

**INSTRUCTORS' PERCEPTION OF USEFULNESS OF  
STUDENT EVALUATION FOR TEACHER AND  
COURSE DEVELOPMENT: A CASE STUDY OF AAU,  
THE DEPARTMENT OF ENGLISH**

**A THESIS SUBMITTED TO THE SCHOOL OF GRADUATE  
STUDIES**

**IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF MASTER OF ARTS IN TEACHING  
ENGLISH AS A FOREIGN LANGUAGE (TEFL)**

**BY  
AKALEWOLD DAMTEW**



**SEPTEMBER, 2011**

ADDIS ABABA UNIVERSITY  
SCHOOL OF GRADUATE STUDIES

INSTRUCTORS' PERCEPTION OF USEFULNESS OF  
STUDENT EVALUATION FOR TEACHER AND COURSE  
DEVELOPMENT: A CASE STUDY OF AAU, THE  
DEPARTMENT OF ENGLISH

By  
**Akalewold Damtew**



**Approved by board of examiners:**

\_\_\_\_\_  
**Chairman, Department Graduate  
Committee**

*Amanuel G*

**Advisor**

\_\_\_\_\_  
**Examiner, External**

*Abebe Gebretsadine*

**Examiner, Internal**

\_\_\_\_\_  
*Amanuel*

\_\_\_\_\_  
*Abebe G/B*

## **Acknowledgements**

First of all, my thanks go to my Lord, Alfa and Omega, Omni-potent and Omni-present, the creator of the visible and the invisible, the Almighty, God without whom nothing is possible.

Second, I am thankful to Amanuel Gebru (Assistant professor) for his great assistance, responsiveness, sense of responsibility, friendly and motivating approach as an advisor.

Next, I am grateful to AAU's English instructors for devotion of their precious time by completing the questionnaire and being interviewed.

Then, I would like to thank my friend, Dereselign Hailu, for his useful comments and suggestions.

Then after, I am indebted to Tihitna Wubneh, the secretary of English department, for supplying me with documents related to my study.

I am also thankful to the librarians of ILS, Education, Graduate and Kennedy Libraries for serving me without boredom.

Last but not least are my thanks to Zebib Kibrom for typing the manuscripts carefully and patiently.

## Table of Content

	<b>Page</b>
Acknowledgement .....	i
Table of content .....	ii
List of table .....	v
List of aberration.....	vi
Abstract .....	vii
<b>CHAPTER ONE: INTRODUCTION</b>	
1.1 Background .....	1
1.2 Statement of the Problem .....	2
1.3 Objective of the Study .....	3
1.3.1 Main Objective .....	3
1.3.2 Specific Objectives .....	3
1.4 Research Questions .....	4
1.5 Significance of the Study .....	4
1.6 Delimitation of the Study .....	4
1.7 Limitations of the Study .....	4
1.8 Operational Definition .....	4
<b>CHAPTER TWO: REVIEW OF RELATED LITERATURE</b>	
2. Perception .....	5
2.1 Origin and Definition .....	5
2.2 The Meaning of Perception .....	6
2.3 Evaluation.....	7
2.4 Student Evaluation .....	9
2.4.1 Formative Student Evaluation .....	10
2.4.2 Summative Student Evaluation .....	11
2.5 Specific, Global and Self – Reported Items .....	13
2.6 Instructors’ Perception of Student Evaluation .....	15
2.7 Student Perception of Student Evaluation .....	17
2.8 Factors Affecting Instructors’ Perception of Student Evaluation.....	23

2.8.1 Factors affecting the quality of Student Evaluation Instruments ...	23
2.8.1.1 Reliability .....	23
2.8.1.1.1 Factors Affecting the Reliability of Student Evaluation Questionnaire (Instrument) .....	24
2.8.1.2 Validity .....	24
2.8.1.2.1 Types of Validity .....	25
2.8.1.2.1.1 Content Validity .....	25
2.8.1.2.1.2 Criterion-Related Validity .....	26
2.8.1.2.1.3 Construct validity .....	26
2.9 Factor Affecting the Validity of Students Evaluation Data .....	26
2.9.1 Class Size .....	25
2.9.2 Instructor Gender .....	27
2.9.3 The Issue of Grade .....	27
2.9.4 Instructor Rank .....	28
2.9.5 Instructor Characteristics .....	29
2.9.6 Course Type .....	30
2.9.7 Workload .....	31
2.9.8 Course Meetings .....	33
2.9.9 Meeting Time .....	33
2.9.10 Maturity and Expertise in the Subject Matter .....	34
2.9.11 Methods of the Student Evaluation Administration .....	34
2.10 Student Evaluation and Teacher Development .....	37
2.11 Student Evaluation and Course Development .....	39
2.12 Conclusion .....	40

## **CHAPTER THREE: METHODOLOGY**

3.1 Research Design .....	41
3.2 Sampling Method .....	41
3.3 Sample Size .....	41
3.4 Description of Study Participants .....	41
3.5 Data Gathering Tools .....	42

3.6 Data Analysis Procedures .....	42
3.7 The Setting of the Research .....	42
3.8 Operational Definition .....	43
3.9 Conclusion .....	44

**CHAPTER FOUR: DATA ANALYSIS AND INTERPRETATION**

4.1 Questionnaire Data .....	45
4.1.1 Instructor Teaching Performance Evaluation .....	45
4.2 Interview Data .....	63
4.3 Discussion of Results .....	67
4.3.1 Results of the Questionnaire .....	67
4.3.2 Results of the Interview .....	70
4.3.3 Comparison between the Results of the Questionnaire and Interview.....	72

**CHAPTER FIVE: CONCLUSIONS AND RECOMMENDATIONS**

5.1 Conclusions .....	73
5.2 Recommendations .....	75

Bibliography

Appendixes

## List of Table

	<b>Page</b>
I. Instructors Teaching Performance Evaluation .....	45
Table 1. Class size, Grade Leniency and Rank .....	45
Table 2. Workload, Instructor Characteristics and Meeting Time .....	47
Table 3. Number of class meetings, Ethnicity and Major course .....	48
Table 4. Minor and Common Courses .....	50
II. Course Evaluation .....	51
Table 5. Class size, Grade Leniency and Rank .....	51
Table 6. Workload, Instructor Characteristics and Meeting Time .....	52
Table 7. Number of class meetings, Ethnicity and Major Course .....	54
Table 8. Minor and Common Courses.....	55
III student-related Question .....	56
Table 9.....	56
Table 10.....	58
IV. Purposes of student evaluation .....	59
Table 11.....	59
VI. Instructor-related Questions .....	62
Table 12.....	62

## **List of Abbreviations**

<b>AAU-</b>	Addis Ababa University
<b>IAS-</b>	Instructional Assessment System
<b>SEEQ-</b>	Student Evaluation of Educational Quality
<b>SEI-</b>	Student Evaluation of Instruction
<b>SET-</b>	Summative Evaluation of Teaching
<b>SETE-</b>	Student Evaluation of Teaching Effectiveness
<b>TEFL-</b>	Teaching English as a Foreign Language



## **Abstract**

*The main objective of this study is to determine AAU's English instructors' perception of usefulness of student evaluation for teacher and course development. The study also attempts to find out factors affecting instructors' perception of student evaluation. In addition, it attempts to discover whether it is used as an input for teacher and course development.*

*In collecting the data, 28 instructors were purposely taken out of 52 instructors from the Department of English excluding instructors of literature. However, the sample was largely-male dominated because the number of male instructors was much greater than the number of female instructors. The number of male instructors was 25 and the number of female instructors was 3. In the researcher's view, the sample is good because it involves different characteristics i.e. sex, rank and specialization area. The data gathering tools were a questionnaire and an interview.*

*The data obtained from the data gathering tools were analyzed using both quantitative and qualitative approaches. Accordingly, the major findings were: most English instructors of the university believe that student evaluations are useful for teacher and course development if both the data gathering instruments and the data collected using them are valid and reliable. In addition, there are factors perceived by the instructors that affect the validity of student evaluation which are different from those mentioned by researchers i.e. lack of skill (reading and writing), lack of knowledge of methodology of teaching, lack of knowledge of usefulness of student evaluation, receiving misinformation from previous students and prejudice about the instructors and courses. However, instructors are recommended: even though student evaluations have some limitations, they should conduct both formative and summative student evaluation because it is useful for teacher and course development. Educational experts are recommended: when student evaluations are prepared, they should be rigorously validated and reliable to gather relevant data. With regard to students, they should evaluate instructors and courses accurately and in a way which is free from any bias.*

# CHAPTER ONE: INTRODUCTION

## 1.1 Background

The evaluation of teaching performance is not a new phenomenon. Doyle (1983), quoted by Marsh in Rutherford (2003), reminds us that “In Antioch in Ad350,... any father who was dissatisfied with the instruction given to his son could examine the boy, file a formal complaint to a panel and laymen, and ultimately transfer his son to another teacher if the teacher could be shown to have neglected his duty.”

In the 20<sup>th</sup> century, we find that the first student questionnaire was prepared by students at Harvard (1926), then at the University of Washington, Purdue University and the University of Texas, also in the mid-1920's. Thousands of studies have been published on students' evaluation of teaching effectiveness with a sharp increase in 1970's and 1980's. Marsh (1987) reported that 1055 publications on the topic appeared between 1976 and 1984. Remmer was the father of research into students' evaluation of teaching effectiveness with the first systematic program started in 1927. However, in 1970's, many studies conducted and published were methodologically unsound even though many of their conclusions were used as the basis for tenure and promotion policies and subsequent research. In 1980's changes took place when research paradigms and methodological standards evolved.

Rutherford (2003) reports that Bernard (2002) distinguishes three periods in the research of student evaluation since the 1970's:

- 1970-80, the focus was more on the idea of a formative evaluation and studies were looking at indicators of teaching efficiency.
- 1980-1990, the focus was more on administrative use of evaluation and studies were examining the validity of student ratings, as well as were looking for multiple ways to conduct a comprehensive evaluation.

- 1990-2000, the focus was more on the notion of teaching Dossier and studies look at the impact and the effectiveness of students' evaluation.

Arreola (2002), cited in Rutherford (2003), notes "In the last decades of the 20<sup>th</sup> century, the demands for accountability in Higher Education reached an all-time high".

What is student evaluation? Student evaluation can be defined as a process in which students rate courses and their instructors' teaching performances to give feedback to both the instructors and the management for the improvement of the quality of education. Generally, student evaluation is important for teacher and course development. However, how do instructors perceive it? Of course, perception varies from person to person (The meaning of perception, 2011). No two persons perceive things exactly in the same way. Perception depends upon one's value system. Value system, in turn, depends up on one's philosophy. However, a person should have a good perception and should also perceive things from different perspectives. A good perception is a cause for action. A person is initiated for action if and only if he/she has a good perception. Therefore, if instructors have good perceptions in student evaluation, they can be interested in it.

## **1.2 Statement of the Problem**

There are various strategies that higher institutions can employ to bring about quality education. These are: student rating, portfolios, peer review, consultation, self-reports, alumni ratings, student focus groups and interviews, observation, department review (Instruction perception of student evaluation, 2011).

As it has been mentioned above, student evaluation is one of the strategies that can be employed by higher institutions. It is also the most useful strategy since students are central and most important in the class. Because students are

central and important in the class, they observe many aspects of the teaching-learning process such as instructors' teaching effectiveness like presentation, organization and planning, the course nature like relevance of objectives, organization, breadth, sequence, scope of the course and instructors' behaviors. As a result, they want to evaluate both the course and their instructors while the course is in progress and at the end of the course. However, as perception varies from person to person, instructors may have either a positive or negative perception of student evaluations.

With regard to the practice, student evaluation has been practiced for several years in the university. However, we don't know how instructors perceive it and whether they make use of it. So far any study has not appeared on this area in the university. Thus, a study is needed to answer certain questions including the questions mentioned above.

### **1.3 Objective of the Study**

#### **1.3.1 Main Objective**

The main objective of this study is to determine instructors' perception of usefulness of student evaluations for course and teacher development.

#### **1.3.2 Specific Objectives**

The specific objectives of this study are:

1. To identify instructors' perception of student evaluation for course development.
2. To identify instructors' perception of student evaluation for teacher development.
3. To identify factors affecting instructors' perception of student evaluation.

## **1.4 Research Questions**

1. Do instructors perceive student evaluations as useful for teacher development?
2. Do instructors perceive student evaluations as useful for course development?
3. What are the factors affecting instructors' perception of student evaluation?
4. Do instructors use student evaluation as an input?

## **1.5 Significance of the Study**

This study is important to show how instructors perceive student evaluations since perception affects utility of student evaluations. And based on the findings, an overhaul may be conducted on the practice of student evaluation in the Department of English.

## **1.6 Delimitation of the Study**

As this is a case study, it is delimited to AAU, the department of English.

## **1.7 Limitations of the Study**

This study has two limitations. One is that there is no observation in the data gathering process. There are only a questionnaire and an interview. The second one is that there is not financial support granted for the research. The cost for the research was spent by the researcher.

## **1.8 Operational Definition**

For the purpose of this study, the following words and phrases are used in the following ways. Some have more than one meaning.

**Evaluation**

- Determining the worth or value of teaching performance or course effectiveness.
- Data or information obtained from data gathering instruments.
- Data gathering instrument (questionnaire, interview etc.)

**Rating**

- Providing value for teaching performance or course effectiveness.
- Value obtained for teaching performance or course effectiveness.

**Total agreement**

- The number or percentage of participants that responded strongly agree and agree together.

**Total disagreement**

- The number or percentage of respondents that responded strongly disagree and disagree together.

## CHAPTER TWO: REVIEW OF RELATED LITERATURE

### 2. Perception

#### 2.1 Origin and Definition

Perception is a word which derives from the Latin word 'perceptio, percipio' which means 'receiving, collection, perception, apprehension, a taking' (perception, 2011). The word perception has been defined by different scholars (dictionaries) in different ways. Here are some of the definitions:

Perception is way of understanding or thinking about something in a particular way. (Macmillan English dictionary, 2002)

Perception is intuitive recognition of a truth, aesthetic qualities etc.; way of seeing or understanding, (oxford current English dictionary, 1999).

Perception is a process by which we become aware of changes through the senses of sight, hearing, etc. (Oxford Advanced Learners Dictionary of the Current English, 1975)

Perception, according to perception (2011), is:

*The act or effect of perceiving, insight or intuition gained by perceiving; the ability or capacity to perceive; way of perceiving; awareness or consciousness; view; the process by which an organism detects and interprets information from the external world by means of the sensory receptors; recognition and interpretation of sensory stimuli based chiefly on memory.*

"Perception... is the process by which an organism attains awareness or understanding of its environment by organizing and interpreting sensory information." (<http://wikipedia.org/wiki/perception>).

