

**EXPLORING ENGLISH LANGUAGE ASSESSMENT IN ETHIOPIAN
UNIVERSITIES: ASSUMPTIONS, PRACTICES AND CHALLENGES**

Ph.D. DISSERTATION

DAGIM ENDALE GALATA

MARCH 2025

HARAMAYA UNIVERSITY, HARAMAYA

**Exploring English Language Assessment in Ethiopian Universities:
Assumptions, Practices and Challenges**

**A Dissertation Submitted to the School of Foreign Languages and Journalism,
Directorate for Postgraduate Programs
HARAMAYA UNIVERSITY**

**In Partial Fulfillment of the Requirements for the Degree of DOCTOR OF
PHILOSOPHY IN TEACHING ENGLISH AS A FOREIGN
LANGUAGE (TEFL)**

Dagim Endale Galata

March 2025

Haramaya University, Haramaya

STATEMENT OF THE AUTHOR

By my name and signature below, I declare and confirm that this Dissertation is my work. I have considered all ethical and technical principles of scholarship in the preparation, data collection, data analysis, and compilation of this Dissertation. All scholarly materials included in this dissertation were recognized through citations. This dissertation is submitted in partial fulfillment of the requirements for a Doctor of Philosophy degree in Teaching English as a Foreign Language (TEFL), under the School of Foreign Languages and Journalism at Haramaya University. The Dissertation is deposited in the Haramaya University Library and Information Services and is made available to borrowers under the rules and regulations of the library. I utterly declare that this dissertation has not been submitted to any other institution anywhere for the award of any academic degree, diploma, or certificate.

Brief quotations from this dissertation may be made without special permission provided that accurate and complete acknowledgments of the sources were made. Request for permission for extended quotations from or reproduction of this dissertation in whole or in part may be granted by the head of the School of Foreign Languages and Journalism when in his or her judgment, the proposed use of the material is in the interest of scholarship. In all other instances, however, permission must be obtained from the author of the dissertation.

Name: Dagim Endale Galata Signature: _____

Date: _____

Department: School of Foreign Languages and Journalism

BIOGRAPHICAL SKETCH

Mr. Dagim Endale Galata was born in Eastern Wallagga zone, Jimma Arjo District. He started his formal education at Andinet Primary School in Arjo and went to Jimma Arjo secondary school to finish high school.

Completing his high school education, he joined Jimma Teachers Training College to continue higher education program. Then, with very great distinction, he successfully graduated with diploma in teaching English in 2004. After his graduation, he was assigned to Gida Ayana district, where his career began, and taught at Gutin secondary school for one and half years, from 2005 to March 2006.

As he had great ambition to continue in his education, again he decided to join university and pursue his degree. Fortunately, he succeeded and continued his education at Jimma University as advance standing student and graduated in Bachelor of education in teaching English in 2009.

Since he graduated with great distinction, he got the chance to sit for proficiency exam prepared by MoE to select instructors for university. Successfully passing the exam, he was given another opportunity to pursue his master's degree in TEFL. Accordingly, in 2010, he joined Addis Ababa University and completed his study successfully in 2011.

As soon as he graduated in master's degree, the ministry assigned him to Wolkite University, where he taught for about three years, from 2012-2014. After that, he joined Wallagga University and has been teaching at the university for the past eleven years.

After his long journey, he again decided to go back to university and pursue his PhD degree, and accordingly, in 2018, he started his PhD study in TEFL at Haramaya University.

HARAMAYA UNIVERSITY

POSTGRADUATE PROGRAMS DIRECTORATE

I hereby certify that I have read and evaluated this dissertation entitled **Exploring English language assessment in Ethiopian universities: assumptions, practices and challenges**, prepared under my guidance by Dagim Endale and recommend it be submitted as fulfilling the dissertation requirement.

