Andrews-Guillen, C. (2003). Experiences of Differential Treatment among College Students of Color. *The Journal of Higher Education, 74*(4), 428-444. doi:10.1080/00221546.2003.11780855

Subbaye, R., & Vithal, R. (2017). Teaching criteria that matter in university academic promotions. *Assessment & Evaluation in Higher Education*, *42*(1), 37-60.

Summers, M. (1996). The camera adds more than pounds: Gender differences in course satisfaction for campus and distance learning students. *Journal of Research and Development in Education*, *29*(4), 212-19.

Tatro, C. N. (1995). Gender effects on student evaluations of faculty. *Journal of Research & Development in Education*.

Thawabieh, A. M. (2017). Students Evaluation of Faculty. *International Education Studies, 10*(2), 35. doi:10.5539/ies.v10n2p35

Theall, M., & Franklin, J. (2001). Looking for Bias in All the Wrong Places: A Search for Truth or a Witch Hunt in Student Ratings of Instruction? *New Directions for Institutional Research, 2001*(109), 45-56. doi:10.1002/ir.3

Thompson, C. G., Kim, R. S., Aloe, A. M., & Becker, B. J. (2017). Extracting the variance inflation factor and other multicollinearity diagnostics from typical regression results. *Basic and Applied Social Psychology*, *39*(2), 81-90.

Thornton, B., Adams, M., & Sepehri, M. (2010). The Impact of Students' Expectations of Grades and Perceptions of Course Difficulty, Workload, and Pace on Faculty Evaluations. *Contemporary Issues in Education Research*, *3*(12), 1-6.

Tran, T. T. T., & Do, T. X. (2020). Student evaluation of teaching: Do teacher age, seniority, gender, and qualification matter? *Educational Studies,* 1-28. doi:10.1080/03055698.2020.1771545

Vatcheva, K. P., Lee, M., McCormick, J. B., & Rahbar, M. H. (2016). Multicollinearity in regression analyses conducted in epidemiologic studies. *Epidemiology (Sunnyvale, Calif.)*, *6*(2), 227.

Wagner, N., Rieger, M., & Voorvelt, K. (2016). Gender, ethnicity and teaching evaluations: Evidence from mixed teaching teams. *Economics of Education Review*, *54*, 79-94.

Wang, L., & Gonzalez, J. A. (2020). Racial/ethnic and national origin bias in SET. *International Journal of Organizational Analysis, 28*(4), 843-855.

Weinkle, L. J., Stratford, J. M., & Lee, L. M. (2020). Voice in Digital Education: The Impact of Instructor's Perceived Age and Gender on Student Learning and Evaluation. *Anatomical Sciences Education*, *13*(1), 59-70.

Williams, W. M., & Ceci, S. J. (1997). "How'm I doing?" Problems with student ratings of instructors and courses. *Change: the Magazine of Higher Learning*, *29*(5), 12-23.

Wood, K., Linsky, A. S., & Straus, M. A. (1974). Class Size and Student Evaluations of Faculty. *The Journal of Higher Education, 45*(7), 524-534. doi:10.1080/00221546.1974.11776994

Young, S., & Duncan, H. E. (2014). Online and face-to-face teaching: How do student ratings differ? *Journal of Online Learning and Teaching*, *10*(1), 70.

Young, K., Joines, J., Standish, T., & Gallagher, V. (2019). Student evaluations of teaching: The impact of faculty procedures on response rates. *Assessment & Evaluation in Higher Education, 44*(1), 37-49. doi:10.1080/02602938.2018.1467878

Young, S., Rush, L., & Shaw, D. (2009). Evaluating Gender Bias in Ratings of University Instructors' Teaching Effectiveness. *International Journal for the Scholarship of Teaching and Learning, 3*(2), . doi:10.20429/ijsotl.2009.030219

Zafar, M., Ghazal, L., Parpio, Y. N., & Amirali, M. (2017). Students' Evaluation on Teachers' Performance. *i-Manager's Journal of Educational Technology*, *14*(2), 34.

Appendix

College of Education Course Evaluation

Course:
Instructor:

Your input is very important in the assessment of both course content and instructor effectiveness. Although your instructor will be able to view the names of students who participate in the survey, the ratings you are about to complete will be anonymous and will not be available to your instructor until after final grades have been posted.

Course/instructor evaluation data will be used for both course and program improvement and for instruction evaluation purposes. Such information has a significant impact on tenure, promotion, pay, and retention decisions. Therefore, please be thoughtful, thorough, and accurate in completing the evaluation. The information from the survey also aids the individual faculty member in course planning and improvement of teaching.

Please make your selections from the following.

	Very Good	Good	Average	Poor	Very Poor	No Opinion
1. The course as a whole was	○	○	○	○	○	○
2. The clarity of course objectives was	○	○	○	○	○	○
3. The usefulness and relevance of course content to course objectives was	○	○	○	○	○	○
4. The organization of the course was	○	○	○	○	○	○
5. The instructor's overall effectiveness in teaching was	○	○	○	○	○	○
6. The variety and appropriateness of learning activities in this course was	○	○	○	○	○	○
7. The instructor's ability to present and explain course content was	○	○	○	○	○	○
8. The instructor's use of examples and illustrations was	○	○	○	○	○	○
9. The instructor's enhancement of students' interest in course content was	○	○	○	○	○	○
10. My confidence in the instructor's knowledge was	○	○	○	○	○	○
11. The instructor's enthusiasm was	○	○	○	○	○	○
12. My interest level in each session (either online or face-to-face) was	○	○	○	○	○	○
13. The availability of extra help when requested was	○	○	○	○	○	○
14. The use of course time (either online or face-to-face) was	○	○	○	○	○	○
15. The instructor's interest in student learning and success was	○	○	○	○	○	○
16. The amount I learned during the course was	○	○	○	○	○	○
17. The evaluation and grading techniques (tests, papers, projects, etc.) were	○	○	○	○	○	○
18. The amount and appropriateness of assigned work was	○	○	○	○	○	○
19. The clarity of my responsibilities and requirements was	○	○	○	○	○	○
20. Additional Comments:						

VITA

Tyesha Stewart graduated from Stephen F. Austin State University with a Bachelor of Arts degree in Psychology and a minor in Child and Family Development in December of 2014. In spring of 2015, Tyesha enrolled at Stephen F. Austin State University and obtained her Master of Arts degree in School Psychology in August of 2018. She then went on to pursue her Doctor of Philosophy degree. The program of School Psychology has provided Tyesha with the opportunity to gain valuable experience assessing children across a wide range of ages, populations, and needs within several school districts. During her doctorate experience, she completed practicums and internships in schools, private practice, and inpatient/outpatient settings within the hospital. These enriching opportunities has provided her a wide range of clinical experiences in a highly diverse mental health population. In the future, Tyesha plans to start a private practice and work toward faculty appointment in higher education settings.

Permanent Address: 1931 Blue Sage Dr.
Lancaster, TX 75146

Publication Manual of the American Psychological Association, Seventh Edition

This dissertation was typed by Tyesha D. Stewart