"Perception is our sensory experience of the world around us and involves both the recognition of environmental stimuli and actions in response to these stimuli" (What is perception? 2011).

Among the definitions given above, 'way of understanding' is preferred for this study. Therefore, it is possible to say that perception is way of understanding by means of the mind even though sensation and perception (the effect of sensation in the mind) are inseparable. According to <http://sapdesignguild.org/resources/optical/illstrutions/intro definition.html>:

*Sensation usually refers to the immediate, relatively unprocessed result of stimulation of sensory receptors in the eyes ears, nose, tongue or skin. Perception, on the other hand, better describes one's ultimate experience of the world and typically involves future processing of sensory input. In practice, sensation and perception are virtually impossible to separate.*

Hence, in the title of this study the term or the phrase "instructors' perception" means "the way in which instructors understand."

## **2.2 The Meaning of Perception**

Perceptions vary from person to person (The meaning of perception, 2011). Different people perceive different things about the same situation. What we perceive can have different meanings according to the way we perceive things. One might change one's perspective or simply make things mean something else. The meaning of something will change when we look at it differently. There is no fixed meaning to anything. We can always change perspectives and change meanings.

Moreover, perception in human society is manifested in their interactions. It is an understanding among people-the way one understands the other or others. According to <http://en.wikipedia.org/wiki/perception>, "social perception is the part of perception that allows people to understand individuals and groups of their social world, and thus an element of social cognition." However, the cause for initiation for action is a good perception. So, one should have a good perception in order to be initiated for action.



## **2.3 Evaluation**

Evaluation is perhaps the most complex and least understood of the terms in measurement and evaluation (Mulu and Daniel, 2005). Inherent in the idea of evaluation is 'value'. In evaluation, we consider how worth something is we are generally interested in the relative value of the thing such as a watch, a shirt, a book or an idea. In education, evaluation is concerned with making judgment about the relative or absolute worth of entities as varied as an essay, a dance performance, class project, a specific curriculum, a new teaching method or a state-wide school system. When we evaluate, we are engaged in some process that is designed to get information that will help make a judgment about a given situation. According to Grondlund (1971), "Evaluation is the systematic process of collecting, analyzing and interpreting information to determine the extent to which pupils are achieving instructional objectives." Evaluation (2011) defines evaluation as "... systematic determination of merit, worth and significance of something or someone using criteria against a set of standards." As it is defined by Introduction to Evaluation (2011), "Evaluation is the systematic acquisition and assessment of information to provide useful feedback about some object i.e. a program, policy, technology, person, need, activity and so on." According to Evaluation Trust (2011), "Evaluation is assessing and judging the value of a piece of work, an organization or a service."

Generally, any evaluation process requires information about the situation in question. A situation is an umbrella term that takes into account such as objectives, goals, standards, procedures, and so on (Mulu and Daniel, 2005). When we evaluate, we are saying that the process will yield information regarding the worthiness, appropriateness, goodness, validity etc, of something for which a reliable measurement or assessment has been made.

There are various types of evaluation. In terms of the purpose, evaluation is divided into formative, summative, diagnostic and placement evaluation. In terms of the interpretation of assessment information, evaluation is divided into norm-referenced and criterion-referenced evaluation. However, for this study, we deal with formative and summative evaluation since courses and instructors are evaluated while the course is in progress and at the end of the time allocated to decide whether it is effective or ineffective.

The main purpose of evaluation is to help an organization reflect on what it is trying to achieve, assessing how far it is succeeding, and identify required changes (Evaluation Trust, 2011). Evaluation is often used to characterize and appraise subjects of interest in a wide range of human enterprise, including the arts, criminal justice, foundations and non-profit organizations, government, health care, and other human service. As Introduction to Evaluation (2011):

*The generic goal of most evaluations is to provide “useful feedback” to a variety of audiences including sponsors, donors, client-groups, administrators, staff, and other relevant consequences. Most often, feedback is perceived as “useful” if it aids in decision-making. But the relationship between an evaluation and its impact is not a simple one... studies that initially seem critical sometimes fail to influence short-term decisions, and studies that seem to have no influence can have a delayed impact when more congenial conditions arise. Despite this, there is broad consensus that the major goal of evaluation should be to influence decision-making or policy formulation through the provision of empirically-driven feedback.*

## **2.4 Student Evaluation**

Student evaluation is a process in which students rate courses and their instructors' performance to give feedback to both instructors and administrators. It is one of the strategies academic institutions use to bring about quality education. The two main types of student evaluations are formative and summative. Centra (1993), cited in Hobson and Talbot (2001),

says, "There exist today two very different purposes of student evaluation of teaching effectiveness within higher education (SETE) i.e. formative and summative." Similarly, Neath (1996) states that there are two main uses of student evaluation ratings: formative and summative.

#### **2.4.1 Formative Student Evaluation**

Formative student evaluation is conducted while the course is in progress. The purpose of formative student evaluation is to make instructors aware of the improvements they need in their teaching. In relation to this, Mohanty, Gretes, Flowers, Algozzine, and Spooner (2005), cited in Course Evaluation (2011), state, "Students evaluations are formative when their purpose is to help faculty members improve and enhance their teaching skills (presentation, organization, planning)." Similarly, Instruction perception of student evaluation, (2011) says, "using evaluations to inform instructors of their teaching effectiveness and to aid them in improving or enhancing their teaching constitute the formative purpose of teaching evaluation". However, the effectiveness of formative evaluation is determined by the organizational location of evaluation system. In particular, a departmental approach to evaluation is advocated by certain researchers. Advantages with this approach are: instruction and its employment are often discipline specific and faculty most closely identify with their department and are more likely to seek guidance and information within their department (Hativa, 1995), cited in Instruction perception of student evaluation, (2011). As described by Neath (1996), formative evaluations are generally used as a tool to alert the instructors to provide possible problems in their techniques. For example, the students may provide feedback such as the instructor may be speaking too quickly for the students to understand the material, their handwriting may be illegible or more in-class practice, and problems may be useful to grasp the abstract concepts. Formative evaluations have been shown to provide useful suggestions and may improve the overall quality of undergraduate education

(Griffin and Pool, 1998; Marsh and Roche, 1993). Hobson and Talbot (2001) report:

*SETE may serve formative purpose when we use informal evaluations during a semester to determine what is working well and not-so-well, to pinpoint needed changes, and to guide those changes. Teachers may have students complete a written evaluation, participate in a facilitated "classroom" dialogue (Anglo and Cross 1993) or participate in discussions with an instructor over the course of a semester (Nuhfler, 1992).*

However, Centra (1993) indicated evaluations serve a formative purpose only if the following four conditions are met. First, teachers must learn something (new information). Second, they must value the new information. Third, they must understand how to make improvements. And, finally, teachers must be motivated to make the improvements either intrinsically or extrinsically.

#### **2.4.2 Summative Student Evaluation**

Summative student evaluation is a type evaluation which takes place at the end of a semester. Summative evaluation is commonly used for summative decisions such as tenure, merit increase, retention for non-tenured faculty, promotion and course assignment decisions.

Summative evaluations are more serious than formative evaluation since there are decisions on instructors' and courses' effectiveness or ineffectiveness. Therefore, the validity of summative evaluation instruments should rigorously be tested and various stakeholders should participate in the construction of student evaluation questionnaire and the evaluation should focus on learning outcomes and should be used with other criteria. In relation to this, Instruction perception of student evaluation, (2011) reports:

*Researchers recommend that when such data are going to be used for summative purposes, various stakeholders, including administrators, faculty and students should collaborate in determining a proper evaluation system. The focus of evaluation should be on desired educational outcomes and whether these outcomes are being met. If students rating forms are to be used, the instruments must be subjected to rigorous validity tests and analysis. Further, student ratings data should be used in combination with other criteria in order to provide a better assessment of teaching, which is inherently a multidimensional construct. Multiple sources of data collected over a span of time and covering a variety of courses taught would be most effective in informing summative decision-making.*

Michelle (2011) states that summative evaluations are used to assist in major personnel decisions such as hiring, firing, promotion, tenure, and distribution of merit pay. Further, in many institutions, the sole evaluation of teaching performance may be based upon the opinions and comments of students often solicited a single time at the end of the semester. Neath (1996: 1, 363, emphasis in original), cited in Michelle (2011), states, "Instead of publish or perish, faculty can now publish and perish if their ratings on teaching evaluations are not sufficiently high." When the student evaluations become a primary measure of teaching effectiveness and ultimately a determinant of continued employment, faculty members can understandably become skeptical and fearful of the student evaluation. Hobson and Talbot (2001) report:

*The summative purpose of SETE is for use in evaluating the overall effectiveness of an instructor, particularly for tenure and promotion decisions (Centra 1993). As stated, universities rely on several sources of information about a faculty members teaching, but only SETE tend to be administered systematically (Cashin 1988), generally as end-of-course evaluation forms that are used by an entire university community.*

Summative evaluations should also be made by the current students of the class and should be confidential and anonymous, and can be performed on papers or online technology (Course Evaluation, 2011).

## **2.5 Specific, Global and Self – Reported Items**

Hobson and Talbot (2001), cited in Instruction perception of student evaluation, (2011), state the three most widely used student evaluation forms are published instruments based on specific, global and self-reported student rating questionnaires. Each system attempts to measure teaching effectiveness in terms of student learning and may be used for either summative or formative purposes. In addition to items designed to target specific dimensions of behaviors, the end-of-course evaluation forms also tend to include global items on the overall effectiveness of the teacher or the quality of the course. Although global ratings tend to correlate highly with a number of specific factors (Centra, 1993), there exists an ongoing debate about the appropriate uses of each.

As Cashin and Downey (1992) observed, “one of the continuing debates concerning the use of student ratings of teachings is the debate revolving around what kind of measure should be used for summative evaluation of faculty, in making personnel decisions for retention, promotions, tenure or salary increases, and of course, to assess their effectiveness.” Some researchers have argued that only global items should be used for summative decisions regarding tenure and promotion. Others assert either those only specific ratings should be used for summative and formative purposes (Cashin, 1988) or that both global and specific items need to be considered in summative decisions (Marsh 1991a, 1991b; Marsh and Bailey 1993). On the other hand, Instruction perception of student evaluation, (2011) reports:

*Instruction and course evaluation system created by the University of Illinois at Urbana, uses a cafeteria based model that allows instructors to choose from over four hundred globally based questions. The form is computer based and only recommended for formative use. Student instructional rating system, created by Michigan State University in 1982, focuses on specific questions regarding instructor effectiveness, but includes one global question regarding the general enjoyment of the course (Hobson and Talbot, 2001 p.29). Instructional Assessment system (IAS), created by the Educational Assessment Center at the University of Washington, uses both specific and global questions on several forms that are designed specifically for different class sizes.*

In addition, student rating forms generally contain both global (or overall rating) items and specific items, which assesses specific aspects of the instructor and course. Research is split on the value of both of these types of items. Some argue that teaching is multi-dimensional and therefore requires specific items to accurately assess different facets of teaching. Others show that when specific items are factor analyzed, they essentially reduce down to one or two items that are global in nature. Studies also reveal that responses on specific and global items are highly correlated. Gallager, 2000; d'Appollonia and Abrania and Abrami, 1977; Young and Show, 1999; Bain, 2004, cited in Instruction Perception of Student Evaluation, (2011) state that with regard to the uses of these types of items, researchers warn against making summative decision based solely on ratings on global items. In addition, formative purposes seem better informed by having data on specific areas that faculty can target in order to improve their teaching.

From the above literature of various scholars, we can understand that there is a debate about the use of global and specific items for formative and summative evaluations. With regard to the debate, Hobson and Talbot (2001) state:

*Because this debate about items use is not yet resolved, it is important for faculty members to beware of its existence and the implications of each perspective. Professors will want to inquire about the specific student evaluation form used by their departments or universities and ascertain the relative importance placed by the instrument and the institution on global and specific items.*

## **2.6 Instructors' Perception of Student Evaluation**

Instructors have various perceptions about student evaluation. These perceptions are both positive and negative. However, the negative perceptions dominate over the positive ones.

Instruction perception of student evaluation, (2011) states:

*Traditionally, faculties have been very skeptical of teaching evaluations and have previously opposed such practices. The negative feelings often spring from fear that student ratings will be used or misused for summative decision-making purposes. Moreover, faculty often believes that students do not take evaluations seriously and the ratings encourage grade leniency.*

In the above quotation, it is indicated that instructors are suspicious about student evaluation because they think it might be used for summative decisions. But if it is not used for summative decisions such as tenure, promotion, salary increase and retention and rigorously validated, they feel it is welcome. It also shows that students are not matured enough to evaluate the instructors' performance and the course. In addition, it causes grade leniency.

