

Analysis of Factors Affecting Student Evaluation of Teaching Effectiveness in Saudi Higher Education: The Case of Jubail University College

Tayfour Abdalla Mohammed¹, Siraj Muhammed Pandhiani^{2,*}

¹Department of Business Administration, Jubail University College, Jubail, Saudi Arabia

²Department of General Studies, Jubail University College, Saudi Arabia

*Corresponding author: pandhianis@ucj.edu.sa

Abstract The attributes of effective teaching in higher education remains controversial and has never been conclusive. The purpose of this study is to determine the factors affecting the students' perceptions of teaching effectiveness, and how the instructor and course attributes can significantly influence teaching effectiveness as measured by students in course evaluation surveys. The study analyzed 3,798 student evaluations of faculty at Jubail University College using factor analysis to find out the factors loading and average extract variance value. The study predicted that there is a significant relationship between the five dimensions of teaching and students' ratings of teaching effectiveness (i.e. instructor's personality, knowledge, teaching ability, marking and grading policy, and course attributes and learning outcomes). The findings support the hypothesis that there is a significant relationship between effective teaching dimensions and students ratings. The study contributes to the body of literature on evaluation of teaching effectiveness in Saudi higher education.

Keywords: student evaluation of teaching effectiveness, sete, higher education, factor, validity, reliability, jubail university college

Cite This Article: Tayfour Abdalla Mohammed, and Siraj Muhammed Pandhiani, "Analysis of Factors Affecting Student Evaluation of Teaching Effectiveness in Saudi Higher Education: The Case of Jubail University College." *American Journal of Educational Research*, vol. 5, no. 5 (2017): 464-475. doi: 10.12691/education-5-5-2.

1. Introduction

In the recent past years the expansion of higher education sector in the Kingdom of Saudi Arabia has raised the attention of policy makers to issues of educational quality, rather than mass production of graduates, and that there is a need to move the issue to the forefront of the educational agenda. Shevlin [28] asserted that, student evaluation of teaching (SET) is ubiquitous in the UK and the US universities. Moreover, in the UK, information from SET is considered as not only important evaluative information, but also a guide for potential changes in course material and method of delivery, which is clearly noted by the Quality Assurance Agency for Higher Education (QAA) in UK. In Australia, the move towards performance-based funding as part of quality assurance in Australian universities has driven a shift from voluntary to mandatory use of course evaluation surveys to evaluate and reward teaching staff performance [27]. Accordingly, higher education quality assurance agencies in Saudi Arabia such as National Commission for Academic Accreditation and Assessment (NCAAA) are given much autonomy to monitor the quality of higher education and to prescribe the necessary tools to ensure compliance with quality standards.

Considering this, higher education institutions in the Kingdom are mandated to utilize different instruments to examine and evaluate the quality of their educational infrastructure and the quality of teaching and learning [20]. The quest for excellence in college and university teaching is also becoming a worldwide concern and universities pay increasing attention to the quality of the pedagogy practice in their classrooms to assess how effectively professors are teaching [31].

Student evaluation of teaching effectiveness (SETE) has long history in higher education and used by many universities throughout the world as part of their quality assurance [8,22]; yet it has been described by many researchers as one of the most difficult and contentious task, lacking consensus on the attributes of effective teaching, and a cast of doubt concerning the validity and reliability of the measures used so far [17,18,19]. Previous studies in regional universities (Arabian Gulf) suggested that many colleges and universities have adopted the American educational model, hence, have implemented student evaluation of teaching as an instrument for assessing the effectiveness of their course offering and evaluating their academic staff [1]. On the other hand, prior research in the areas has also suggested that many faculty believe that SETE is simply a "popularity contest and have no relation to effective teaching" [19]. In fact, these authors argued that many faculty believe that factors

beyond their direct control, such as the nature of the course and section size can impact student evaluations of faculty. There is a plethora of factor analysis literature which attempted to examine the determinant dimensions of SETE effectiveness [8,13,31], however, the literature revealed a serious lack of consensus and heated debate on the attributes and dimensions that constitute SETE [17].

The purpose of this paper is to contribute to the existing body of empirical literature on student evaluation of teaching effectiveness by empirically investigating a number of dimensions of teaching effectiveness and their ratings by college students in the Saudi higher education context.

The rest of the paper is organized as follows: The next section is the literature review and hypotheses development. This is followed by data and methodology. The analysis and discussion of the results follow this. The final section is the summary and conclusion.

2. Literature Review and Hypothesis

The literature on SETE is substantial, including comprehensive reviews of research on the subject by Al-Issa and Sulieman [1] who found 2988 articles on the subject published in highly ranked scholarly journals between 1990 and 2005. There exist multiple definitions in the literature of effective teaching and the measures that can accurately capture it [17]. Despite its popularity, several researches indicated that SETE should not be considered as the only possible source of information about faculty teaching effectiveness, and *“it is certainly not the best source of that information”* [10]. The arguments between the proponents and opponents are never conclusive. To demonstrate the paradox and inconclusive nature of SETE as a measure of teaching effectiveness [13] argued that; on one side, students regularly interact with their instructors so their opinions and evaluations are valid and relevant; on the other side, if students are *“carelessly and mindlessly”* completing an evaluation form, then SETE is invalid.

The proponents of SETE suggest that it is a source of rich insight on teaching effectiveness and student learning experience, and that it is potentially efficacious and authentic indicator of teaching performance as long as the data generated are used as primary rather than generalized [12]. According to Mary [16] SETE are widely used in tenure and promotion decisions, and in assessing merit. Despite the concern about their legitimacy, however, if properly designed SETEs are both valid and reliable measures to capture the faculty teaching effectiveness. Chulkov and Alstine [5] highlighted that, SETE's proponents utilized the influence of international accrediting bodies such as the Association of Advancement of Collegiate Schools of Business (AACSB) to promote the use of SETE. According to these authors, AACSB states that business schools should use instructional evaluations

as a basis for professional development efforts for individual faculty members and for the faculty in general.

Opponents and critics of SETE argue that, students do not have the knowledge or the experience to evaluate the complex dimensions of teaching [1]. Furthermore, student assessment of teaching quality of their tutors' might be used for punitive purposes such as management control over promotion, professional advancement and dismissal [20]. In another study Moragan [19] found that many faculty believe that students' evaluation has no relation to effective teaching, and that many factors beyond their control, such as the difficulty of course material and section size can impact student evaluation of faculty. Crumbley and Reichelt [8] further asserted that empirical evidence has confirmed that student evaluations are misleading and dysfunctional and that SETE provides incentives for faculty to manipulate their student assessment grading and marking to enhance their SETE scores. For example, this same study by Crumbley and Reichelt [8] found that 53.5% of faculty surveyed knew of other colleagues who have reduced grading standards and course content to improve their SETE scores. To these authors, this represents a “power shift” and a move from professors running universities and colleges to students/administrators controlling higher education; where students simply “punishing difficult professors on their evaluation forms” (ibid: p.382). According to Tanner [30] the current trends in SET have greatly contributed to undermining the teaching profession. Smithson John [29] argued that course evaluation surveys completed by students touch on only few aspects of course and teaching quality.

Despite of the substantial body of literature around SETE, there is a little agreement as to which factors to be used to capture and measure the valid and reliable evaluations. Shevlin [28] asserted that *“despite the perceived importance of SET, there appears to be little agreement on the nature and number of dimensions/factors that represent teaching effectiveness”* (p.398). The following Table 1 summarized some of the factors employed by authors in education research in an attempt measure the attributes of effective teaching.