Adinew Tadesse Degago (PhD, Asso. Prof.)	_____	_____
Chairperson, Advisory Committee	Signature	Date
Abera Admasu Endashaw (PhD, Asst. Prof.)	_____	_____
Member, Advisory Committee	Signature	Date
Alemayehu Getachew Tsegaye (PhD, Asst. Prof)	_____	_____
Member, Advisory Committee	Signature	Date

As a member of the Board of Examiners of the PhD dissertation open defense examination, I certify that I have read and evaluated the dissertation prepared by Dagim Endale Galata, and examined the candidate. I recommend that the dissertation be accepted as fulfilling the dissertation requirements for the degree of ***Doctor of Philosophy in TEFL***

_____	_____	_____
Chairperson	Signature	Date
_____	_____	_____
Internal Examiner	Signature	Date
_____	_____	_____
External Examiner	Signature	Date

ACKNOWLEDGMENTS

First of all, I would like to express my deepest gratitude to my major advisor Dr. Adinew Tadesse for his continued guidance, invaluable support, devoted time, and valuable feedback during each stage of this dissertation. I am also profoundly indebted to Dr. Abera Admasu and Dr. Alemayehu Getachew, my co-advisors, for their every support throughout the development of this dissertation.

I would also like to extend my thanks to the EFL instructors and students who spent their valuable time participating in this study. Many thanks also go to Jimma University, Wallagga University, and Metu University for allowing me to carry out this study in their campuses and to Ministry of Education and Haramaya University that financially supported the study. I am highly indebted to my colleagues Mekuria Zewude, Ashenafi Bekele and Gezagn Alemu who spent time to read and give comments and accompanied me to the end of my research.

Further, I am grateful to my wife Chaltu Alemu who gave me unreserved support and encouragement and for everything she did to keep me focused and driven, as well as for her patience while I was taking much time from home for the dissertation work. My many heartfelt thanks also go to my parents, my mother, my brothers and sister for their emotional and material support, which played a major role in the success of this study.

Finally, I would like to take this opportunity to thank my friends, Ibsa Wakwaya, Habte Gebeyehu, Firew Zerihun, Tadele Zerihun and Solomon Moga, who have been supporting me unreservedly through internet, printing, copying, and even financially.

ACRONYMS AND ABBREVIATIONS

EEDRM	Ethiopian Education Development Road Map
EETA	Ethiopian Education and Training Authority
ESL	English as a Second Language
ETP	Education and Training Policy
FDRE	Federal Democratic Republic of Ethiopia
FL	Foreign Language
HE	Higher Education
HEIs	Higher Education Institutions
MoE	Ministry of Education
MoSHE	Ministry of Science and Higher Education
NRC	National Research Council
PD	Professional Development
SL	Second Language
TIIE	The Institute of International Education

TABLE OF CONTENTS

STATEMENT OF THE AUTHOR	II
BIOGRAPHICAL SKETCH	III
ACKNOWLEDGMENTS	V
ACRONYMS AND ABBREVIATIONS.....	VI
TABLE OF CONTENTS.....	VII
LIST OF TABLES	XI
LIST OF FIGURES	XII
ABSTRACT.....	XIII
CHAPTER ONE	1
INTRODUCTION	1
1.1. BACKGROUND OF THE STUDY	1
1.2. STATEMENT OF THE PROBLEM.....	4
1.3. OBJECTIVES OF THE STUDY	9
1.3.1. General Objective	9
1.3.2. Specific Objectives	9
1.4. SIGNIFICANCE OF THE STUDY	9
1.5. SCOPE OF THE STUDY	10
1.6. LIMITATIONS OF THE STUDY	10
1.7. OPERATIONAL DEFINITIONS OF BASIC TERMS.....	11
CHAPTER TWO	13
REVIEW OF RELATED LITERATURE	13
2.1. INTRODUCTION.....	13
2.2. THE ENGLISH LANGUAGE IMPROVEMENT COURSES.....	13
2.3. ASSESSMENT AND ITS REFORMS IN ETHIOPIAN HE.....	15
2.4. ASSESSMENT AND ITS ROLE IN ESL/EFL TEACHING	17
2.5. ASSESSMENT PRACTICES	19
2.5.1. Purposes of Assessment.....	20
2.5.2. Assessment Methods.....	23
2.5.3. Provisions of assessment feedback	29
2.5.4. Involving Students in the Assessment Process	34
2.6. ASSUMPTIONS ABOUT ASSESSMENT	35
2.6.1. Conceptualizing assumptions.....	35
2.6.2. Influences of Assumptions.....	38