Rutherford (2003) states, "Faculty generally believe that students' evaluations are biased by a number of factors which they believe to be unrelated to teaching effectiveness." Hazari (2011) reports, "A study was conducted to determine perceptions of department of management faculty about the student evaluation of instruction (content and process). It was found that majority of



faculty respondents were not satisfied with SEI forms and process.” In addition, Dulz and Lyons (2011) say that student evaluation affects instructor’s classroom behavior. Therefore, it is a hindrance than a help and they add:

*Although the term “student evaluation” is regularly used, it is more realistic to talk about student rating, because that is typically what each student does. That is, the student chooses a number using a rating scale of some sort. There is little evidence in the literature of any wide use of student-volunteered evaluation criteria. In other words, students respond to the instrument presented to them. Evaluations are not often conducted by students.*

The above scholars show us that students do not evaluate their instructors and the courses seriously. They consider it as an obligation. As a result, they merely fill the questionnaire by choosing a number or a letter without understanding the questions. Moreover, Course Evaluation (2011) says, “Summative student evaluations of teaching (SETE) have been widely criticized, especially by teachers, for not being accurate measures of teaching effectiveness” and reports:

*surveys have shown that a majority of teachers believe that a teacher’s raising the level of standards and/or content would result in worse SETE for the teachers and that students in filling out SETE are biased in favor of certain teacher’s personalities, looks, disabilities, gender and ethnicity. The evidence that some of these critics cite indicates that factors other than effective teaching are more predictive of favorable ratings. Many of those who are critical of SETE have suggested that they should not be used in decisions regarding faculty hires, retentions, and tenure. Some have suggested that using them for such purposes leads to the dumbing down of educational standards. Others have said that the typical way SETE is now used at most universities is demeaning to instructors and has a corrupting effect on students’ attitude towards their teachers and higher education in general.*

On the other hand, there are some scholars that perceive student evaluation positively. Burnard (2002), cited in Rutherford (2003), surveying 1200 faculty, found that the faculty's perception is that all faculty and course must be evaluated but students should not be the only source of information and that faculty must be allowed to take part. Evaluations help to improve teaching. Students written comments are an interesting source of information. Instruction perception of student evaluation, (2011) reports:

*Studies have emerged showing how institutions and individual faculty members have begun using evaluations, consultations, and portfolios to quantitatively improve instruction. When faculty are well informed about the purposes of evaluation, much of their anxiety dissipates and willingness to learn from students feedback increases and adds "... evaluations are most effective in improving teaching when faculty members understand and evaluate the importance of such processes, and an institutional and departmental culture that supports and respects teaching is evident.*

## **2.7 Student Perception of Student Evaluation**

Students are central and most important in the class. The presence of the instructor and all other things is for students. As they are central and most important in the teaching-learning process, they observe many aspects of the teaching-learning process such as instructor teaching performance, their behaviors, course objectives, organization, breadth, depth, sequence and relevance of topics. Therefore, their feedback is very important. Instruction perception of student evaluation, (2011) reports:

*Research shows that students tend to take teaching evaluations more seriously than faculty and institutional members commonly believe students are more willing to participate and offer meaningful feedback when they believe and can see their input is being considered and incorporated by their instructors and institution. In general, however, students do not perceive that their feedback is being used. Some studies show that students believe their input should be considered for summative purposes. Students would like to see more specific items related to teaching effectiveness on student evaluation of teaching instruments (Sojka and Deeterschmetz, 2002; Chen and Hoshower, 2003).*

Burnard (2002), cited in Rutherford (2003), reports that most students surveyed out of a 400 sample thought that evaluation is very important to improve teaching and formation of students. However, they thought that it has little impact on formation of students and little influence on faculty practices; faculty don't read their comments; they are not considered for promotion; universities are not interested in teaching; teaching is not important to them and evaluation is important but not useful. As a result, they call it: "operation bidon". In addition, TOEFL iBT-Sample Essay (2011) declares:

*Student can feel more involved in their studies because this will encourage teachers to prepare for their lessons more effectively. By being asked to evaluate their teachers, the students are able to feel that they are directly involved in their education because they will be paying closer attention to their teachers and material. In other words, this involvement presents a better chance for students to realize their full potential. If students spend their time merely sitting at their desks with no feeling of direct involvement, they benefit less from classes and do not achieve their full potential.*

Moreover, Rutherford (2003) states that student ratings are good predictors of instructor's effectiveness. It is, therefore, necessary to be realistic in our expectations of what student ratings can achieve in terms of feedback and information to interpret carefully the results, but also to supplement this evaluation with additional information. TOFL iBT-sample Essay (2011) states, "Teachers have always a sense of superiority in class because the job of the students is to sit and listen to them. However, if the teachers know that the students are going to make a judgment concerning their abilities and lesson, they will make a stronger effort to prepare more effective lesson."

<http://blog-ecollege.com/wordpress/?p=5> states that most instructors think they are fair in their instructional behavior. However, studies have shown that students often do not rate their instructors as high on the "fairness scale".

Perception of fairness, the very definition of fairness, often differs between instructors and their students. According to Dr. Rodabaugh's typology of perceived fairness provides an excellent starting point: interactional fairness-the nature of the interaction between instructor and student(s), procedural fairness the policies for assessment and course administration, and outcome fairness-the distribution of grades among students.

### **Interactional Fairness**

The majority of students consider interactional fairness as key to a fair course environment. The prevalent view appears to be that if interactional fairness is present, then fairness in grades and other important (to students) areas will follow. Fairness in the interactional area is perceived to include:

**Respect:** students commonly enter a course respecting the instructor and they want to be respected in turn respect includes being polite to all students in the tone used in an announcements and syllabus, and calm and professional with individual students in postings to discussions, emails, and feedback on assignments and exams. If a student must be corrected in any way, it should be done in a private manner.

**Impartiality:** students expect an instructor to treat everyone in the course equally. In the online environment, equality of treatment is the most apparent in the tone and frequency of an instructor responding to individual students in a threaded discussion or chat room interaction.

**Concern:** students want their instructors to care about them and their success in pursuit of their educational goals. Concern can begin with an effort to get to know students in the course syllabus. Concern can also be shown by having an online office, emailing students who seem to be struggling, and providing constructive feedback on assignments.

**Integrity:** integrity revolves around clarity in presenting course policies and procedures and consistency in applying them to all students in all situations. This argues for a well-written, thorough course syllabus that ensures students understand policies and procedures and any consequences for failure to comply with course requirements.

**Propriety:** students expect instructors to act in socially acceptable ways and to follow the rules. In the online environment, propriety is most clearly applicable in communications with students whether in discussion postings, emails to individuals or teams, or comments on written assignments and exams.

### **Procedural Fairness**

From the students viewpoint procedural fairness also ranks above outcome fairness. Again, if procedures are fair as established and followed fairly, then students believe that fair outcomes will logically follow. Fairness in procedural area is perceived to include:

**Policies:** policies should be seen as fair by students. To this end, policies and procedures should be clearly set forth and explained in the course syllabus. Policies of particular concern to most students involve such items as participation in discussions, plagiarism, other students' participation in team assignments, and the handling of late assignments. The degree to which policies are applied equally reflects upon perceived instructor integrity, a factor in interactional fairness.

**Workload:** students must see the required workload as reasonable given the type of course overview /objectives as stated in the school's catalog. What is reasonable will depend, to some degree, on the type of students in courses and the level of students in the school. Commonly, the work load will be higher for

high school students than middle school and higher in advanced courses than lower level courses.

**Assessments:** students appear to consider a number of factors when deciding if assessments (exams and quizzes) are fair. For example:

- Does the assessment questions align with the learning objectives in the syllabus?
- Do the assessment questions relate to content presented in the course?
- Is the level of difficulty of assessment questions appropriate for the level of course and the students?
- Are the assessment questions well designed?

**Feedback:** students consistently provide high ratings to online courses where the instructor presence is visible. Providing constructive feedback on assignments and assessments help ensure that students

1. View these activities as fair and
2. View the course in a positive light.

**Student input:** student involvement in the educational process increases the likelihood that they will view a course as “fair”. One way to involve students is to periodically solicit their input on the course. This holds true at both the middle and high school levels.

### **Outcome Fairness**

Grades, obviously, are important to students. In general, students want

1. To know the policies and rubrics by which they are graded and
2. Their grades to reflect their actual performance. Fairness in the outcome area is perceived to include:

**Information:** students believe that fairness requires they be given detailed information, preferably written in the course syllabus, specifying:

- The assessment tools and graded assignments that will be used to determine their course grade.
- The weight each assessment tool and graded assignment will have in determining their course grade.
- The grading rubric in a percentage or point basis as well as information about grade determination. For example, is a present cut-off score used or relative standing in the student group?

**Policy change:** students expect that the grading policies and rubric set in the course syllabus will be firm and applied equally to all students. If a grading policy is altered for any reason, the rationale should be explained fully to the students.

**Multiple assessments:** students commonly feel that the use of multiple assessment tools is “fairer” than reliance on a single type of assessment, such as a mid-term and final exam. In the online environment, instructors frequently use a mix of exams and quizzes, written assignments, and participation in discussions. This “portfolio” approach to assessment is generally well-received by students.

**Individual performance:** almost universally, students believe that fairness requires their grade to be based on their individual performance. This holds true for team or group projects as well. Students feel that their grade should reflect their contribution to the team assignment and not all.



## **2.8 Factors Affecting Instructors' Perception of Student Evaluation**

There are two factors that affect instructor's perception of student evaluation. These are: factors affecting the quality of student evaluation instruments (lack of reliability and validity of student evaluation) and factors affecting the validity of student evaluation data (class size, instructor gender, the issue of grade, instructor rank, instructor characteristics, course type, workload, course meeting, meeting time, maturity and expertise in the subject matter and methods of student evaluation administration)

### **2.8.1 Factors affecting the quality of Student Evaluation Instruments**

The quality of student evaluation instrument can be affected if it lacks reliability and validity.

#### **2.8.1.1 Reliability**

Instruction perception of student evaluation, (2011) defines reliability as, "the consistency of ratings among raters and also the stability of such ratings over time" and reports:

*Research has shown that student ratings show an acceptable level of consistency, or inter-rater reliability, given a class size of at least 15. The level of consistency among raters increases as class size increases. Longitudinal studies and studies of alumni ratings of an instructor/course have found that ratings show high levels of stability over time. Further, cross-sectional studies show that student ratings reliably reflect instructor versus course effectiveness (Hobson and Talbot, 2001; Aleamoni, 1999; Marsh and Roche, 1997).*

According to the quotation of the above scholar, student evaluation can be affected due to class size. That is, when the number of students in a class decreases, the reliability of the evaluation decreases and vice-versa. Besides,





studies that gather data over a long period of time repeatedly enables to obtain high levels of stability. However, studies that gathers data only at a time show student rating which is course versus instructor. Therefore, student evaluation instrument should have reliability in order to obtain reliable data.

#### **2.8.1.1.1 Factors Affecting the Reliability of Student Evaluation Questionnaire (Instrument)**

There are a number of factors that can cause student evaluation data to fluctuate. The main ones are (Wiersam and Jurs, 1985): error of measurement, range of students' performance on the questionnaire, length of the questionnaire, item similarity and item difficulty. Error of measurement can be caused by student rate of work on the questionnaire, fatigue (instability in individual performance on the questionnaire), lack of consistency in measuring the data, directions and use of answer sheet, if the same questionnaire is administered twice or parallel forms are used. Overall measurement conditions should be so structured to minimize errors of measurement as much as possible. Length of the questionnaire is the other factor that affects the reliability of student evaluation. That is as the number of questions decreases, the reliability decreases and vice-versa. Similarity of items and item difficulty can also be problems for the reliability of student evaluation. If items are similar, it is impossible to get the required amount of data. If items are difficult, it is difficult to get the correct data from students (Mulu and Daniel, 2005).

#### **2.8.1.2 Validity**

Instruction perception of student evaluation, (2011) defines reliability as "the extent to which student evaluation questionnaire measures what it is intended to measure." and states:

*The validity of rating is difficult to confirm, however, as there is no agreed upon definition of what constitutes effective teaching, as a result, most researchers have adopted a construct validation or convergent validity approach, which involves examining the correlation between student ratings and other partial measures of teaching effectiveness, such as student learning, observer, peer, and alumni ratings. Studies in this vein have reported generally positive and significant correlations between these measures and students ratings. In addition, research on the validity of ratings has aimed to identify potential biasing factors not directly related to teaching effectiveness (Hobson and Talbot, 2001; Aleamuni, 1999; Kulik, 2001; Greenwald and Gillmore, 1976b; Mckeachie, 1997).*

From the above quotation, it is understood that there is no conventional effective teaching criteria, so it is difficult to assume a given student evaluation as valid. Therefore, instructors are forced to use what they think (their constructs) make student evaluation valid and it should be used with other measures of teaching effectiveness to check their correlations. However, there are potential biasing factors that are not related to effective teaching and affect the validity of the evaluation. So, this should be taken in to consideration and ways should be found to eliminate them.

#### **2.8.1.2.1 Types of Validity**

There are three types of validity. These are content, criterion-related and construct validity.

##### **2.8.1.2.1.1 Content Validity**

Content validity may be defined as the extent to which a student evaluation questionnaire contains a representative sample of the content under consideration Mulu and Daniel (2005). The content of the questionnaire may be defined in terms of the content of the items and their objectives.

The focus of content validity then is on the adequacy of the sample and not simply on the appearance of the questionnaire. A questionnaire that appears to be relevant measure, based on superficial student evaluation, is said to have face validity. Face validity should not be considered a substitute for content validity. As might be expected, content validity is of primary concern in achieving reliable data from the questionnaire.

#### **2.8.1.2.1.2 Criterion-Related Validity**

Whenever data from questionnaires are to be used to predict future performance (predictive validity) or to estimate current performance (concurrent validity) on some valued measure other than the questionnaire itself, we are concerned with criterion related validity (Mulu and Daniel, 2005). Criterion-related validity may be defined as the extent to which data from questionnaire is related to some other valued measures of performance (called criterion).

#### **2.8.1.2.1.3 Construct validity**

Construct validity is used to interpret student evaluation data in terms of some general psychological quality (Mulu and Daniel 2005). It refers to certain degree of students' reasoning ability on issues.

### **2.9 Factor Affecting the Validity of Student Evaluation Data**

There are certain factors researchers believe might affect the validity of student evaluations. These are: class size, instructor gender, the issue of grade, instructor behavior, course type, workload, instructor rank, time of class meeting, number of class meetings, maturity and expertise in the subject matter and methods of student evaluation administration.

### 2.9.1 Class Size

One of the factors researchers believe might affect student evaluation is class size. The problem with large class size is that instructor-student interaction will become low since instructors can not interact with large number of students. As a result, they might provide low ratings. Rutherford (2003) reports:

*A review of the literature (Marsh, 1987, Feldman, 1984, Centra, 1979 and Arreola, 2002) shows that class size is moderately correlated with the most logically related variables (for instance group interaction and individual rapport for the SEEQ instrument), but not with other dimensions, nor with the overall ratings of the course instructor.*

On the other hand, in his study Centra (1979) found that classes in the 35-100 ranges were found to receive the lowest ratings while high rating were received by larger and smaller classes. According to Arreola (2002):

*Low rating for large classes are not due to the class-size effect but to the fact that required and first year courses tend to be offered in large class setting since research tells us that required course are usually rated harsher, and that 1<sup>st</sup> year students tend to rate also more harshly, this lower rating is expected.*

The results with regard to class size have not been absolute. The literature quoted in Lersch (2011) reported:

*Marsh and Roche (1979) found that while most researches report that smaller classes are rated somewhat more favorably, the results of studies have been mixed. Gleason (1979) reported that of thirty studies that examined the effects of course size on students evaluations; twenty found that smaller classes received higher rating. However, these relationship were typically modest; class size only counted for anywhere from one to eight percent of the variance in ratings. Gleason argued that low student evaluation ratings given to instructors in large classes should not be dismissed simply due to the size of the class.*

### 2.9.2 Instructor Gender

The second factor that researchers believe might affect student evaluation is the gender of the instructor. Lersch (2011) reports:

*The research exploring the effects of instructor attributes on student evaluation has been equivocal. With respect to the gender of the instructor, a number of studies have found that female instructors are evaluated lower than their male peers (Bruschke and Gartner, 1991; Hearn, 1985; Sadler, 1991). In a review essay, Neath (1996) provided a number of tips that a faculty member could use to increase their teaching effectiveness ratings. The first choice was to be male. Citing the work of Bennet (1982) and Unger (1979), Neath argues that students expect a high level of support and nurture from female faculty members. When this expectation is not met, the students provide low ratings of instructional effectiveness to female faculty.*

*However, the findings concerning the effects of instructor gender on student ratings is not absolute. Marsh and Roche (1997), in a review of studies, found mixed support for the hypothesis that women would be evaluated lower than men and concluded instructor gender would ultimately have little or no effect on overall student ratings. A few studies, for example, (Rowden and Carlson, 1996) have found that female instructors were rated significantly higher than male instructors.*

Rutherford (2003) indicates that the literature on the gender of instructor hasn't any influence on student ratings-both male and female instructors are rated equally and tells us:

*A heated debate continues about the influence on student evaluations of a teacher's sex and gender-role orientation (Masculine, feminine, androgynous, or undifferentiated) (Basow and Howe 1987; Hobson 1997; Sidanius and Crane 1989). With respect to student characteristics, an interaction effect appears to exist between students and instructor's sex/gender-role orientation, but it is not yet convincingly documented.*

Depending up on the above scholars, it is understood that although the role of instructor's gender is doubtful, female instructors receive low rating if they don't behave as a mother does since this is expected of mothers in most cultures. Moreover, the sex or gender-role between students and instructor may have potential bias on the validity of student evaluation. For instance, a female instructor may like male student, as a result, female students may tend to provide low ratings to female instructors.