Despite the wide range of SET factor analysis literature and empirical studies [28] questioned, whether we really measuring the relevant variables of teaching effectiveness, and how far these measures are valid and reliable in making critical personnel decisions. Marsh [14] suggested that, student evaluations designed to measure and reflect effective teaching, should be “multidimensional/multifaceted” using multiple factors.

Moreover, d'Apollonia and Abrami [9] explained that, majority of the studies in the area focused on the relationship between ratings in SETE surveys and variable related to student characteristics, lecturer behavior and attributes, and course attributes [28]. Findings from the literature on SET factors and variables analysis revealed diverse and contradictory results.

Table 1. Summary of SETE Factors Considered in Previous Studies

Study	SETE factors considered.
Shevlin et al (2000)	Lecturer ability, module attributes and lecturer charisma
Jepsen et al (2015)	Learning styles
Smithson et al (2015)	Learning outcomes, delivery, organization, assessment, feedback and overall satisfaction.

Study	SETE factors considered.
Schumacher et al (2015)	Managing classrooms, organizing for instructions, implementing instructions, monitoring student progress, and potentials.
Al-Issa and Suleiman, (2007)	Student's origin, gender, language, school/college, GPA, and academic standing.
Van der Merwe (2015)	Lecturer knowledge, lecturer personality and lecturer attitude
Emery et al (2003)	Punctual in meeting class, available outside class, grading fair and reasonably, prepared for class, and knowledge of the subject.
Chulkov and Alstine, (2012)	Knowledge of the subject, enthusiasm about subject, clarity of explanations, attitude towards students, opportunity for questions, organization of topics, organization of lectures, textbooks.
Ngware & Ndirangu, (2005)	Presentation/teaching style, knowledge of subject matter and personal attributes.
Mittal et al (2015)	Lecturer ability, module attributes and lecturer charisma
Liaw and Goh (2003)	Lecture organization, knowledge of subject, presentation/teaching style, clarity in delivering the lecture, and appropriateness of examples and illustrations
Badri et al (2006)	Instructor's knowledge and performance in teaching and assessment, instructor's role in exams, grading and in relationship with students, overview of the course, course requirements and efforts, course outcomes
Otani et al (2012)	Class preparation, clear explanation, enthusiasm, stimulating course materials, instructor availability, student understanding of expectations, comprehensive exam, fair assessment, course objectives, positive learning environment, encouragement in participation, effective use of class time, gaining new knowledge.

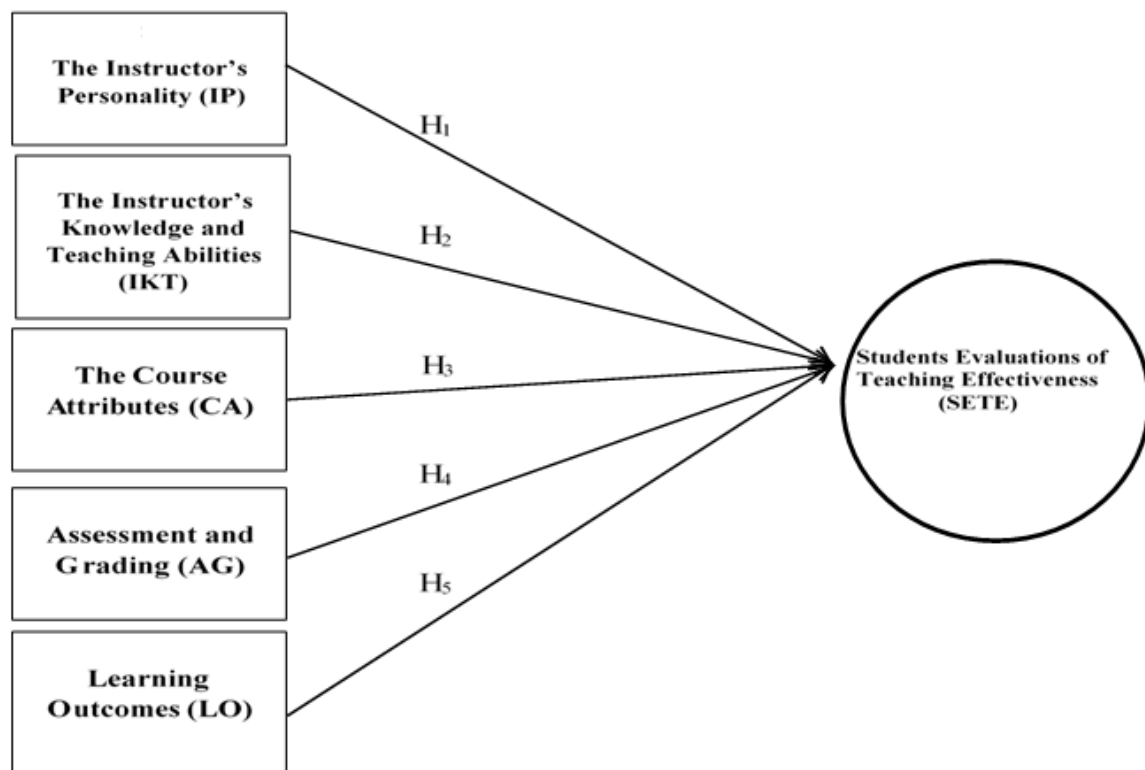


Figure 1. Research Hypothesis

From these various literature discussed above, this study was devised to examine the relationship between SETE ratings and the instructor's personality traits, his/her behavior in marking and grading, his/her knowledge and teaching ability, the course attributes and the course learning outcomes. It is predicated that:

Hypothesis:

- H1. There is a significant relationship between instructor's personality traits and SETE ratings.
- H2. There is a significant relationship between students' perceived fairness in assessments marking, grading and workload and SETE ratings.
- H3. There is a significant relationship between instructor's knowledge and teaching abilities and SETE ratings.
- H4. There is a significant relationship between the course attributes and SETE ratings.
- H5. There is a significant relationship between students' perception of learning outcomes and SETE ratings.

3. Data and Methodology

To examine the factors that affect Jubail University College (JUC) student evaluation of teaching effectiveness, data from course evaluation survey (CES) were analyzed (see appendix 1). JUC is a small university college which has two separate campuses for male and female students. In the academic year 2016/2017, there were 3669 students enrolled in the JUC seven academic programs that offer bachelor degrees in business administration, MIS, computer science, interior design, English language, mechanical engineering and civil engineering.

For our study we used data obtained from CES conducted during the first and second semesters of the academic year 2015/2016. The survey consisted of 26 specific question divided into five sections in which students evaluate their faculty performance at the start of the course, during the course, after the course, and overall

rating with the last section (3 questions) giving opportunity for comments. The responses in 23 questions were measured on a five-point scale, ranging from 1 (strongly disagree) to 5 (strongly agree). The survey CES questions were re-grouped under five factors that predicted to significantly influence the rating of SETE, namely: the *instructor's personality traits, his/her behavior in marking and grading, his/her knowledge and teaching ability, the course attributes and the course learning outcomes*. (see appendix 2)

The survey is administered electronically where all enrolled full-time undergraduate students were mandatory required to complete the survey in a pre-designated computer labs. During the academic year 2015/2016 there were 2267 students enrolled (*preparatory year students are excluded*) in full-time courses during the first semester of which, 1331 were female student and 936 were male. 2000 students completed the survey with 88% response rate. The number of female respondents is 1160 giving a response rate of 87%, while 840 male students completed the survey with a response rate of 90%. During the second semester there were 2080 students enrolled, and 1798 responded (86%). 1221 were female student out of which 999 responded (82%) and 859 male students; around 799 responded (93%).