### **2.9.3 The Issue of Grade**

The third factor that researchers believe might affect student evaluation is the issue of grade. Most researchers believe that high ratings have positive relationships with students' receiving high grades. As quoted in Lersch (2011):

*A frequently stated idea among faculty is that instructors who grade more leniently than others may be doing so in order to receive better student evaluations. Gift grades put students in a reciprocity situation. As phrased by Grenwald and Gillmore (1997:1209), "any consideration of strategies to increase ratings is likely to focus quickly on the very simplest strategy that is suggested by academic folklore-just give high grades." Grades have been described as 'the basic currency of our educational system' (Tata, 1999: 263) and numerous studies have explored the effects of grade leniency on student evaluation.*

A number of researches have concluded that higher grades have an effect on student evaluations: the higher the grades awarded, the higher the ratings (Chako, 1983; Neath, 1996; Tata, 1999). However, the strength of this relationship has been debated while a number of studies have reported a positive correlation were relatively low-from .10 to .30 (Cashin, 1995). In addition, Course Evaluation (2011) states, "Grades assigned by instructors have positive relationship with student evaluations of teaching (SET)". Similarly, the same idea is reported by Dulz and Lyons (2011). It reads,

“Students who scored higher on exams tended to give their teachers higher ratings... students place a higher value on instructor behavior which enables them to achieve higher evaluations from the instructor.”

To sum up, from the most of the above researchers, it is possible to understand that grades assigned to students have positive relationship with student ratings- as grades assigned increase, student ratings increase, and vice-versa.

#### **2.9.4 Instructor Rank**

The fourth factor that researchers believe might affect the validity of student evaluation is instructor rank (assistant lecturers, assistant professors, full professor, and Ph.D holders). Regarding the effects of the instructor rank on student evaluation, mixed results have been obtained (Lersch, 2011). In a review of 50 studies on student evaluations ratings, Eison (1999), cited in Lersch (2011), reported that while about half found that faculty of higher rank received better evaluations, an equal number found no relationship between rank and evaluation. Similar results were reported by Marsh and Roche (1997), who noted that instructor rank has little or no impact on student ratings. Marsh (1987), cited in Rutherford (2003), notes no relation between rank and global ratings, but a slightly positive correlation with the dimension “Breadth of coverage” and slightly negative correlation with the dimension group interaction.

The above scholars show us that instructor rank is not generally connected with student evaluations. However, they indicate that instructors with high rank receive, to some extent, high ratings with regard to breadth of coverage but receive, to some extent, less ratings with regard to group interaction.

### **2.9.5 Instructor Characteristics**

Instructor characteristics is the fifth factor that researchers think might affect the validity of student evaluation. Students tend to provide high ratings to instructors who knows a lot, lectures in tandem with the textbook and is very entertaining (Dulz and Lyons, 2011). In a similar way, Instruction perception of student evaluation, (2011) states, "student evaluation is a popularity contest." Hamermesh and West (2005) indicate:

*Student at the University of Texas at Austin gave attractive instructors higher SET (student Evaluation of Teaching) scores than less attractive instructors. However the author concluded that it may not be possible to determine if attractiveness increases the effectiveness of an instructor, possibly resulting in better learning outcomes. It may be the case that students pay more attention to attractive instructors... some papers has examined online course evaluations, such as ratemy professor. Com and found them to be heavily influenced by the instructor's attractiveness and reported easiness.*

As cited in Rutherford (2003), Aleamoni and Specer (1973), Aleamoni (1976), and Marsh (1987) state "... the instructor's characteristics (Warm, friendly, approachable) had no effect on the lack of organization of the instructor, or the poor stimulating strategies used in the course".

From the above scholar, it is understood that it is difficult to know the effect of attractiveness on student evaluations .it is difficult to know whether attractiveness adds to instructors' effectiveness or attractiveness causes students to pay more attention so that leaning outcomes increase. This shows that there are debates about the relationship between instructor characteristics and student evaluation.

### **2.9.6 Course Type**

The sixth factor that researchers believe might affect the validity of student evaluation is course type. Lersch (2011) reports:



*One general belief shared by many instructors is that students enrolled in certain courses, such as research methods or statistics, may evaluate the overall effectiveness of the instructor lower than other courses that are viewed as more appealing. And, there is some research to support this belief: students do tend to give higher evaluations in classes in which they expressed a prior interest or were taking as electives (Marsh and Cooper, 1981; Marsh and Dukins, 1992). High evaluations are given when the reason for taking a course is 'general interest' or 'major electives' rather than 'major requirement' or 'general educational requirements' (Marsh and Roche, 1997).*

Reasons for taking a course: major requirement, major elective, general interest, general education requirement, minor /related field and other can also affect student evaluation. In addition, Rutherford (2003) reports:

*Marsh (1980) reports that taking a course as a major requirement or as a general education requirement had a negative effect on all factors.*

*In fact, 'general interest' had the strongest effect on all variable (although usually less than 0.20). A review of the literature shows 1<sup>st</sup> year students tend to rate a course more harshly than 2<sup>nd</sup> year students, 2<sup>nd</sup> year students tend to rate a course more harshly than 3<sup>rd</sup> year students, etc. And that graduate level courses tend to receive slightly higher ratings.*

The above scholar shows us that it is not only the case of the type of the course but also the case of the level of students that affects student evaluation. That is, as the level of the students increases, the severity of student evaluation decreases. Because, in the lower level, students expect the level of assistance they were provided by teachers when they were in high school. So, when they observe this is absent, they tend to provide low ratings.

### **2.9.7 Workload**

The seventh factor researchers believe might affect the validity of student evaluation is workload. In his studies, Marsh (1980, 1982), cited in Rutherford (2003), states that workload/difficulty was also found to have relationship with

student rating. This means that if students are given too many assignments, they think they are suffering from course work and deprived of their freedom for things apart from their education. As a result, they might provide low ratings.

### **2.9.8 Course Meetings**

The eighth factor that researchers believe affect student evaluation is the issue of number of course meetings (contact hours). Eison (1999), cited in Lersch, states that course meetings has not received a great deal of attention in the filled of research.

In the case of course meetings, the assumption is that students think that they have learned a lot if they have a large number of course meetings. And course meetings increase instructor-student interaction so that awareness about the instructor increases. Therefore, they may rate their instructors and the course high when there are high course meetings and vice-versa.

### **2.9.9 Meeting Time**

The ninth factor that researchers believe might affect student evaluation is meeting time i.e. morning, afternoon, evening, winter summer, etc. It is obvious that there is a difference between meeting in the morning and in the after-noon and meeting in the winter and in summer. For example, in the afternoon the temperature of the air increases and suffocation is created. This time students may not feel comfortable. As a result, they may develop hatred towards the instructor as well as the course. However, in a review of 50 studies, Eison (1999), cited in Lersch (2011), reports, "Only three of the studies examined the effects of course meeting time on course evaluations. In these studies, no significant relationship between meeting time and students' evaluation ranking was found."

### **2.9.10 Maturity and Expertise in the Subject Matter**

The tenth factor that researchers believe might affect the validity of student evaluations is maturity of students. As cited in Lersch (2011), Schemelik, Spencer and Gellman (1997) state, "It has been also argued that students are not qualified to render evaluations concerning the overall effectiveness of their instructors because of their lack of maturity, experience, and expertise in the subject matter". In relation to this, Rutherford (2003) says that there is a difference in ratings between former students and students who are just completing the course because former students of a certain course have added perspective of time and experience on the job; therefore, they rate more efficiently than those that are just completing the course. Instruction perception of student evaluation, (2011) States, "... students are not able to make informed and consistent judgments about their instructors; students rating are essentially a popularity contest."

In the above scholars' literature, it is understood that evaluation requires maturity (age) which is a factor for experience and knowledge in the subject matter. Therefore, students are not efficient enough to evaluate the instructor performance and the course. Therefore, due to lack of maturity, their evaluation is a popularity contest which is not directly related to learning outcomes.

### **2.9.11 Methods of the Student Evaluation Administration**

The eleventh factor that researchers believe affect the validity of student evaluation is method of student evaluation administration. There are mainly two types of student evaluation methods. These are paper and pencil which is conducted at class and student evaluation form to be completed electronically. Some researchers say that paper and pencil student evaluation creates stress in students while completing student evaluation data forms in class. As a result, the validity and reliability of student evaluation data is affected. But

when students complete student evaluation electronically (online), they feel free and this contributes to the validity and reliability of the student evaluation. There are also some researchers who state that time and location in which student evaluation are administered affects the validity of student evaluation.

In relation to this, Course Evaluation (2011) declares:

*Beginning in the mid-1970's, course evaluations were typically conducted using stamper paper forms that students completed individually in class at, or near, the end of a quarter or semester. Among the enormous advances in technology that have been made since then, the single most important change in the course evaluation area has been the option of completing evaluation forms electronically. However, whichever method of collecting evaluation data is used, the undisputed goal is to obtain reliable and credible information about the quality of faculty instruction. Further, to the extent that the process by which these are collected is flawed, achieving the crucial objective of "good data" will similarly be compromised. Another way of putting this is to say that an institution must place a strong emphasis on the quality control measures deployed throughout the course evaluation process. The term quality control is being used here to emphasize the point that the procedures and condition that under-gird the collection of course evaluation data should be used such that students have the opportunity to evaluate each of their courses in a consistent firmament. If this is achieved, course evaluation results within an institution can be readily compared across different academic departments, between courses within a specific academic department, and between courses for a given instructor.*

*The ideal conditions for collecting "clean" course evaluation data are to have all students in a given course complete the evaluation form at the same time on the same date, and in the same location. This ideal can be accomplished via-paper or electronically (in the latter case, with a suitable internet connection, students could access the evaluation form in class using their laptops, tablets, or smart phones), they can be closely approximated. By making as uniform as possible the condition under which evaluations are completed, unwanted sources of variability in the process can be kept to the*

that the faculty's perception is that all faculty and course must be evaluated but students should not be the only source of information and that faculty must be allowed to take part. Evaluations help to improve teaching. Students' written comments are interesting source of information.

With regard to student perception, Sojka and Deeterschmetz, 2002; Chen and Hoshower, 2003) state that students tend to take teaching evaluations more seriously than faculty and institutional members commonly believe students are more willing to participate and offer meaningful feedback when they believe and can see their input is being considered and incorporated by their instructors and institutions. In general, however, students do not perceive that their feedback is being used. Some studies show that students believe their input should be considered for summative purposes. Students would like to see more specific items related to teaching effectiveness on student evaluation of teaching instruments. Burnard (2002) reports that most students surveyed out of 400 sample thought that evaluation is very important to improve teaching and formation of students. However, they thought that it has little impact on formation of students and little influence on faculty practices; faculty don't read their comments; and they are not considered for promotion. Rutherford (2003) states that student ratings are good predictors of instructors' effectiveness. It is, therefore, necessary to be realistic in our expectations of what student ratings can achieve in terms of feedback and information to interpret carefully the results, but also to supplement this evaluation with additional information.

## **CHAPTER THREE: METHODOLOGY**

### **3.1 Research Design**

This is a study conducted to determine instructors' perception of usefulness of student evaluation for teacher and course development. In the study, both quantitative and qualitative approaches were used. However, as it was a cross-sectional study, the data were collected once. First, data were collected using a questionnaire. Then, data were collected using an interview by way of triangulating the data collected using the questionnaire.

### **3.2 Sampling Method**

The sampling method for both the questionnaire and the interview was purposive sampling. This means instructors who were well-experienced, and who were thought to give relevant data were chosen to fill the questionnaire and to be interviewed.

### **3.3 Sample Size**

The sample size for the questionnaire was 28 instructors. These 28 instructors were chosen out of 52 instructors from the department of English excluding instructors of literature. The sample size for the interview was 4 instructors who were from those that filled the questionnaire.

### **3.4 Description of Study Participants**

In this study, there were both male and female participants. However, it was largely male-dominated because the number of male instructors was much greater than the number of female instructors-nearly all were male instructors. The number of male participants was 25 but the female ones were 3. Among the 25 male participants, there were 9 Ph.D holders, 1 assistant professor and 15 assistant lecturers. Among the three female participants, one was a Ph.D holder and the other two were assistant lecturers. These 28 instructors were taken out of 52 instructors from the department of English. The number of

participants for the interview was 4 who were from those that completed the questionnaire. Among these, one was a Ph.D holder but the rest, 3 instructors, were assistant lecturers. All the participants for the interview were male instructors.

### **3.5 Data Gathering Tools**

As it has already been mentioned in the research design, the data gathering tools for the study were a questionnaire and an interview.

### **3.6 Data Analysis Procedures**

In this study, there were both quantitative and qualitative data. The questionnaire had questions to collect both quantitative and qualitative data. But the interview had questions to collect only qualitative data.

Therefore, the quantitative data were tabulated using numbers and then changed in to percent for interpretation. However, for the qualitative data, what were commonly responded by the instructors was taken as data for interpretation.

### **3.7 The Setting of the Research**

The setting of the research was Addis Ababa University (AAU), Sidist Kilo Campus, the department of English. AAU is a university which is found in the north of Addis Ababa. It is the first university in Ethiopia. It was the palace of Emperor Haile Selassie I. Then it became a university in 1961 G.C. It is one of the biggest universities in Ethiopia. There are three main buildings: OCR (Old classroom), NCR (New classroom) and NBR (New building room) in addition to administration and academic offices, student dormitories and buildings under construction.