The study measured 23 attributes of teaching effectiveness which were administered to students electronically. The 23 attributes were designed to measure five dimensions of teaching effectiveness. Four attributes related to the instructor's personality measuring traits such as punctuality, commitment, enthusiasm and caring. Four attributes related to the instructor's knowledge and teaching ability. Six attributes related to course characteristics, four attributes related to the students' perception of the instructor's behavior in assessments marking and grading, and four attributes related to students' perception of the course learning outcomes. With the last items being the overall rating of the student satisfaction with the teaching effectiveness considered as a dependent variable measured by calculating the overall mean the scores in the above mentioned factors. Responses to the items were made on a 5-point Likert scale anchored with (strongly agree =5), (agree=4), (neutral=3), (disagree=2), and (strongly disagree =1).

To measure the reliability of the CES questionnaire we used Cronbach's alpha. According to Chen [4], Cronbach's alpha is one of the most popular methods to assess reliability. It is useful when the measuring tool has many similar questions or items to which the participant can respond [7]. The value of alpha measures the internal consistency or homogeneity between the items. According to Raykov [24], high correlations among the items are associated with high alpha value [4]. A value of alpha ranges from 0 to 1 and the higher the alpha the higher the reliability.

The reliability was tested using SPSS software and

values were calculated for each of the multi-item factors, which were included in our model. All the items/questions in our CES showed high Cronbach's alpha values, which indicated that our measurements were reliable. Table 2 presents Cronbach's alpha, factor loading and average extract variance values for each construct in the research model. The lower accepted value of Cronbach's alpha is 0.7 [21] even though Robinson [25] argued that a value of 0.6 could be accepted. All Cronbach's alpha values listed in Table (2) are greater than 0.7.

Although Cronbach's alpha is a widely used technique and indicator of reliability, it has been criticized for underestimation [3,6,23,24]. The problem of underestimation results from the underlying assumption for Cronbach's alpha where, all measured items are equally weighted or the path coefficients from the latent factor to the measured items are assumed to be equal. Failure to meet this assumption results in Cronbach's alpha underestimating reliability. Alternatively, Wertz [32] developed the composite reliability in assessing the reliability for a set of indicators. The composite reliability relaxes the assumption in Cronbach's alpha assessment and it is "a closer approximation under the assumption that the parameter estimates are accurate" [6], and thus it has been considered a more accurate measurement than Cronbach's alpha [11]. Table 2 also shows the reliability assessment using the composite reliability. The accepted value for composite reliability is 0.70 or higher [21].

To test our hypothesis we used SPSS software to run the t-test for the data to calculate the t-value/s and p-value/s. If the calculated p-value is < 0.05 the null hypothesis is rejected, otherwise it is accepted.

4. Results Analysis & Discussion

Figure 1 hypothesized a five-factor measurement model for the 22 questions in the CES questionnaire measuring students' evaluation of teaching effectiveness. The instructor's personality (IP) is measured by respective items IP1 to IP4, instructor's knowledge and teaching abilities (IKT) is measured by items IKT1 to IKT4, the course attributes (CA) is measured by items CA1 to CA6, the instructor's assessment and grading behavior (AG) is measured by items AG1 to AG4, and the course learning outcomes is measured by items LO1 to LO4. The reliability was specified at Cronbach's alpha ≥ 0.7 and for average variance extract it is ≥ 0.6 of the items in each factor.

Table 2 indicates that, in the male branch the factor loading values range between 0.911 (IKT) and 0.961 (IP & LO) and variance extract values range between 0.788 (CA) and 0.895 (IP). In the female branch, on the other hand, the factor loading values range between 0.914 (AG) and 0.978 (IP) and variance extract values range between 0.866 (AG) and 0.936 (IP).

Table 2. Factor Loading, Average Extract Variance (AEV) & Cronbach's alpha Values

	Factor loading Values	AEV Values	Cronbach's alpha
Male Branch	0.911 - 0.961	0.788 - 0.895	0.913 – 0.964
Female Branch	0.914 - 0.978	0.866 - 0.936	0.946 – 0.984
Significant value >	0.5	0.6	0.7

The Factor loadings indicate that the questions used in the CES for student evaluation of teaching effectiveness are good indicators and explanatory of the instructor's personality, his/her knowledge and teaching abilities, the course attributes, the instructor's assessment and grading behavior, and students' perceptions of the course learning outcomes. All the factor loadings and average variance extract are positive, relatively high (> 0.70) therefore, statistically significant.

The following Figure 2 & Figure 3 summarize the p-values for male and female branches at ($\alpha = 0.05$) for all

items used by end of second semester. The figures below indicate that the p-values are less than α (0.05), therefore, our study reject all the null hypothesis in favor of our alternate hypothesis stated earlier. This result indicates that there is a significant relationship between SETE and each of the five dimensions used in the evaluation tool.

However, when looking at the *p-value* test from a very close glance, we will find some variations. For example, the analysis of the data on departmental basis revealed that in particular departments some of the null hypothesis are accepted.

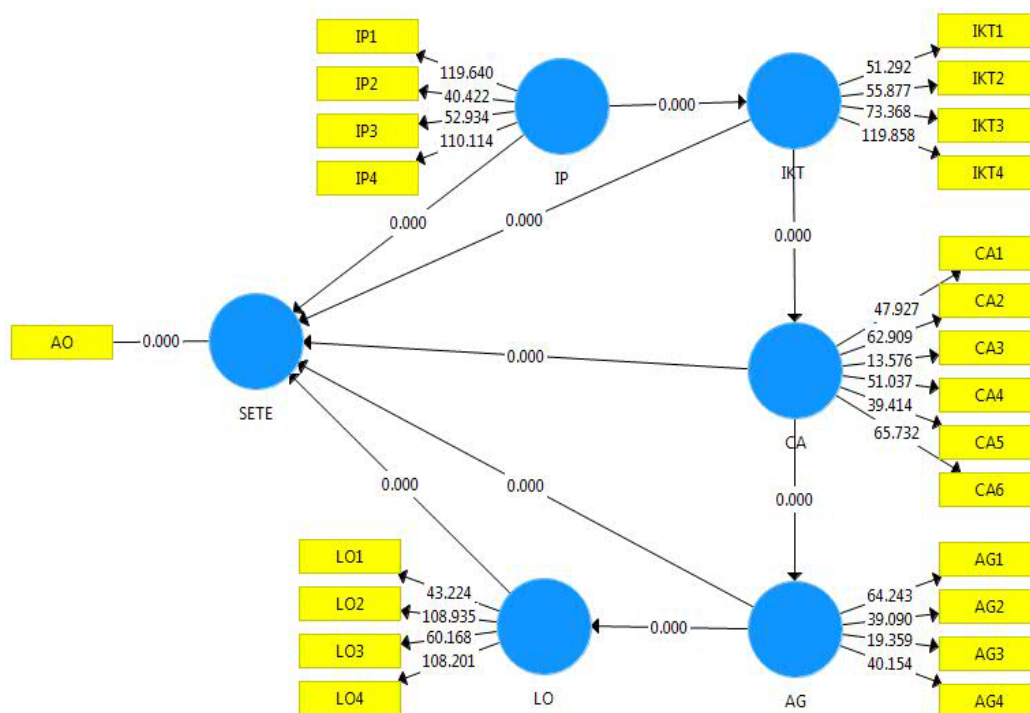


Figure 2. Male Branch P-Values

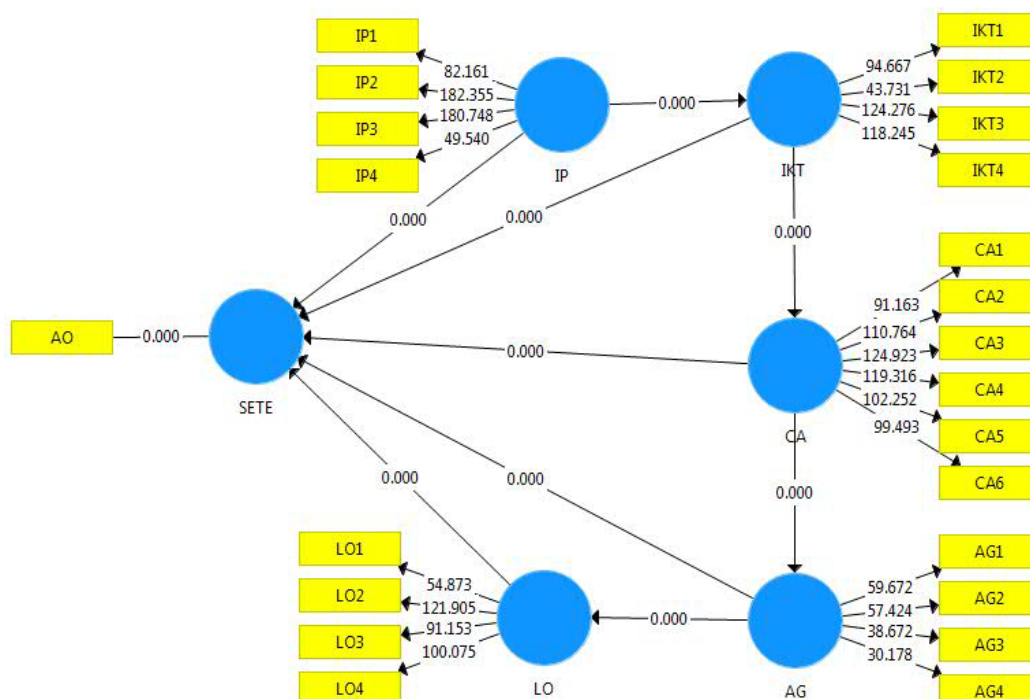


Figure 3. Female Branch P-Values

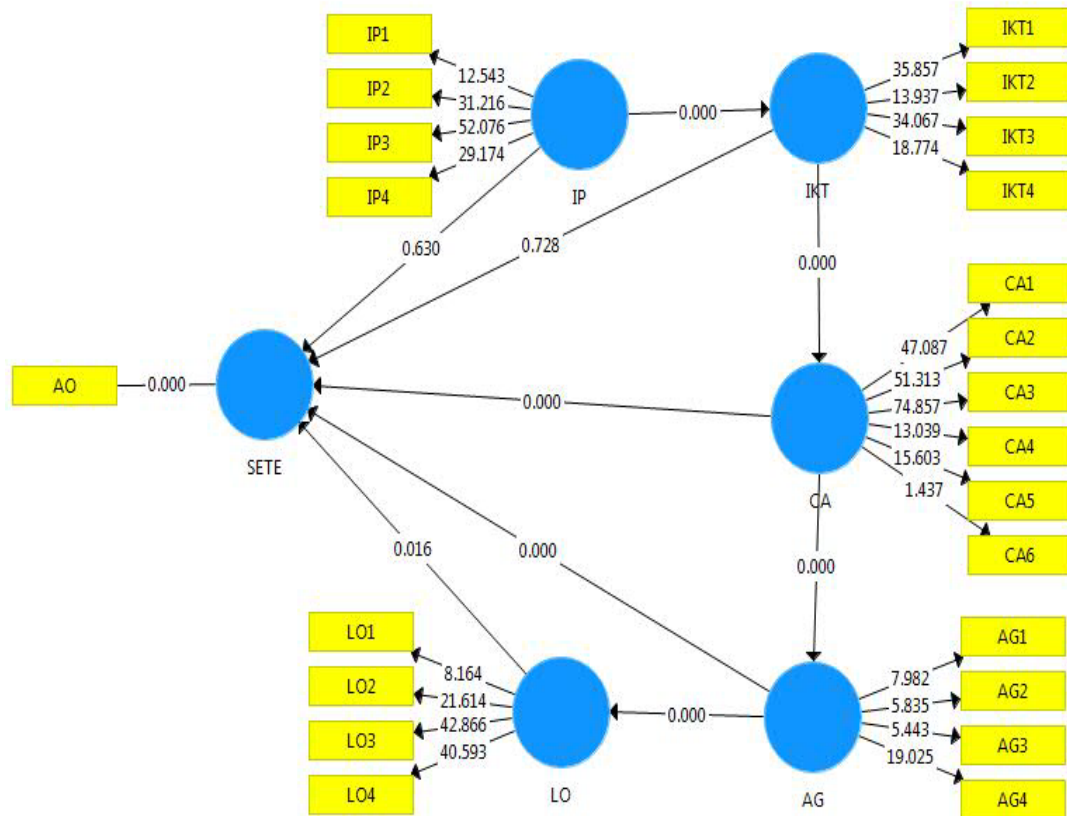


Figure 4. T-value & P-value for students in Mechanical Engineering Discipline

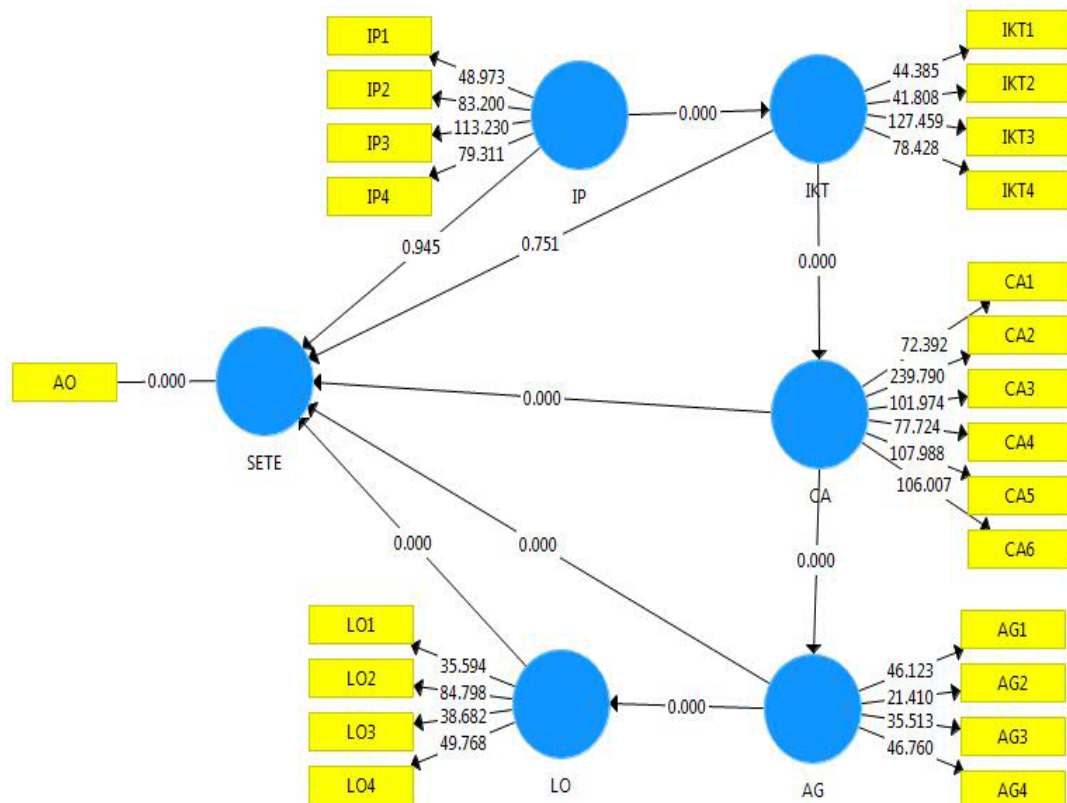


Figure 5. T-value & P-value for students in Civil Engineering Discipline

For example, the students in the Engineering discipline (Mechanical & Civil) in the male branch both rejected the alternate hypothesis H1 & H2 (see Figure 4 & Figure 5).

The P-values are IP = 0.630, IKT = 0.728 for Mechanical engineering and IP = 0.945, IKT = 0.751 for

Civil; and all are greater than α (0.05). The above results indicate that Mechanical and Civil engineering students, do not consider the instructor's personality or his knowledge of subject and teaching ability as determinant factors when they rate their instructors in the CES at the

end of each semester. In other words, they do care about assessments and grading, course characteristics and whether they achieved the course learning outcomes.

The same result revealed in the t-test results for the Interior Department in the female branch [Figure 6](#). The

P-values for the ID department are IP = 0.778, IKT = 0.999 and CA = 0.068 and the rest are null. The ID department findings reject the alternate hypothesis in three factors, therefore, they did not see any significant relationship between students' ratings in CES and the above mentioned factors.