AAU, sidist kilo campus, is a place for social science. There are certain faculties such as Faculty of Language Studies, Faculty of social Sciences (Geography, History, Sociology, etc.), Faculty of Law, Faculty of Business and Management and Faculty of Informatics. It offers both undergraduate and post-graduate programs in various fields. In the Faculty of Language Studies, Foreign Language Studies and Ethiopian Languages are found. Among the departments in Foreign Language Studies, one is the Department of English. The Department of English offers both undergraduate and post-graduate programs. Among the post graduate programs, one is TEFL (Teaching English as a Foreign Language).

### **3.8 Conclusion**

In the researcher's view, the sample size was sufficient because the research literature states that from large population, a sample size of 25% is enough but for this study whose population is small, it should be much more than 25%. The sample size for this study was 28 instructors which was taken out of the population of 52 instructors i.e. nearly 54%. With regard to the nature of the participants, it was good because it involved different characteristics i.e. sex, rank and specialization area. As to the data gathering instruments, a questionnaire and an interview were used because these were the data gathering instruments that were believed enable to gather relevant data. The setting was suitable since the participants were not scattered to gather the data.



## CHAPTER FOUR: DATA ANALYSIS AND INTERPRETATION

This section presents the data analysis and interpretation. The data collected from the questionnaire are tabulated and then are analyzed and interpreted. The data from the interview is presented in reported and non-reported forms.

### 4.1 Questionnaire Data

#### 4.1.1 Instructor Teaching Performance Evaluation

##### Rating scales for Interpretation

The following rating scales show the interpretation of the difference of the total agreement and disagreement in number and percent.

a difference of 0 (0%)= equal

a difference of 1-2 (3.6-7.1%)= negligible (nearly equal)

a difference of 3-7 (10.7-25%)= relatively most

a difference of 8-15 (28.6-57.1%)= most

a difference of 16-27 (60.7-96.4%)= nearly all

a difference of 28 (100%)= all

##### Key

**S. agree- Strongly agree**

**S. disagree-Strongly disagree**

**Table 1. Class size, Grade Leniency and Rank**

No	Item	Number of instructors that responded				
		S. Agree	Agree	Undecided	Disagree	S.Disagree
1	Instructors with large class sizes receive low ratings	1	8	11	6	2
2	Students provide high ratings to instructors who provide grades leniently.	6	18	4	0	0
3	Instructors with high rank receive high ratings.	0	1	13	12	2

Table 1 shows that 3.6% of the instructors strongly agree that instructors with large class sizes receive low ratings, 28.6% affirmed their agreement to the statement, 39.3% expressed their indecision; however, 21.4% disagreed to it and 7.1% confirmed their disagreement. Totally, 32.2% of the instructors expressed their agreement but 28.5% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 3.7%. Therefore, this means that some instructors believe that instructors with large class sizes receive low ratings but others don't.

On the same Table, 21.4% of the instructors strongly agree that students provide high ratings to instructors who provide grades leniently, 64.3% affirmed their agreement to the statement, 14.3% expressed their indecision; however, 0% disagreed to it and 0% confirmed their disagreement. Totally, 85.7% of the instructors expressed their agreement but 0% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 85.7%. Therefore, this means that nearly all instructors think that instructors that provide grades leniently receive high ratings.

Again on the same Table, 0% of the instructors strongly agree that instructors with high rank receive high ratings, 3.6% affirmed their agreement to the statement, 46.4% expressed their indecision; however, 42.9% disagreed to it and 7.1% confirmed their disagreement. Totally, 3.6% of the instructors expressed their agreement but 50% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 46.4%. Therefore, this means that most instructors don't believe instructors with rank receive high ratings.

**Table 2. Workload, Instructor Characteristics and Meeting Time**

No	Item	Number of instructors that responded				
		S. Agree	Agree	Undecided	Disagree	S.Disagree
4	Instructors who are attractive, entertaining and approachable receive high ratings.	4	21	1	2	0
5	Instructors who teach courses with high work load receive low ratings	4	10	9	5	0
6	Instructors who teach courses in an unfavorable time receive low ratings.	0	12	16	0	0

Table 2 indicates that 14.3% of the instructors strongly agree that instructors who are attractive, entertaining and approachable receive high ratings, 75% affirmed their agreement to the statement, 3.6% expressed their indecision; however, 7.1% disagreed to it and 0% confirmed their disagreement. Totally, 89.3% of the instructors expressed their agreement but 7.1% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 82.2%. Therefore, this means that nearly all instructors believe that instructors who are attractive, entertaining and approachable receive high ratings.

On the same Table, 14.3% of the instructors strongly agree that instructors who teach courses with high work load receive low ratings, 35.7% affirmed their agreement to the statement, 32.1% expressed their indecision; however, 17.9% disagreed to it and 0% confirmed their disagreement. Totally, 50% of the instructors expressed their agreement but 17.9% confirmed their disagreement. This shows that the difference between the percentage of the respondents that

expressed their agreement and disagreement is 32.1%. Therefore, this means that most instructors think that instructors that teach courses with high work load receive low ratings.

Again on the same Table, 0% of the instructors strongly agree that instructors who teach courses in an unfavorable time receive low ratings, 42.9% affirmed their agreement to the statement, 57.1% expressed their indecision; however, 0% disagreed to it and 0% confirmed their disagreement. Totally, 42.9% of the instructors expressed their agreement but 0% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 42.9%. Therefore, this means that nearly all instructors believe that instructors that teach in an unfavorable time receive low ratings.

**Table 3. Number of class meetings, Ethnicity and Major course**

No	Item	Number of instructors that responded				
		S. Agree	Agree	Undecided	Disagree	S.Disagree
7	Instructors who teach courses with less contact hours receive low ratings.	2	4	12	9	1
8	Students provide high ratings to instructors who are in the same ethnic group they belong to	1	6	11	7	2
9	Instructors who teach major courses receiver high ratings.	1	9	11	7	0

Table 3 shows that 7.1% of the instructors strongly agree that instructors who teach courses with less contact hours receive low ratings, 14.3% affirmed their

agreement to the statement, 42.9% expressed their indecision; however, 32.1% disagreed to it and 3.6% confirmed their disagreement. Totally, 21.4% of the instructors expressed their agreement but 35.7% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 14.3%. Therefore, this means that relatively most instructors don't think that instructors that teach courses with less contact hours receive low ratings.

On the same Table, 3.6% of the instructors strongly agree that students provide high ratings to instructors who are in the same ethnic group they belong to, 21.4% affirmed their agreement to the statement, 39.3% expressed their indecision; however, 25% disagreed to it and 0% confirmed their disagreement. Totally, 25% of the instructors expressed their agreement but 32.1% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 7.1%. Therefore, this means that some instructors believe that students provide high ratings to instructors who are in the same ethnic group students belong to but others don't.

Again on the same Table, 3.6% of the instructors strongly agree that instructors who teach major courses receive high ratings, 32.1% affirmed their agreement to the statement, 39.3% expressed their indecision; however, 25% disagreed to it and 0% confirmed their disagreement. Totally, 35.7% of the instructors expressed their agreement but 25% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 10.7%. Therefore, this means that relatively most instructors think that instructors that teach major courses receive high ratings.

**Table 4. Minor and Common Courses**

No	Item	Number of instructors that responded				
		S. Agree	Agree	Undecided	Disagree	S.Disagree
10	Instructors who teach minor courses receive low ratings	1	7	13	7	0
11	Instructors who teach common courses receive low ratings	1	1	8	18	0

Table 4 indicates that 3.6% of the instructors strongly agree that instructors who teach minor courses receive low ratings, 25% affirmed their agreement to the statement, 46.4% expressed their indecision; however, 25% disagreed to it and 0% confirmed their disagreement. Totally, 28.6% of the instructors expressed their agreement but 25% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 3.6%. Therefore, this means that some instructors believe that instructors that teach minor courses receive low ratings but others don't

On the same Table, 3.6% of the instructors strongly agree that instructors who teach common courses receive low ratings, 3.6% affirmed their agreement to the statement, 28.6% expressed their indecision; however, 64.3% disagreed to it and 0% confirmed their disagreement. Totally, 7.2% of the instructors expressed their agreement but 64.3% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 57.1%. Therefore, this means that nearly all instructors don't think that instructors that teach common course receive low ratings.

## II Course evaluation

**Table 5. Class size, Grade Leniency and Rank**

No	Item	Number of instructors that responded				
		S. Agree	Agree	Undecided	Disagree	S.Disagree
1	Courses with large class sizes receive low ratings.	2	7	8	9	2
2	Courses taught by instructors who provide grades leniently receive high ratings.	7	10	5	4	2
3	Courses taught by instructors with high ranks receive high ratings	1	2	9	16	0

Table 5 shows that 7.1% of the instructors strongly agree that courses with large class sizes receive low ratings, 25% affirmed their agreement to the statement, 28.6% expressed their indecision; however, 32.1% disagreed to it and 7.1% confirmed their disagreement. Totally, 32.1% of the instructors expressed their agreement but 39.2% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 7.1%. Therefore, this means that some instructors believe that courses with large class sizes receive low ratings but others don't.

On the same Table, 25% of the instructors strongly agree that courses taught by instructors who provide grades leniently receive high ratings, 35.7% affirmed their agreement to the statement, 17.9% expressed their indecision; however, 14.3% disagreed to it and 7.1% confirmed their disagreement. Totally, 60.7% of the instructors expressed their agreement but 21.4% confirmed their disagreement. This shows that the difference between the percentage of the

respondents that expressed their agreement and disagreement is 39.3%. Therefore, this means that most instructors believe that courses taught by instructors that provide grades leniently receive high ratings.

Again on the same Table, 3.6% of the instructors strongly agree that courses taught by instructors with high ranks receive high ratings, 7.1% affirmed their agreement to the statement, 32.1% expressed their indecision; however, 57.1% disagreed to it and 0% confirmed their disagreement. Totally, 10.7% of the instructors expressed their agreement but 57.1% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 46.4%. Therefore, this means that most instructors don't believe that courses taught by instructors with high rank receive high ratings.

**Table 6. Workload, Instructor Characteristics and Meeting Time**

No	Item	Number of instructors that responded				
		S. Agree	Agree	Undecided	Disagree	S.Disagree
4	Courses taught by attractive, entertaining and approachable instructors receive high ratings.	8	16	0	4	0
5	Courses with high workload receive low ratings	1	10	10	7	0
6	Courses taught in unfavorable time receive low ratings.	0	14	3	10	1

Table 6 indicates that 28.6% of the instructors strongly agree that courses taught by attractive, entertaining and approachable instructors receive high ratings, 57.1% affirmed their agreement to the statement, 0% expressed their



indecision; however, 14.3% disagreed to it and 0% confirmed their disagreement. Totally, 85.7% of the instructors expressed their agreement but 14.3% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 71.4%. Therefore, this means that nearly all instructors think that courses taught by attractive, entertaining and approachable instructors receive high ratings.

On the same Table, 3.6% of the instructors strongly agree that courses with high workload receive low ratings, 35.7% affirmed their agreement to the statement, 35.7% expressed their indecision; however, 25% disagreed to it and 0% confirmed their disagreement. Totally, 39.3% of the instructors expressed their agreement but 25% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 14.3%. Therefore, this means that relatively most instructors believe that courses with high workload receive low ratings.

Again on the same Table, 0% of the instructors strongly agree that Courses taught in unfavorable time receive low ratings, 50% affirmed their agreement to the statement, 10.7% expressed their indecision; however, 35.7% disagreed to it and 3.6% confirmed their disagreement. Totally, 50% of the instructors expressed their agreement but 39.3% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 10.7%. Therefore, this means that relatively most instructors believe that courses taught in an unfavorable time receive low ratings.

**Table 7. Number of class meetings, Ethnicity and Major Course**

No	Item	Number of instructors that responded				
		S. Agree	Agree	Undecided	Disagree	S.Disagree
7	Courses with less contact hours receive low ratings.	0	8	9	9	2
8	Courses taught by instructors who are in the same ethnic group students belong to receive high ratings.	1	6	10	9	2
9	Major courses receive high ratings.	1	7	8	11	1

Table 7 shows that 0% of the instructors strongly agree that courses with less contact hours receive low ratings, 28.6% affirmed their agreement to the statement, 32.1% expressed their indecision; however, 32.1% disagreed to it and 7.1% confirmed their disagreement. Totally, 28.6% of the instructors expressed their agreement but 39.2% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 10.6%. Therefore, this means that relatively most instructors believe that courses with less contact hours don't receive low ratings.

On the same Table, 3.6% of the instructors strongly agree that Courses taught by instructors who are in the same ethnic group students belong to receive high ratings, 21.4% affirmed their agreement to the statement, 35.7% expressed their indecision; however, 32.1% disagreed to it and 7.1% confirmed their disagreement. Totally, 25% of the instructors expressed their agreement but 39.2% confirmed their disagreement. This shows that the difference

between the percentage of the respondents that expressed their agreement and disagreement is 14.2%. Therefore, this means that relatively most instructors don't think that course taught by instructors who are in the same ethnic group students belong to receive high ratings.

Again on the same Table, 3.6% of the instructors strongly agree that major courses receive high ratings, 25% affirmed their agreement to the statement, 28.6% expressed their indecision; however, 39.3% disagreed to it and 3.6% confirmed their disagreement. Totally, 28.6% of the instructors expressed their agreement but 42.9% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 14.3%. Therefore, this means that relatively most instructors don't think that major courses receive high ratings.

**Table 8. Minor and Common Courses**

No	Item	Number of instructors that responded				
		S. Agree	Agree	Undecided	Disagree	S.Disagree
10	Minor courses receive low ratings	1	7	7	13	0
11	Common courses receive low ratings	1	5	7	15	0

Table 8 indicates that 3.6% of the instructors strongly agree that minor courses receive low ratings, 25% affirmed their agreement to the statement, 25% expressed their indecision; however, 46.4% disagreed to it and 0% confirmed their disagreement. Totally, 28.6% of the instructors expressed their agreement but 46.4% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 17.8%. Therefore, this means that relatively most instructors don't believe that minor courses receive low ratings.

On the same Table, 3.6% of the instructors strongly agree that common courses receive low ratings, 17.9% affirmed their agreement to the statement, 25% expressed their indecision; however, 53.6% disagreed to it and 0% confirmed their disagreement. Totally, 21.3% of the instructors expressed their agreement but 53.6% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 32.3%. Therefore, this means that most instructors don't think that common courses receive low ratings.

### III student-related Question

**Table 9**

No	Item	Number of instructors that responded				
		S. Agree	Agree	Undecided	Disagree	S.Disagree
1	Students are too immature to evaluate their teachers.	0	6	2	10	10
2	Students are too immature to evaluate courses.	0	6	2	10	10
3	Students lack expertise in the subject matter to evaluate their instructors.	2	12	0	11	3

Table 9 shows that 0% of the instructors strongly agree that students are too immature to evaluate their teachers, 21.4% affirmed their agreement to the statement, 7.1% expressed their indecision; however, 35.7% disagreed to it and 35.7% confirmed their disagreement. Totally, 21.4% of the instructors expressed their agreement but 71.4% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 50%. Therefore, this means that most instructors don't believe that students are too immature to evaluate their instructors.