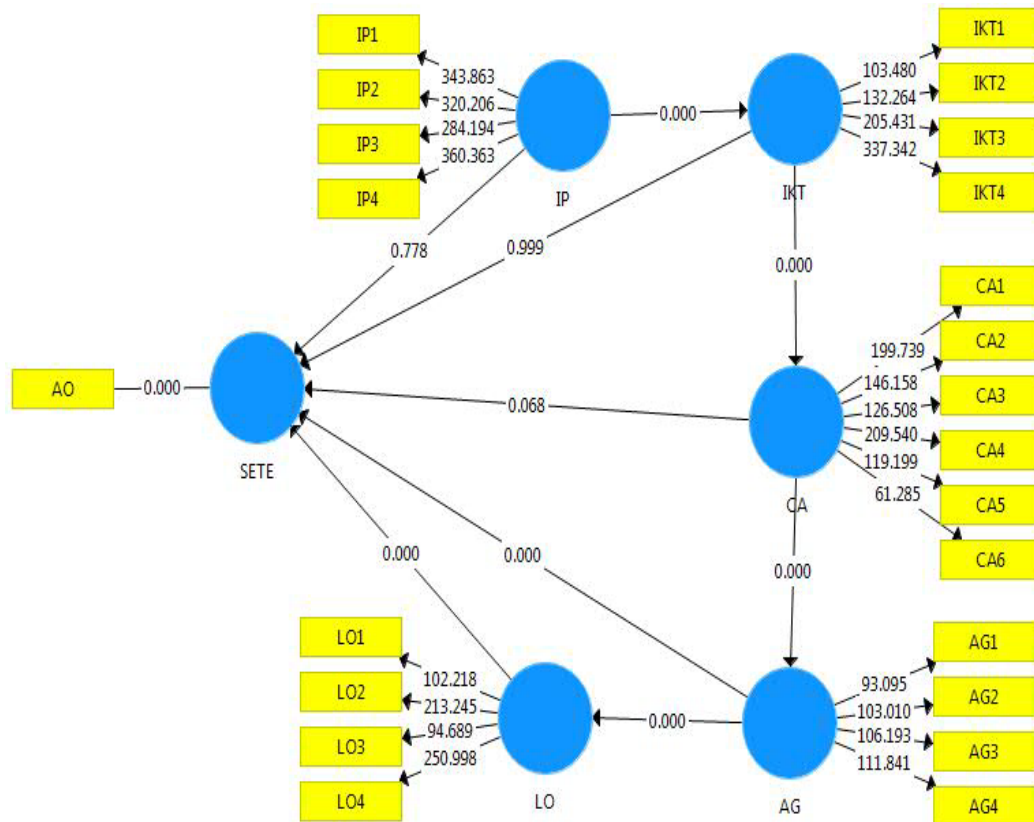


Figure 6. T-value & P-value for students in ID Department

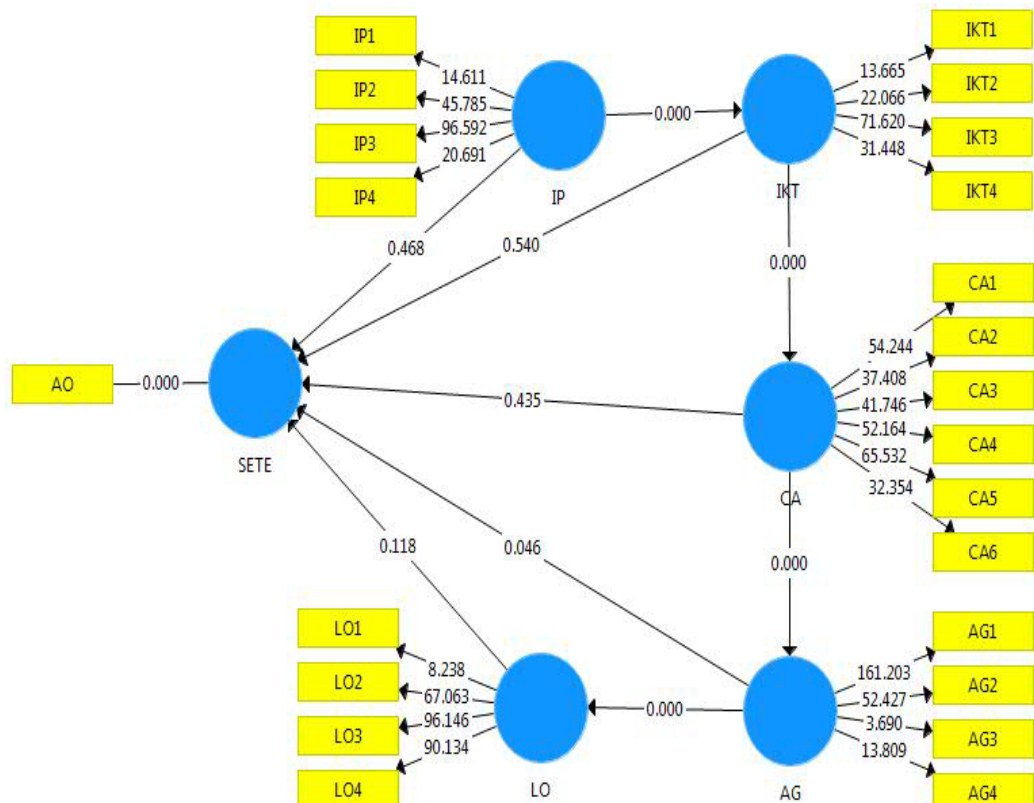


Figure 7. T-value & P-value for students in MIS Program

When looking at the P-value in the MIS program in the business department at the female branch [Figure 7](#) the result is slightly different. The P-values are IP = 0.468, IKT = 0.540, CA = 0.435, AG = 0.046, and LO = 0.118 all are greater than α (0.05) except AG = 0.046. The above results indicate that MIS students evaluate their instructors' teaching effectiveness on the basis of their perception of instructors' behavior in assessment and grading, and, give no consideration to the other dimensions of teaching effectiveness. While their counterparts in the MIS male branch [Figure 8](#) consider all dimensions of teaching effectiveness except the IP which indicates no significant relationship with students rating of teaching effectiveness ($P\text{-value} = 0.490$).

The descriptive statistical data presented in [Table 3](#) gives an overview of the mean scores of all the dimensions/factors of teaching effectiveness as ranked by JUC students.

The statistics in the table above show that the instructors were rated high on their personality attributes by both male and female students with an overall mean of

4.120 and 3.913 respectively. This is followed by course attributes (mean 4.071 & 3.853), then knowledge and teaching ability which has a mean of (4.015 & 3.829) for male and female respectively. Followed by perception of learning outcomes (3.986 & 3.826) and finally the instructor's assessment and grading behavior (mean 3.985 & 3.781).

Thus, it can be concluded that the students are most impressed with the instructors personality, course attributes, instructors knowledge and teaching abilities, course learning outcomes and assessment and grading respectively.

The descriptive statistics relating of the attributes of teaching effectiveness under each factor /dimension are presented in [Table 4](#). In the male branch, the statistics show that the instructor **commitment**, was rated highest with an overall mean of 4.202, the attribute related to the '**usefulness**' of the course learning outcomes was rated second highest with an overall mean of 4.162 and the attribute "*this course helped me to improve my skills in working as a member of a team*" was the lowest rated with an overall mean of 3.886.

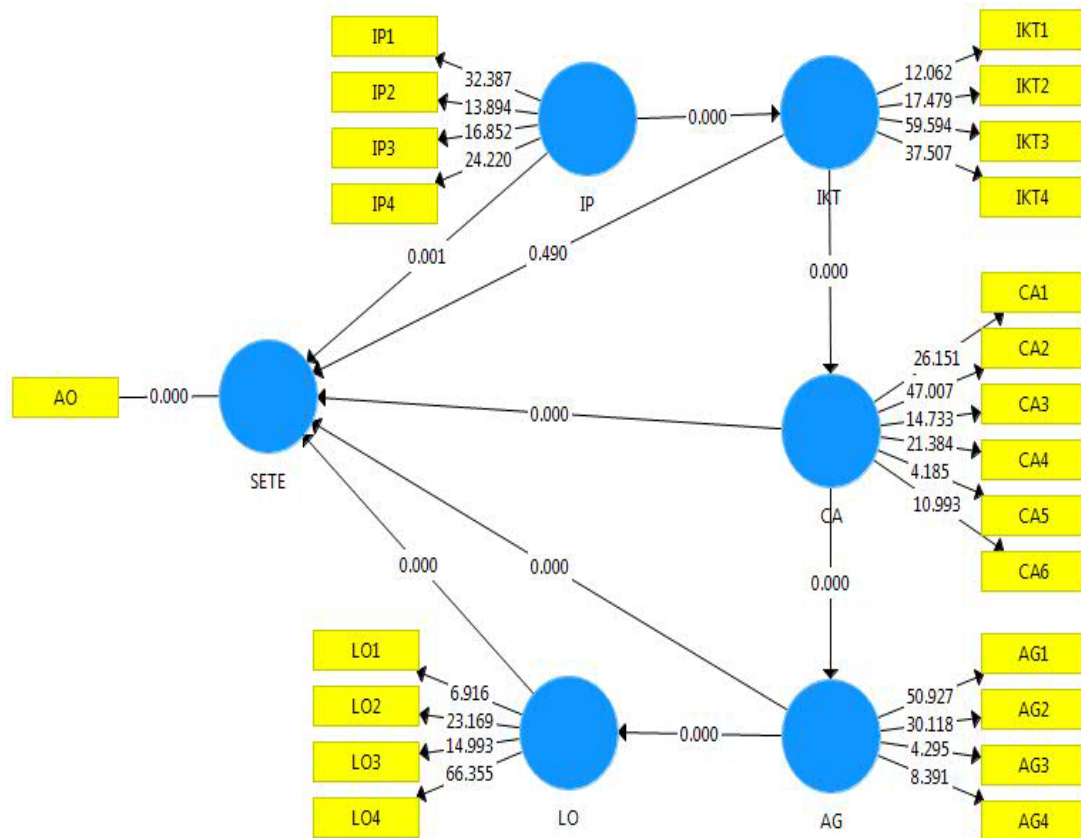


Figure 8. T-value & P-value for students in MIS Program

Table 3. Descriptive Statistics of the Dimensions of SETE

Attributes	Overall Average Male	Rank	Overall Average Female	Rank	Overall Average Male/Female	Rank
The Instructor's Personality (IP)	4.120	1	3.913	1	4.016	1
The Instructor's Knowledge and Teaching Abilities (IKT)	4.015	3	3.829	3	3.922	3
The Course Attributes (CA)	4.071	2	3.853	2	3.962	2
Assessment and Grading (AG)	3.985	5	3.781	5	3.883	5
Learning Outcomes (LO)	3.986	4	3.826	4	3.906	4