On the same Table, 0% of the instructors strongly agree that students are too immature to evaluate courses, 21.4% affirmed their agreement to the statement, 7.1% expressed their indecision; however, 35.7% disagreed to it and 32.1% confirmed their disagreement. Totally, 21.4% of the instructors expressed their agreement but 71.4% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 50%. Therefore, this means that most instructor don't believe that students are too immature to evaluate courses.

Again on the same Table, 7.1% of the instructors strongly agree that students lack expertise in the subject matter to evaluate their instructors, 42.9% affirmed their agreement to the statement, 0% expressed their indecision; however, 39.3% disagreed to it and 10.7% confirmed their disagreement. Totally, 50% of the instructors expressed their agreement and again 50% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 0%. Therefore, this means that some instructors believe that students lack expertise in the subject matter to evaluate their instructors but others don't.

**Table 10**

No	Item	Number of instructors that responded				
		S. Agree	Agree	Undecided	Disagree	S.Disagree
4	Students lack expertise in the subject matter to evaluate courses.	2	12	0	11	3
5	Student evaluations are measures of teachers' popularity, not teacher effectiveness.	2	10	4	11	1
6	Students are inaccurate in their judgment.	3	10	6	8	1

Table 10 indicates that 7.1% of the instructors strongly agree that students lack expertise in the subject matter to evaluate courses, 42.9% affirmed their agreement to the statement, 0% expressed their indecision; however, 39.3% disagreed to it and 10.7% confirmed their disagreement. Totally, 50% of the instructors expressed their agreement and again 50% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 0%. Therefore, this means that some instructors believe that students lack expertise in the subject matter to evaluate courses and others don't.

On the same Table, 7.1% of the instructors strongly agree that Student evaluations are measures of teachers' popularity, not teacher effectiveness, 35.7% affirmed their agreement to the statement, 14.3% expressed their indecision; however, 39.3% disagreed to it and 3.6% confirmed their disagreement. Totally, 50% of the instructors expressed their agreement and again 50% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and

disagreement is 0%. Therefore, this means that some instructors believe that student evaluations are measures of teachers' popularity but others don't.

Again on the same Table, 10.7% of the instructors strongly agree that students are inaccurate in their judgment, 35.7% affirmed their agreement to the statement, 21.4% expressed their indecision; however, 28.6% disagreed to it and 3.6% confirmed their disagreement. Totally, 46.4% of the instructors expressed their agreement but 32.2% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 14.2%. Therefore, this means that relatively most instructors believe that students are inaccurate in their judgments.

#### IV. Purposes of student evaluation

**Table 11**

No	Item	Number of instructors that responded				
		S. Agree	Agree	Undecided	Disagree	S.Disagree
1	If student evaluations are rigorously validated and reliable, they are good sources of feedback.	10	15	0	2	1
2	Student evaluations are important for teacher professional development.	7	19	1	0	1
3	Student evaluations are important for course development.	7	17	3	0	1
4	Student evaluation is a hindrance, not a help	0	2	3	15	8

Table 11 shows that 35.7% of the instructors strongly agree that if student evaluations are rigorously validated and reliable, they are good sources of feedback, 53.6% affirmed their agreement to the statement, 0% expressed their indecision; however, 7.1% disagreed to it and 3.6% confirmed their disagreement. Totally, 89.3% of the instructors expressed their agreement but 10.7% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 78.6%. Therefore, this means that nearly all instructors believe that if students' evaluations are rigorously validated and reliable, they are good sources of feedback.

On the same Table, 25% of the instructors strongly agree that student evaluations are important for teacher professional development, 67.9% affirmed their agreement to the statement, 3.6% expressed their indecision; however, 0% disagreed to it and 3.6% confirmed their disagreement. Totally, 92.1% of the instructors expressed their agreement but 3.6% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 88.5%. Therefore, this means that nearly all instructors believe that student evaluations are important for teacher professional development.

Again on the same Table, 25% of the instructors strongly agree that student evaluations are important for course development, 60.7% affirmed their agreement to the statement, 10.7% expressed their indecision; however, 0% disagreed to it and 3.6% confirmed their disagreement. Totally, 85.7% of the instructors expressed their agreement but 3.6% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 82.1%. Therefore, this means that nearly all instructors believe that students' evaluations are important of course development.



Again on the same Table, 0% of the instructors strongly agree that student evaluation is a hindrance, not a help, 7.1% affirmed their agreement to the statement, 10.7 expressed their indecision; however, 53.6% disagreed to it and 28.6% confirmed their disagreement. Totally, 7.1% of the instructors expressed their agreement but 82.2% confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 75.1%. Therefore, this means that nearly all instructors don't believe that student evaluations are a hindrance.

## **V. What do you think are the advantages and disadvantages of student evaluation?**

### **Advantages**

The respondents mentioned their perceived advantages of student evaluation as follows:

1. It helps for self-sensation that invites for more exploration of professional development.
2. If its purpose is developmental, it can help the teacher improve himself/herself in terms of knowledge, teaching skills, and attitudes.
3. It makes teachers work hard.
4. It can be a base to plan for the future
5. It may prevent instructors from mistreating students for fear of being evaluated adversely.
6. It can help the teacher re-examine his/her teaching and relationship with his/her students.
7. It can help instructors improve their courses.
8. It keeps teachers "on track"
9. It helps instructors handle classes more effectively and efficiently.

### **Disadvantages**

The respondents mentioned the disadvantages of student evaluation as follows:

1. It may be influenced by factors unrelated to teaching effectiveness.

2. Students do not evaluate objectively. Their evaluation is directly connected with grades.
3. It is usually based on instructors' behaviors.
4. Students give immature and un-informed decisions.
5. It may not be useful if it is not properly set by educational experts.
6. If not properly done, it de-motivates hard-working teachers.
7. Students' may be influenced by the type of the course.
8. It is sometimes unfairly used by ineffective instructors and faculty administrators to attack hard working instructors.

## VI. Instructor-related Questions

**Table 12**

No	Item	Number of instructors that responded	
		Yes	No
	Do you feel comfortable with student evaluation of your course?	24	4
	Are you satisfied with the current student evaluations?	11	17

Table 11 indicates that 85.7% of the instructors strongly agree that they feel comfortable with student evaluation of their courses; however, 14.3% affirmed their disagreement to the question. Totally, 85.7% of the instructors expressed their agreement but the rest, 14.3%, confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 71.4%. Therefore, this means that nearly all instructors feel com for table with student evaluation of their courses.

On the same Table, 39.3% of the instructors strongly agree that they satisfied with the current student evaluations; however, 60.7% affirmed their

disagreement to the question. Totally, 39.3% of the instructors expressed their agreement but the rest, 60.7%, confirmed their disagreement. This shows that the difference between the percentage of the respondents that expressed their agreement and disagreement is 21.4%. Therefore, this means that relatively most instructors are not satisfied with the current students' evaluations.

## **4.2 Interview Data**

The following is an interview between the researcher and four respondents. There are questions followed by their responses respectively.

Q1. Do you think student evaluation is useful? Why?

Respondent 1 replied, "Yes, because I must get feedback about my methodology of teaching in order to make students feel comfortable with my teaching and enable them to achieve the objectives of the lesson (course). So, student evaluation helps me to look for other ways to improve my teaching, accomplish my tasks on time, check, balance and create integrity."

Respondent 2 answered yes and explained that it is important to improve his way of teaching, to get feedback on the content, assessment and evaluation if it is constructive or valid and reliable.

Respondent 3 responded, "Yes, because it provides me with feedback about my way of teaching and the course provided they evaluate properly."

Respondents 4 explained it is useful if it is done properly; it helps him as a source of feedback.

Q2. What are factors that affect the validity of student evaluation? How do they affect it?

Respondent 1 replied, "The factors are: lack of the skill of students, lack of maturity, and lack of knowledge in the subject matter, lack of knowledge about

methodology of teaching and irrelevance of their evaluation. For example, they want to learn grammar because they don't know why they learn English."

Respondent 2 answered that the factors are: lack of knowledge of the usefulness of student evaluation, weakness of how to evaluate and receiving misinformation from previous students and added that if student evaluation is valid and reliable, it is very important.

Respondent 3 responded, "The factors are: lack of objectivity, lack of relevance of their evaluations, lack of knowledge about student evaluation, lack of seriousness and receiving misinformation from previous students about the instructor. Therefore, the information obtained from them can not show me the weaknesses of my teaching and can't be used for course development."

Respondent 4 replied, "Most of the items in the evaluation form do not reflect the actual work of the instructor. This means their evaluations are irrelevant."

Q3. Do you think students evaluate instructors reliably? Why?

Respondent 1 answered no because students give ratings in exchange for grades. Therefore, laissez-faire instructors receive good ratings because they provide grades leniently.

Respondent 2 responded, "It depends on what kind students are. Weak students provide ratings in relation to grades provided to them. They have also prejudice about their instructors and courses. In addition, they don't understand the usefulness of student evaluation."

Respondent 3 replied no because they don't understand the usefulness of student evaluation.

Respondent 4 answered, "No, students may be biased against or in favor of an instructor due to different reasons."

Q4. How do you respond to student evaluation?

Respondent 1 responded, "I strongly believe that instructors should be evaluated by students but the way they are evaluated is not often good. Moreover, student evaluation is not used- it is not applicable. I have taught in the university for more than 5 years but I have received the evaluation copy only once."

Respondent 2 replied that it is good because it shows him to see his weaknesses and strengths.

Respondent 3 answered, "It is very important but students don't have an understanding about the usefulness of student evaluation although I myself tell them to show me my weaknesses."

Respondent 4 responded that it depends on what students say.

Q5. What are the purposes of student evaluation?

Respondent 1 replied, "It is important for teachers' professional and course development in general. It enables them identify their strengths and weaknesses. It is also important to bring about quality education and distinguish between the efficient and inefficient teachers there by upgrading the inefficient ones."

Respondent 2 answered that it helps the teacher with his professional development- it is important to improve what to teach, how to teach and how to handle students. It is even important for the development of the country in one way or the other.

Respondent 3 replied, "It is important for teacher development: teaching skill, presentation, assessment, evaluation and career development (salary increase, position, etc)."

Respondent 4 answered that it is important to give them feedback about their teaching performances."

Q6. What do you think are the advantages and disadvantages of student evaluation?

Respondent 1 responded, "It is generally important for teacher and course development but it has certain disadvantages. For example the teacher will become unmotivated if he/she is not evaluated reliably. It also causes teachers to give grades leniently. In addition it causes teachers to consider it useless because they don't get anything (salary increase, opportunities, positions, etc) for their good student evaluation achievement.

Respondent 2 replied that the advantage is that it shows instructors the weaknesses of their teaching. But the disadvantage is that student evaluation is not reliable. Students rate instructors in exchange for grades.

Respondent 3 answered, "If it is validated and reliable, it has only advantages. It is useful for well-informed decisions and good teaching. If it is not valid and reliable, it damages the image and morale of the instructor."

Respondent 4 responded that its advantage is that it is a very important source of feedback on every aspect of the teaching and learning. But its disadvantage is that students may not base their evaluation on the actual work of their instructors.

Q7. Do you feel comfortable with student evaluation of your course? Why?

Respondent 1 replied, "I can't say yes or no because I have never seen any evaluation copy of mine."

Respondent 2 responded that they are evaluated at the end of courses. However, there is no student evaluation since two years ago. Besides they don't get the feedback.

Respond 4 replied of course and added that he wants to know how they find his teaching.”

Q8. Have you been recently evaluated? If your answer is yes, have you been satisfied with it? Why?

Respondent 1 answered, “My name has been submitted for evaluation. But I don't know whether I have been evaluated or not because I have not received the feedback.”

Respondent 2 responded yes and explained that the evaluation made by the evening and summer students was good because they are matured, but the regular students don't evaluate reliably. They associate it with grades they are provided. Besides if the teacher is to give them a course in the future, they don't evaluate reliably.

Respondent 3 replied, “Yes, because it was fair.”

Respondent 4 answered, “I have not been recently evaluated.”

### **4.3 Discussion of Results**

In this section, the results of the questionnaire in comparison with the review literature, the results of the interview in summary and comparison between the results of the questionnaire and the interview are presented.

#### **4.3.1 Results of the Questionnaire**

In this study, there are factors supposed to affect student evaluation of teaching performance and course effectiveness and the results are as follows:

Based on the data collected, some instructors agree that instructors with large class sizes receive low ratings and others don't and this is the same as what is stated in the review literature with regard to grade, in similarity to the review

literature, almost all the instructors that provide grades leniently receive high ratings. Relatively most instructors don't perceive instructors with high rank receive high ratings but in the review literature, the effect of rank is neutral-some research findings tell us that it affects student evaluation and others tell us otherwise. Concerning instructor characteristics, like the review literature, nearly all instructors affirm that instructors who are attractive, entertaining and approachable receive high ratings. In the same way as we can understand from the review literature; relatively most instructors believe that instructors with high work load receive low ratings. When we look at the data on time of meeting, nearly all instructors confirm that instructors that teach courses in an unfavorable time receive low ratings but in the review literature its impact is neutral. Relatively most instructors perceive that instructors that teach courses with less contact hours receive low ratings but the review literature tells us that contact hours is given little attention-there are no research findings. With reference to ethnicity, some instructors believe that students provide high ratings to instructors who are in the same ethnic group they belong to but others don't but in the review literature, it says that students provide high ratings to instructors who are in the same ethnic group they belong to and vice-versa. When we see the data on major course, relatively most instructors agree that instructors that teach major course receive high ratings and this is in agreement with the review literature. As to minor courses, some instructors affirm that instructors that teach minor courses receive low ratings but other don't but this is a bit in contrast to the review literature because required courses are evaluated harsher. Hence, low ratings are expected. Nearly all instructors believe that instructor that teach common courses receive high ratings in contrast to the review literature because required courses are evaluated harsher like minor courses.

Above, the results of factors affecting student evaluation of teaching performance have been shown. However, surprisingly, although the degrees vary, most of the ratings provided to factors affecting students evaluation of



instructors' teaching performance was also provided to the evaluation of course effectiveness and vice-versa. There are differences on major course, ethnicity and minor course. However, there is a big difference on major course but slight differences on ethnicity and minor course. However, the ratings provided to instructors' teaching performance may not be provided to the effectiveness of courses and vice-versa; this may be lack of seriousness of the students. Or the instructors might have not taken this in to consideration when they completed the questionnaire; instructors and courses may not be equally effective or ineffective.