Table 4. Descriptive Statistics of the Attributes of SETE and Overall Ranking

Attributes	Variables	Overall mean Male	Rank Male	Overall mean Female	Rank Female	Overall mean M/F	Overall Rank
The Instructor's Personality (IP)	My instructor(s) were fully committed to the delivery of the course.	4.202	1	3.933	3	4.068	1
	My instructor(s) were available during office hours to help me.	4.139	4	3.929	4	4.034	4
	My instructor(s) were enthusiastic about what they were teaching.	4.112	7	3.940	2	4.026	6
	My instructor(s) cared about my progress and were helpful to me.	4.028	10	3.850	10	3.939	10
The Instructor's Knowledge and Teaching Abilities (IKT)	My instructor(s) had thorough knowledge of the content of the course.	4.130	6	3.964	1	4.047	2
	In this course effective use was made of technology to support my learning.	4.005	14	3.784	19	3.895	15
	In this course I was encouraged to ask questions and develop my own ideas.	4.000	15	3.801	13	3.901	14
	In this course I was inspired to do my best work.	3.927	20	3.769	21	3.848	20
The Course Attributes (CA)	The course outline (including the knowledge and skills the course was designed to develop) was made clear to me.	4.147	3	3.911	6	4.029	5
	The things I had to do to succeed in the course, including assessment tasks and criteria for assessment, were made clear to me.	4.139	5	3.863	8	4.001	7
	Sources of help for me during the course including faculty office hours and reference material, were made clear to me.	4.099	8	3.860	9	3.980	8
	The conduct of the course and the things I was asked to do were consistent with the course outline.	4.083	9	3.869	7	3.976	9
	Course materials were of up to date and useful. (texts, handouts, references etc.).	4.012	13	3.821	12	3.917	12
	The resources I needed in this course (textbooks, library, computers etc.) were available when I needed them.	3.947	18	3.796	16	3.872	17
Assessment and Grading (AG)	The things I had to do in this course (class activities, assignments, laboratories etc) were helpful for developing the knowledge and skills the course was intended to teach.	3.923	21	3.771	20	3.847	21
	The amount of work I had to do in this course was reasonable for the credit hours allocated.	4.023	11	3.787	18	3.905	13
	Marks for assignments and tests in this course were given to me within reasonable time.	4.014	12	3.849	11	3.932	11
	Grading of my tests and assignments in this course was fair and reasonable.	3.981	16	3.717	22	3.849	19
Learning Outcomes (LO)	What I learned in this course is important and will be useful to me.	4.162	2	3.915	5	4.039	3
	This course helped me to improve my ability to think and solve problems rather than just memorize information.	3.967	17	3.796	15	3.882	16
	This course helped me to develop my skills in working as a member of a team.	3.886	22	3.795	17	3.841	22
	This course improved my ability to communicate effectively.	3.928	19	3.798	14	3.863	18

In the female branch on the other hand, the attribute “my instructor had thorough knowledge of the content of the course” was rated highest with an overall mean of 3.964, the attribute “my instructor/s were enthusiastic about what they were teaching” was rated second highest with an overall mean of 3.940, and the attribute related to “the grading of tests and assignments in this course was fair and reasonable” was the lowest rated with an overall mean of 3.717.

The results indicate that the SETE ratings were significantly influenced by the students’ perception of

instructions dimensions considered in this study. The five-factor analysis with high factor loading, average extract variance and Cronbach’s alpha values, suggest that meaningful and reliable factors related to teaching effectiveness were being measured. The relatively close values of the factors measured, which seems surprising to many instructors who always predict that their assessment and grading policy is by large the major determinant of their ratings in SETE surveys. However, this explanation coincides with the major findings in the literature as concluded by Shevlin [26] that “good and effective

teaching is not a one-dimensional skill, teaching is shown to be multi-dimensional" (p. 403).

5. Summary and Conclusion

The purpose of this study was to determine the factors affecting students' perceptions of teaching effectiveness and whether the instructor and course attributes have significant influence on effectiveness of teaching as measured by students' course evaluation survey. The results driven from the study demonstrate that CES is a useful and reliable tool for evaluating teaching effectiveness in Jubail University College. The findings suggest that instructor's personality, knowledge and teaching ability as well as course attributes, learning outcomes and the grading behavior have a significant relationship with effectiveness of teaching as perceived by students. Among the factors, instructor's personality was found to be the most important factor in the teaching evaluations for both male and female students.

The research design and data analysis is limited to one academic year, the validity and reliability of the results could be enhanced if the data analysis is extended to three years back. Moreover, this study has not considered the way students completed the online survey and the level of training they received in filling the questionnaire.

Despite the limitations, the results of the study should be useful to JUC instructors, academic departments and college administration to better understand students' course-instructor requirements. For example, course learning outcomes were generally considered low priority by both male and female students. This result should provide useful information for both instructors and academic departments for better management of course syllabus and content. For JUC faculty it is evident the need to highlight the importance of course learning outcomes as they introduce their courses or specific topics. Academic departments should demonstrate to students that the results of CES are integral part of measuring their learning outcomes. The results of this study could also contribute to the growing body of theoretical and empirical literature on students' perception of teaching effectiveness in Saudi higher education.

Acknowledgements

The authors would like to extend their sincere appreciation to the Managing Director of Jubail University College who facilitated the access to the data. We would also like to thank and acknowledge the Research and Applied Studies Center for their support during the research process.

References

- [1] Al-Issa, Ahmad and Sulieman Hana, (2007), "Student evaluations of teaching: perceptions and biasing factors", *Quality Assurance in Education*, Vol. 15 (3), pp. 302-317.
- [2] Badri, Masood A., Mohamed Abdulla, Mohammed A. Kamali, and Hamzeh Dodeen, (2006), "Identifying potential biasing variables in student evaluation of teaching in a newly accredited business program in the UAE", *International Journal of Educational Management*, Vol. 20 (1), pp. 43-59.
- [3] Bollen, K.A., & P.J. Curran (2005). *Latent curve models: A structural equation perspective*. Hoboken, NJ: Wiley.
- [4] Chen, G. M. (2009). Intercultural effectiveness. In L. A. Samovar, R. E. Porter, & E. R. McDaniel (Eds.), *Intercultural communication: A reader* (pp. 393-401). Boston, MA: Wadsworth.
- [5] Chulkov, Dmitriy V. and Alstine, Jason Van, (2012). "Challenges in designing student teaching evaluations in a business program", *International Journal of Educational Management*, Vol. 26 (2) pp. 162-174.
- [6] Chin, W. W. 1998. "Issues and Opinion on Structural Equation Modeling," *MIS Quarterly* Vol. 22 (1), pp. 7-16.
- [7] Cooper, D. R., & Schindler, P. S. (2008). *Business Research Methods* (12th ed.). Retrieved from The University of Phoenix eBook Collection database.
- [8] Crumbley, D. Larry and Reichelt, Kenneth J. (2009). "Teaching effectiveness, impression management, and dysfunctional behavior Student evaluation of teaching control data", *Quality Assurance in Education*, Vol. 17 (4), pp. 377-392.
- [9] d'Apollonia, S., & Abrami, P. C. (1997). Navigating student ratings of instruction. *American Psychologist*, 52, 1198-1208.
- [10] Emery, C. R., Kramer, Tracy R. and Tian, Robert G. (2003). "Return to academic standards: a critique of student evaluations of teaching effectiveness", *Quality Assurance in Education*, Vol. 11 (1), pp. 37-46.
- [11] Fornell, C., Larcker, D.F., 1981. Evaluating structural equation models with unobservable variables and measurement error, *Journal of Marketing Research* Vol.18 (1), pp. 39-50.
- [12] Kinash, Shelley, Vishen Naidu, Diana Knight, Madelaine-Marie Judd, Chenicheri Sid Nair, Sara Booth, Julie Fleming, Elizabeth Santhanam, Beatrice Tucker, Marian Tulloch, (2015). "Student feedback: a Learning and teaching performance indicator", *Quality Assurance in Education*, Vol. 23 (4), pp. 410-428.
- [13] Lakin, Amy L. (2016), "Effective faculty evaluation at the teaching-centered university Building a fair and authentic portfolio of faculty work", *International Journal of Educational Management*, Vol. 30 (6), pp. 976-988.
- [14] Liaw, Shu-Hui and Goh, Kim-Leng (2003), "Evidence and control of biases in student evaluations of teaching", *International Journal of Educational Management*, Vol. 17 (1), pp. 37-43.
- [15] Marsh, H. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research*, II(Who1e Issue No. 3).
- [16] Mary Kelly (2012). "Student Evaluation of Teaching Effectiveness: Considerations for Ontario University", *COU Academic Discussion Paper*, pp. 1-15.
- [17] Mittal, S. Gera, R. and Batra, D. Kumar, (2015), "Evaluating the validity of Student Evaluation of Teaching Effectiveness (SET) in India", *Education & Training*, Vol. 57 (6) pp. 623-638.
- [18] Mittal, Sanjiv and Gera, Rajat (2013), "Student Evaluation of Teaching Effectiveness (SET): An SEM Study in Higher Education in India", *International Journal of Business and Social Science*, Vol. 4 (6) pp. 289-298
- [19] Morgan, Donald Ace, Sneed John, and Swinney, Laurie, (2003). "Are student evaluations a valid measure of teaching effectiveness: perceptions of accounting faculty members and administrators", *Management Research News*, Vol. 26 (7). pp. 17-32.
- [20] Ngware, M. W. and Ndirangu, M., (2005). "An improvement in instructional quality: can evaluation of teaching effectiveness make a difference?", *Quality Assurance in Education*, Vol. 13(3) pp. 183-201.
- [21] Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill, Inc.
- [22] Otani, Koichiro, B. Joon Kim, and Jeong-IL Cho. (2012) "Student Evaluation of Teaching (SET) in Higher Education: How to Use SET More Effectively and Efficiently in Public Affairs Education." *Journal of Public Affairs Education*, Vol.18, (3), PP. 531-44.
- [23] Raykov T (1997a). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement* 21(2): 173-184.
- [24] Raykov T (1997b) Scale reliability, Cronbach's coefficient alpha, and violations of essential tau-equivalence with fixed congeneric