As we can access the data on questions related to students, nearly all instructors don't perceive students are too immature to evaluate their instructors' teaching performances and courses but this is contrary to the review literature. With reference to expertise in the subject mater, some instructors affirm that students lack expertise in the subject matter to evaluate their instructors and courses respectively but others don't; however, this is different from the review literature because in the review literature, it says that students lack expertise in the subject matter to evaluate their instructors and courses. Regarding the validity of student evaluations, relatively most instructors don't believe that student evaluations are measures of teacher's popularity but this is neutral to what is stated in the review literature. Concerning the accuracy of student evaluation, relatively most instructors agree that students are inaccurate in their judgments and this is in conformity to the review literature.

As we can see the data on questions related to the purposes of student evaluation, nearly all instructors confirm that if student evaluations are rigorously validated and reliable, they are good sources of feedback and this is in accordance with the review literature. Again, nearly all instructors perceive that student evaluations are important for teacher and course development and the same is stated in the review literature. Moreover, nearly all instructors

don't agree that student evaluation is a hindrance but this is neutral to the review literature.

As it can be depicted from the data on Yes-No questions, nearly all instructors feel comfortable with student evaluation of their courses because it gives them useful feedback. However, most instructors are not satisfied with the current student evaluations because students do not evaluate their instructors and courses reliably. In addition their evaluation is biased by a number of factors. As to the results of advantages and disadvantages of student evaluation, see "Data Analysis and Interpretation (v)."

### **4.3.2 Results of the Interview**

Below the results of the interview are presented in a summary form. There are questions followed by their summaries.

1. Do you think student evaluation is useful? Why?

Student evaluation is useful to get feedback about teaching methodology and courses. It is also useful to achieve objectives and improve assessment and evaluation skills. Moreover, it is important to make students feel comfortable with the teaching-learning process.

2. What are factors that affect the usefulness of student evaluation? How do they affect it?

Factors that affect the usefulness of student evaluations are: lack of skill, lack of maturity, lack of knowledge in the subject matter, lack of understanding about usefulness of student evaluation, lack of relevance of their evaluation lack of seriousness, prejudice and receiving misinformation about instructors from previous students. However, factors like class size, gender rank, course type and instructor characteristics, contact hours, meeting time, and methods of student evaluation administration have been given little attention by the instructors.

3. Do you think students evaluate instructors reliably? Why?

Most instructors believe that instructors should be evaluated by their students and their feedback is very important if both the student evaluation instruments and the data collected from them are valid and reliable. However, student evaluation is not applicable because the feedback is not communicated to the instructors.

4. What are the purposes of student evaluation?

It is useful to show strengths and weaknesses of instructors. It is also important to identify the efficient and inefficient ones there by improving the inefficient ones. It is, in general, important for teacher and course development; it brings about quality education. At large, it is important for the development of the country in one way or the other.

5. What do you think are the advantages and disadvantages of student evaluations?

Student evaluation has certain advantages and disadvantages. The advantages are: 1) It is important because it shows weaknesses in teaching, 2) It is important for good teaching, 3) It is important to give well-informed decisions, 4) It is important for career development. The disadvantages are, 1) Students rate their instructors and the course in exchange for grades, 2) Students may not base their evaluation on the actual work of the instructors, 3) It damages the image of the instructor, 4) It affects the morale of the instructor 5) It makes instructors consider it useless because they don't get anything for their good student evaluation achievements.

6. How do you respond to student evaluation?

Instructors believe that instructors should be evaluated by students because student evaluation shows them their weaknesses and strengths. However, it is not often good because they don't understand its usefulness. In addition, the feedback is usually not communicated to them.

7. Do you feel comfortable with student evaluation of your course? Why?

Some instructors feel comfortable with student evaluation of their course because some students provide them with relevant information and they want to know how students find their teaching. However, there is no student evaluation since two years ago. Besides, they don't often get the feedback.

8. Have you been recently evaluated? If your answer is yes, have you been satisfied with it? Why?

Some instructors have not been recently evaluated. And a few don't know whether they have been evaluated or not. But others say that the evaluations made by the evening and summer students are good because they are matured. But the regular students don't evaluate reliably. They associate it with grade. Specially, if the instructor is to give them a course in the future, they don't evaluate their instructor genuinely.

#### **4.3.3 Comparison between the Results of the Questionnaire and Interview**

Depending on the data, the results of questionnaire and the interview are almost similar except for the factors affecting the validity of student evaluation data. In the interview, only a few of the factors that scholars believe may affect the validity of student evaluation data have been mentioned i.e. the issue of grade and lack of maturity and expertise in the subject matter. And the other difference is that in the interview, there are a few factors perceived by the respondents which are different from those presented in the review literature. These are: lack of knowledge of teaching methodology, lack of skills (reading, writing, etc), lack of understanding about the usefulness of student evaluation, receiving misinformation from previous students and prejudice about the instructors and courses.

## **CHAPTER FIVE: CONCLUSIONS AND RECOMMENDATIONS**

### **5.1 Conclusions**

In general, based on the data collected, it is possible to say that most English instructors of the university believe that student evaluation is useful for teacher and course development if both student evaluation instruments and the data collected using them are valid and reliable. There are also advantages and disadvantages that they perceive. The advantages are: 1) It helps for self-sensation that invites for more exploration of professional development. 2) It is important in terms of knowledge, teaching skills and attitudes. 3) It is useful as a base to plan for the future. 4) It makes instructors work hard 5) It prevents instructors from mistreating students for fear of being evaluated adversely 6) It helps the teacher with his/her teaching and relationship with his/her students 7) It helps instructors to enrich their courses 8) It keeps teachers "on track" 9) It helps instructors handle classes more effectively and efficiently. The disadvantages are: 1) Students don't evaluate their instructors objectively; their evaluation is directly connected with grades provided to them. 2) It is usually based on instructors' behaviors 3) Students might give immature and unformed decisions 4) It may not be useful if it is not properly set by educational experts 5) If it is not properly done, it un-motivates hard-working instructors 6) Student evaluations may be influenced by the type of the course 7) It is sometimes unfairly used by ineffective instructors and faculty administrators to attack hard-working instructors 8) Students may receive misinformation from previous students. 9) Students may have prejudice about their instructors and courses.

In addition to these perceived advantages and disadvantages, depending on the data, there are certain factors believed by the instructors that affect the validity of student evaluation. Among these, the main ones are the issue of grade, instructor characteristics, (attractiveness, the nature of entertaining and approachability), unfavorable meeting time and major courses, which are not

directly related to learning outcomes. To elaborate, students who are provided high grades provide high ratings and vice-versa. Students also provide high ratings to instructors who are attractive, entreating and approachable. In addition, they provide low ratings to instructors that teach in an unfavorable time. Moreover, students provide high ratings to instructors who teach major courses. There are also certain factors perceived by the instructors that affect the validity of student evaluation which are different from those mentioned in the review literature. These are: lack of skills (reading, writing, etc), lack of knowledge of teaching methodology, lack of knowledge of usefulness of student evaluation, receiving misinformation from previous students and prejudice about the instructors and courses.

With regard to applicability, student evaluation is not usually used for teacher and course development because students don't evaluate their instructors and the courses reliably. This means if students are not serious with their evaluations, their evaluation data will not be valid. In addition, student evaluation has not been practiced for the last two years and the feedback is rarely communicated to the instructors. Furthermore, instructors consider student evaluation useless because they do not get any reward: position or opportunity, etc for their good student evaluation achievements.

## **5.2 Recommendations**

### **To instructors**

1. Although student evaluations have some limitations instructors should conduct both formative and summative evaluations because they are useful for teacher and course development.
2. Instructors should understand that student evaluations are not conducted for the purpose of fault-finding.
3. Even though some students don't evaluate their instructors and courses reliably, there are some that provide reliable information. Therefore, instructors should not hesitate to conduct student evaluation.
4. Although student evaluation may sometimes seem to have more weaknesses than strengths, instructors should re-examine them so that they might find them right.
5. Instructors should understand weaknesses are part of teacher professional development.
6. Instructors should never perceive themselves 100% perfect. Therefore, they should be ready to learn from their weaknesses.
7. Instructors should orientate their students about the usefulness of student evaluation there by informing them to evaluate them and courses reliably.

### **To Educational Experts**

1. When student evaluation instrument are prepared, they should be rigorously validated and reliable in order to gather relevant data.
2. Student evaluation instruments should be prepared in a departmental approach because instruction and its improvements are discipline specific and instructors must closely identify with their department and are more likely to seek guidance and information within their department.

### **To Students**

Since the main purpose of student evaluation is to bring about quality education, students should evaluate their instructors and courses accurately and in a way which is free from any bias.



## Bibliography

- Abrami, P. C. & d'Apollonia, P.A. (1990). "Validity of student ratings of instruction: What we know and what we do not." *Journal of Educational Psychology* 82(2): 219-31.
- Abrami, P.C. (1998). **Improving Judgments about Teaching Effectiveness using Teacher Rating Forms**. Center for the study of learning and performance. Working Paper. Concordia University.
- Aleamoni, L. M. & Spencer, R. E. (1973). **The Illinois Course Evaluation Questionnaire**: A description of its Development and a Report of its Results. *Educational and psychological Measurement*, 33, 669-684.
- Aleamoni, L. M. (1976). **On the Invalidity of Students Ratings for Administrative Personnel Decision**. *Journal of Higher Education*, 47(5), 607-610.
- Aleamoni, L. M. (1999) "Student Rating Myths versus Research Facts from 1024 to 1998." *Journal of Personnel Evaluation in Education* 13(2): 153-166.
- Angelo, Thomas A., and K. Patricia Cross. (1993). **Classroom Assessment Techniques**: A Handbook for College Teachers, Second Edition. San Francisco, CA: Jossey-Bass Inc.
- Arreola, R. A. (2002). **Developing a Comprehensive Faculty Evaluation System**. Anker Publishing company, Inc. Bolton, MA.
- Arreola, R.A. (2000). **Developing a Comprehensive Faculty Evaluation System**: A Handbook for College Faculty and Administrators on Designing and Operating a Comprehensive Faculty Evaluation System. Bolton, MA: Anker Publishing Co., Inc.
- Bain, K. (2004). **What the Best College Teachers Do**. Cambridge, MA: Harvard Press.
- Bernard, H. (2002). **Using Evaluation to improve Teaching**: Address delivered at the University of Ottawa, November 2002.

- Borich, G. D (1988). **Effective Teaching Methods**. New York: Macmillan Publishing Company.
- Cashin, W. E. (1989). "Defining and Evaluating College Teaching." (IDEA Paper No. 21). Manhattan: Kansas State University, Center for Faculty Evaluation and Development.
- Centra, J. A. & Creech, F. R. (1976). **The Relationship between Student, Teacher, and Course Characteristic and Student Ratings of Teacher Effectiveness (Project Report 76.1)**. Princeton, NJ: Educational Teaching Service.
- Centra, J. A. (1979). **Determining Faculty Effectiveness**. San Francisco: Jossey-Bass.
- Centra, J. A. (1993). **Reflective Faculty Evaluation. Enhancing Teaching and Determining Faculty Effectiveness**. San Francisco: Jossey-Bass.
- Centra, J. A. (1993). **Reflective Faculty Evaluations**. San Francisco: Josey-Bass.
- Chambers, B. A. & Schmitt, N. (2002). "Inequity in the Performance Evaluation Process: How you rate me affects how I rate you." *Journal of Personnel Evaluation in Education* 16(2): 103-112.
- Chen, Y. & Hoshower, L.B. (2003). "Student Evaluation of Teaching Effectiveness: An Assessment of Student Perception and Motivation." *Assessment and Evaluation in Higher Education* 28(1): 71-88.
- Chickering, W. A. & Gamson, Z.F. (1987). **Seven Principles for Good Practice in Undergraduate Education**. *The Wingspread Journal*. AAHE Bulletin, March.
- Cohen, P. A. (1981). **Student Ratings of Instruction and Student Achievement: A Meta-Analysis of Multisession Validity Studies**. *Review of Educational Research*, 51, 281-309.
- Cohen, P.A. (1980). "Effectiveness of Student Rating Feedback for Improving College Instruction: A Meta-analysis of Findings." *Research in Higher Education*, 13, 321-341.

- Course Evaluation (2011). Accessed from [en.wikipedia.org/wiki/course\\_evaluation](http://en.wikipedia.org/wiki/course_evaluation) on 5/7/2011.
- Daniel Tefera, Demoz Admassu, Desalegn Chalchisa, Tamirie Andualem and Alemayehu Belachew (2002). **Introduction to Educational Psychology**. Addis Ababa: Addis Ababa University.
- Davis, B. G. (1990). **Tools for Teaching**. San Francisco, CA: Jossey-Bass.
- Diribsa D. Belete K. Tilahun F. Gentnet D. and Phadhan A. (1999). **Principles of Curriculum Inquiry**: Distance Education Material for In-Service Trainees: Addis Ababa, Addis Ababa University.
- Donald, J., & Saroyan, A. (1991). **Assessing the Quality of Teaching in Canadian University Education**: The Quality and Evaluation to Teaching. *Revista Iglu*, 1, 157-173.
- Doyle, K. O. (1983). **Evaluating Teaching**. Lexington, MA: Lexington Books.
- Dulz, T. and Lyons, P. (2011). **Student Evaluations**: Help or Hindrance? Accessed from [www.abesju.edu/proc2000/n038.pdf](http://www.abesju.edu/proc2000/n038.pdf).
- Duncan, D. (2005). **Clickers in the Classroom**. Upper Saddle River, NJ: Pearson Education.
- Dunegan, K. J. & Hrivnak, M. W. (2003). **Characteristics of Mindless Teaching Evaluations and the Moderating Effects of Image Compatibility**: *Journal of Management Education*, 27(3), 280-303.
- Evaluation (2011). Accessed from [www.en.wikipedia.org/wiki/Evaluation](http://www.en.wikipedia.org/wiki/Evaluation) on 9/13/2011 3:46 pm.
- Evaluation Trust (2011). Accessed from [www.evaluationtrust.org/evaluation/evaluate](http://www.evaluationtrust.org/evaluation/evaluate) on 9/13/2011 4:03 pm.
- Feldman, K. A. (1984). **Class Size and Students' Evaluations of College Teacher and courses**: A Closer Look. *Research in Higher Education*, 21, 45-116.
- Feldman, K. A. (1986). **The Perceived Instructional Effectiveness of College Teachers as Related to their Personality and Attitudinal Characteristics**: A Review and Synthesis. *Research in Higher Education*, 24, 139-213.

- Feldman, K.A. (1976). **Class Size and Students' Evaluations of college Teacher and courses:** A Closer Look. *Research in Higher Education*, 21, 45-116.
- Gallagher, T. G. (2000). "Embracing Student Evaluations of Teaching: A case study." *Teaching Sociology* 28, 140-147.
- Germain-Rutherford, A., (2003). **Student Evaluation Questionnaire Dispelling Misconceptions and Looking at the Literature.** Accessed from spiral. Univ-/yon 1.fr/8-based onneemedia/energy/sairi.asp?uti-id. On 4/7/2011 8:22 AM.
- Greenwald, A. G. & Gillmore, G.M. (1997b). "Grading Leniency is a Removable Contaminant of Student Ratings." *American Psychologist* 52(11): 1209-1217.
- Greenwald, A. G. (1997). "Validity Concerns and usefulness of Student Ratings of Instruction." *American Psychologist* 52(11): 1182-1186.
- Greenwald, A.G. & Gillmore, G.M. (1997a). "No pain, no gain? The Importance of Measuring Course Workload in Student Ratings of Instruction." *Journal of Educational Psychology* 89(4): 743-751.
- Gronlund N.E. (1985). **Measurement and Evaluation of Teaching.** New York: Macmillan.
- Hativa, N. (1995). "The Department-wide Approach to Improving Faculty Instruction in Higher Education: A Qualitative Evaluation." *Research in Higher Education* 36(4): 377-413.
- Hazari, S. I (2011). **Faculty perceptions of Student Evaluation of Instruction (SEI).** Accessed from [www.sunilhazari.com/education/documents/sei.htm](http://www.sunilhazari.com/education/documents/sei.htm) on 5/4/2011 7: 33 AM.
- Hobson, S.M. & Talbot, D. M. (2001). "Understanding Student Evaluations: What all faculty should know." *College Teaching* 49(1): 26-31.
- How do your students perceive "fairness" in your courses? "Online B... (2011). Accessed from [blog-ecollege.com/wordpress/?=5](http://blog-ecollege.com/wordpress/?=5) on 5/12/2011 10:41 AM.

- Howard, G. S & Maxwell, S. E. (1980). **The Correlation between Student Satisfaction and Grades: A Case of Mistaken Causation?** *Journal of Educational Psychology*, 72, 810-820.
- Howard, G. S & Maxwell, S.E. (1982). **Do Grades Contaminate Student Evaluations of Instruction?** *Research in Higher Education*, 16, 175-188.
- Hoyt, D.P. & Cashin, W.E. (1977). "Development of the IDEA System" (IDEA Tech.Rep.No.1). Manhattan: Kansas State University, Center for Faculty Evaluation and Development.
- Introduction to Evaluation (2011). Accessed from [www.socialresearchmethods.net/kb/interval.htm](http://www.socialresearchmethods.net/kb/interval.htm) on 9/13/2011 3:56 pm.
- Introduction to Perception. (2011). Accessed from [www. Sapdesignguidl.Org/rsources/optical illustrations/intro definition. htm/](http://www.Sapdesignguidl.Org/rsources/optical%20illustrations/intro%20definition.htm/)
- Johnson & Ryan (2000). "A comprehensive Approach to the Evaluation of College Teaching." *New Directions for Teaching and Learning*, no.83. Fall 2000.
- Kedzior, M. (2004). **Education Policy Brief (volume 15)**. Accessed from [www.rdc.udel.edu](http://www.rdc.udel.edu). On 8/7/ 2011. 6.10 AM.
- Kim, C., Damewood, E., & Hodge, N. (2000). **Professor Attitude: Its Effect on Teaching Evaluations**, *Journal of Management Education*, 24(4), 458-473.
- Kulik, J. A. (2001). "Student Ratings: Validity, utility and Controversy." *New Directions for Institutional Research* 109, 9-25.
- Land M. L. & Combs, A. (1981). **Teacher Clarity, Student Instructional Ratings, and Student Performance**. Paper Presented at the Annual Meeting of the American Educational Research Association, Los Angeles.
- Lersch, K. (2011). **Exploring the Beliefs Surrounding Student Evaluation of Instruction in Criminology/Criminal Justice Undergraduate Courses**. Accessed from [cba.ua.edu/handbook/](http://cba.ua.edu/handbook/) section 2/ student Evaluation of Teachers On 3/7/2011 4.45 AM.

- M. Hobson, S., and M. Talbot D. (2011). **Understanding student Evaluations**. Accessed from [jan.ucc.nau.edu/~s/m/AdjCI/Evaluation/instructor.htm](http://jan.ucc.nau.edu/~s/m/AdjCI/Evaluation/instructor.htm) on 11/7/2011 4:36 AM.
- Marsh, H. W. & Roche, L. (1993). "The use of Students' Evaluations and an Individually Structured Intervention to Enhance University Teaching Effectiveness." *American Educational Research Journal*, 30 (1), 217-251.
- Marsh, H. W. & Roche, L. A. (1997). "Making Students' Evaluation of Teaching Effectiveness: The Critical Issues of Validity, Bias, and utility." *American Psychologist* 52(11): 1887-1197.
- Marsh, H. W. (1980). **The Influence of Student, Course and Instructor Characteristics on Evaluations of University Teaching**. *American Educational Research Journal*, 17-219-237.
- Marsh, H. W. (1981). **The Use of Path Analysis to Estimate Teacher and Course Effect in Student Ratings of Instructional effectiveness**. *Applied Psychological Measurement* 6,47-60.
- Marsh, H. W. (1983). "Multidimensional Ratings of Teaching Effectiveness by Students from Different Academic Settings and their Relation to Student/Course/Instructor Characteristics." *Journal of Educational Psychology*, 75, 150-166.
- Marsh, H.W. (1987). **Students' Evaluations of University Teaching: Research Findings, Methodological Issues, and Directions for Future Research**. *International Journal of Educational Research*, 11(3), 253-388.
- McKeachie, W. (1997). "Student Ratings; The Validity of use." *American Psychologist* V.52, no. 11, 1218-1225.
- McKeachie, W. and Svinicki, M. 2006. **McKeachie's Teaching Tips**. Boston, MA: Houghton Mifflin Company.
- Mohanty, G., Gretes, J., Flowers, C., Algozzine, B., & Spooner, F. (2005). **Multi-method Evaluation of Instruction in Engineering Classes**. *Journal of personnel Evaluation in Education*, 18(2), 139-151.

- Mulu Nega and Daniel Tefera (2005). **Introduction to Measurement and Evaluation**. Addis Ababa: Addis Ababa University.
- Murray, H. (1984). "The Impact of Formative and Summative Evaluation of Teaching in North American universities." *Assessment and Evaluation in High Education*, 9 (2), 117-132.
- Ornstein A.C and Hunkins F. P (2004). **Curriculum Foundations, Principles and Issues**. New Jersey: Prentice Hall.
- Perception (2011). Accessed from <http://en.wikipedia.org/wiki/perception> on 2/7/2011 3:16 A.M
- Perception. (2011). Accessed from [http:// dictionary.reference.com/ browse/ perception](http://dictionary.reference.com/browse/perception) on 6/7/2011 8:05 AM.
- R. Crise, A. (2011). **What makes a good Instructor?** Accessed From [www.virtualfly casting.com/documents/Crise.pdf](http://www.virtualflycasting.com/documents/Crise.pdf) on 6/7/2011 7.23 AM.
- Rahman, K. (2006). **Learning from your Business Lectures**: Using Stepwise Regression to understand Courses Evaluation Data. *Journal of American Academy of Business*, 19(2), 272-279.
- Ramsdon, P. (2003). **Learning to Teach in Higher Education**. London and New York: Routledgefalmer.
- Instructor perception of student evaluation. (2011). Accessed from [www.olduncla.edu/publications/evalu of instruction/eval6](http://www.olduncla.edu/publications/evaluation_of_instruction/eval6) on 1/7/2011 9: 18 AM.
- Roberston, R. (2011). **Resource Teachers' perceptions of Student Evaluation Components of Special Education Class**. Accessed from [www. Saskschoolboards.ca/old/Research](http://www.Saskschoolboards.ca/old/Research) and Development on 2/7/2011 12.07 AM.
- Seldin, P. et al. (1999). **Changing Practices in Evaluation Teaching**: a Practical Guide to Improved Faculty Performance and Promotion/Tenure Decisions. Bolton, MA: Anke Publishing company, Inc.

- Sojka, J., Gupta, A.K. & Deeter-Schmetz, D. R. (2002). "Student and Faculty Perception of Student Evaluations of Teaching: A Study of Similarities and Differences." *College Teaching* 50(2): 44-49.
- Taba, H. (1962). **Curriculum Development: Theory and Practice**. New York: Harcourt, Brace and World, Inc.
- Tang, T. L. (1997). **Teaching Evaluation at a Public Institution of Higher Education: Factors Related to the Overall Teaching Effectiveness**. *Public Personnel Management*, 26(3), 379-391.
- The Meaning of Perception. (2011). Accessed from [www.worldtrans.org/Tp/Tp1/Tp1-9.HTML](http://www.worldtrans.org/Tp/Tp1/Tp1-9.HTML) on 2/7/2011 2: 18 AM.
- Theall, M. & Franklin, J. (2001). "Looking for Bias in all the Wrong Places: A Search of Truth or a Witch-hunt in Student Ratings of Instruction?" *New Directions for Institutional Research* 109: 45-56.
- TOEFL iBT-Sample Essay. (2011). Accessed from [www. Es/study guide. Com/essays/samplessay074.php](http://www.Es/studyguide.Com/essays/samplessay074.php) on 5/12/2011 10:59 AM
- Tyler R. (1949). **Basic Principles of Curriculum and Instruction**. Chicago: University of Chicago Press.
- Weinstock, R. (2010). **Quality Control in the Course Evaluation process: "Course Evaluations"**
- What is Perception? (2011). Accessed from <http://Psychology.com/od/sensationandperception/ss/perceptproc.htm> On 1/7/2011 9:28 AM.
- Young, S. Y. & Shaw, D.G. (1999). "Profiles of Effective College and University Teachers." *The Journal of Higher Education* 70(6):670-686.



**Appendix**  
**Addis Ababa University**  
**School of Graduate Studies**  
**College of Social Studies and Humanities**  
**Department of English TEFL**

**Questionnaire to be completed by English Language Instructors  
of AAU**

**Dear Instructors:**

The purpose of this questionnaire is to determine AAU's English instructors' perception of student evaluations for teacher and course development. The whole data and information obtained from this questionnaire will be utilized for academic issues only. Therefore, you are kindly requested to respond to the questionnaire genuinely and objectively.

**Personal Information**

Sex \_\_\_\_\_ Rank \_\_\_\_\_

**Notes**

1. You do not need to write your names.
2. Please follow the instructions provided for each part.

Thank you

Akaleold Damtew

MA/TEFL student

### I. Instructor Teaching Performance Evaluation

Choose one of the rating scales (agree, disagree, undecided) and put the mark “√” in the space provided.

No	Item	S. Agree	Agree	Undecided	Disagree	S.Disagree
1	Instructors with large class sizes receive low ratings					
2	Students provide high ratings to instructors who provide grades leniently.					
3	Instructors with high rank receive high ratings.					
4	Instructors who are attractive, entertaining and approachable receive high ratings.					
5	Instructors who teach courses with high work load receive low ratings					
6	Instructors who teach courses in an unfavorable time receive low ratings.					
7	Instructors who teach courses with less contact hours receive low ratings.					
8	Students provide high ratings to instructors who are in the same ethnic group they belong to					
9	Instructors who teach major courses receiver high ratings.					
10	Instructors who teach minor courses receive low ratings					
11	Instructors who teach common courses receive low ratings					

## II. Course evaluation

Choose one of the rating scales (Strongly agree, agree, undecided, disagree and strongly disagree) and put the mark "✓" in the space provided

No	Item	S. Agree	Agree	Undecided	Disagree	S.Disagree
1	Courses with large class sizes receive low ratings.					
2	Courses taught by instructors who provide grades leniently receive high ratings.					
3	Courses taught by instructors with high ranks receive high ratings					
4	Courses taught by attractive, entertaining and approachable instructors receive high ratings.					
5	Courses with high workload receive low ratings					
6	Courses taught in unfavorable time receive low ratings.					
7	Courses with less contact hours receive low ratings.					
8	Courses taught by instructors who are in the same ethnic group students belong to receive high ratings.					
9	Major courses receive high ratings.					
10	Minor courses receive low ratings					
11	Common courses receive low ratings					

### III. Student -Related Questions

Choose one of the rating scales (strongly agree, agree, undecided, disagree and strongly disagree) and put the mark "✓" in the space provided

No	Item	S. Agree	Agree	Undecided	Disagree	S.Disagree
1	Students are too immature to evaluate their teachers.					
2	Students are too immature to evaluate courses.					
3	Students lack expertise in the subject matter to evaluate their instructors.					
4	Students lack expertise in the subject matter to evaluate courses.					
5	Student evaluations are measures of teachers' popularity, not teacher effectiveness.					
6	Students are inaccurate in their judgment.					

### IV. Purposes of Students Evaluation

Choose one of the rating scales (strongly agree, agree, undecided, disagree and strongly disagree) and put the mark "✓" in the space provided

No	Item	S. Agree	Agree	Undecided	Disagree	S.Disagree
1	If student evaluations are rigorously validated and reliable, they are good sources of feedback.					
2	Student evaluations are important for teacher professional development.					
3	Student evaluations are important for course development.					
4	Student evaluation is a hindrance, not a help					

**V. What do you think are the advantages and disadvantages of student evaluation?**

Advantages

---

---

---

---

Disadvantages

---

---

---

---

**VI. Yes-No-question**

Put the mark "✓" in the space provided.

1. Do you think student evaluation is important?

Yes \_\_\_\_\_

No \_\_\_\_\_

2. Do you conduct student evaluation in your class?

Yes \_\_\_\_\_

No \_\_\_\_\_

If yes, for what purposes \_\_\_\_\_

---

---

3. Are you satisfied with the current student evaluations?

Yes \_\_\_\_\_

No \_\_\_\_\_

Why?

---

---

## **Interview**

**As per our appointment, I am going to interview you. Therefore, you are requested to answer the questions genuinely.**

- Q1. Do you think student evaluation is useful? Why?
- Q2. What are factors that affect the validity of student evaluation? How do they affect it?
- Q3. Do you think students evaluate instructors reliably? Why?
- Q4. How do you respond to student evaluation?
- Q5. What are the purposes of student evaluation?
- Q6. What do you think are the advantages and disadvantages of student evaluation?
- Q7. Do you feel comfortable with student evaluation of your course? Why?
- Q8. Have you been recently evaluated? If your answer is yes, have you been satisfied with it? Why?

**ration**

I, the undersigned, declare that this thesis is my original work, and has not been presented for a degree in this university and that all sources of material used for the thesis have been acknowledged.

Name: Akalewold Damtew

Signature: 

Place: Faculty of Language Studies Ababa University

Date of Submission: November 1, 2012