- components, *Multivariate Behavioral Research* Vol. 32(4), pp. 329-353.
- [25] Robinson, P.B., Stimpson, D.V., Huefner, J.C., & Hunt, H.K. (1991). "An attitude approach to the prediction of entrepreneurship", *Entrepreneurship Theory & Practice*, Summer, Vol.15 (4), pp.13-30.
- [26] Schumacher, G. Bettye Grigsby, and Winona Vesey, (2015). "Determining effective teaching behaviors through the hiring process", *International Journal of Educational Management*, Vol. 29 (1) pp. 139 -155.
- [27] Shah, Mahsood and Sid Nair, Chenicheri, (2012). "The changing nature of teaching and unit evaluations in Australian universities", *Quality Assurance in Education*, Vol. 20 (3), pp. 274-288.
- [28] Shevlin, M., Banyard, P., Davies, M.N.O. & Griffiths, M.D. (2000). The validity of student evaluations in higher education: Love me, love my lectures? Assessment and Evaluation in Higher Education, 25, 397-405.
- [29] Smithson John, Melanie Birks, Glenn Harrison, Chenicheri Sid Nair, and Marnie Hitchins, (2015). "Benchmarking for the effective use of student evaluation data", *Quality Assurance in Education*, Vol. 23(1), pp. 20-29.
- [30] Tanner, Samuel Jaye (2016). "Storying the classroom: storytelling and teacher evaluation", *English Teaching: Practice & Critique*, Vol. 15(2) pp. 208-220.
- [31] Van der Merwe, Daniel Carl. (2015). "The Usefulness of Student Evaluations for Enhancing the Effectiveness of Teaching of Financial Accounting Students at a South African University" *Accounting in Africa*. Published online: 09 Mar 2015; 107-126.
- [32] Wertz, C., Linn, R., Jöreskog, K., (1974). "Intra-class Reliability Estimates: Testing Structural Assumptions", *Educational and Psychological Measurement*, Vol.34 (1), pp. 25-33.

Appendix 1. Questions for Course Evaluation Survey (CES)

Questions about the start of the course

1. The course outline (including the knowledge and skills the course was designed to develop) was made clear to me.
2. The things I had to do to succeed in the course, including assessment tasks and criteria for assessment, were made clear to me.
3. Sources of help for me during the course including faculty office hours and reference material, were made clear to me.

Questions about what happened during the course

4. The conduct of the course and the things I was asked to do were consistent with the course outline.
5. My instructor(s) were fully committed to the delivery of the course. (Eg. classes started on time, instructor always present, material well prepared, etc).
6. My instructor(s) had thorough knowledge of the content of the course.
7. My instructor(s) were available during office hours to help me.
8. My instructor(s) were enthusiastic about what they were teaching.
9. My instructor(s) cared about my progress and were helpful to me.
10. Course materials were of up to date and useful. (texts, handouts, references etc.).
11. The resources I needed in this course (textbooks, library, computers etc.) were available when I needed them.
12. In this course effective use was made of technology to support my learning.
13. In this course I was encouraged to ask questions and develop my own ideas.
14. In this course I was inspired to do my best work.
15. The things I had to do in this course (class activities, assignments, laboratories etc) were helpful for developing the knowledge and skills the course was intended to teach.
16. The amount of work I had to do in this course was reasonable for the credit hours allocated.
17. Marks for assignments and tests in this course were given to me within reasonable time.
18. Grading of my tests and assignments in this course was fair and reasonable.

Evaluate of the course

19. What I learned in this course is important and will be useful to me.
20. This course helped me to improve my ability to think and solve problems rather than just memorize information.
21. This course helped me to develop my skills in working as a member of a team.
22. This course improved my ability to communicate effectively.

Overall Evaluation

23. Overall, I was satisfied with the quality of this course.

Open Ended Items

24. What did you like most about this course?
25. What did you dislike most about this course?
26. What suggestion(s) do you have to improve this course?

Appendix 2. Factors of JUC student evaluation of teaching effectiveness and their attributes

Factor	Attributes
The instructor's personality	1. My instructor(s) were fully committed to the delivery of the course. (Commitment) 2. My instructor(s) were available during office hours to help me. (Punctuality) 3. My instructor(s) were enthusiastic about what they were teaching. (Enthusiasm) 4. My instructor(s) cared about my progress and were helpful to me. (Caring)
The instructor's knowledge and teaching abilities	5. My instructor(s) had thorough knowledge of the content of the course. 6. In this course effective use was made of technology to support my learning. 7. In this course I was encouraged to ask questions and develop my own ideas. 8. In this course I was inspired to do my best work.
The course attributes	9. The course outline (including the knowledge and skills the course was designed to develop) was made clear to me. 10. The things I had to do to succeed in the course, including assessment tasks and criteria for assessment, were made clear to me. 11. Sources of help for me during the course including faculty office hours and reference material, were made clear to me. 12. The conduct of the course and the things I was asked to do were consistent with the course outline. 13. Course materials were of up to date and useful. (texts, handouts, references etc.). 14. The resources I needed in this course (textbooks, library, computers etc.) were available when I needed them.
Assessment and grading	15. The things I had to do in this course (class activities, assignments, laboratories etc.) were helpful for developing the knowledge and skills the course was intended to teach. 16. The amount of work I had to do in this course was reasonable for the credit hours allocated. 17. Marks for assignments and tests in this course were given to me within reasonable time. 18. Grading of my tests and assignments in this course was fair and reasonable.
Learning outcomes	19. What I learned in this course is important and will be useful to me. 20. This course helped me to improve my ability to think and solve problems rather than just memorize information. 21. This course helped me to develop my skills in working as a member of a team. 22. This course improved my ability to communicate effectively.
Overall Satisfaction	23. Overall, I was satisfied with the quality of this course