2.6.3.	Exploring Assumptions.....	40
2.7.	ASSUMPTIONS UNDERLYING TRADITIONAL AND ALTERNATIVE ASSESSMENTS	43
2.7.1.	Assumptions of Traditional Assessment.....	44
2.7.2.	Assumptions of Alternative Assessment	46
2.8.	FACTORS THAT INFLUENCE EFL INSTRUCTORS' ASSESSMENT ASSUMPTIONS AND PRACTICES.....	49
2.9.	PREVIOUS STUDIES ON ENGLISH LANGUAGE ASSESSMENT	52
2.10.	THEORETICAL ORIENTATION.....	57
2.11.	CONCEPTUAL FRAMEWORK OF THE STUDY	59
CHAPTER THREE		62
THE RESEARCH METHODOLOGY		62
3.1.	THE RESEARCH PARADIGM	62
3.2.	RESEARCH DESIGN.....	63
3.3.	POPULATION OF THE STUDY	63
3.4.	SAMPLE AND SAMPLING TECHNIQUES.....	64
3.5.	DATA COLLECTION METHODS	67
3.5.1.	Interviews.....	67
3.5.2.	Questionnaire	69
3.5.3.	Documents	71
3.6.	VALIDITY AND RELIABILITY OF THE INSTRUMENTS	72
3.7.	PILOT STUDY.....	73
3.8.	PROCEDURES OF DATA COLLECTION.....	78
3.9.	DATA ANALYSIS PROCEDURES.....	79
3.9.1.	The Quantitative Data Analysis.....	79
3.9.2.	The Qualitative Data Analysis	80
3.9.3.	Integrating the Qualitative and Quantitative Results	82
3.10.	ETHICAL CONSIDERATIONS	83
CHAPTER FOUR.....		84
DATA ANALYSIS, FINDINGS AND DISCUSSIONS.....		84
4.1.	INTRODUCTION.....	84
4.2.	ANALYSIS AND RESULTS OF THE QUANTITATIVE DATA.....	84
4.2.1.	Characteristics of participants.....	84
4.2.2.	Assessment Practices	87
4.2.2.1.	Purposes of assessment	87
4.2.2.2.	Methods of assessing.....	90
4.2.2.3.	Provisions of assessment feedback	92
4.2.3.	Assessment assumptions.....	99
4.2.3.1.	Instructors' assessment assumptions.....	99

4.2.3.2. Students' assessment assumptions	102
4.2.4. Differences in practices and assumptions according to instructors' demographic characteristics	105
4.3. ANALYSIS AND FINDINGS OF THE QUALITATIVE DATA	113
4.3.1. Assessment practices	114
4.3.1.1. Purposes of assessment	114
4.3.1.2. Assessment methods	120
4.3.1.3. Provisions of assessment feedback	127
4.3.2. Assessment assumptions.....	131
4.3.2.1. Assumptions about content assessed/knowledge	132
4.3.2.2. Assumptions about values and judgments in assessment.....	137
4.3.2.3. Assumptions about power and control	142
4.3.3. Factors inhibiting instructors from interpreting their assessment assumptions into practice	145
4.4. DISCUSSIONS OF THE FINDINGS	149
4.4.1. Assessment practices	149
4.4.2. Assessment assumptions	155
4.4.3. Assessment practice and assumption comparisons according to instructors' demographic variables	159
4.4.4. Factors inhibiting interpreting assumptions into practices	161
CHAPTER FIVE	163
SUMMARY, CONCLUSIONS AND RECOMMENDATIONS.....	163
5.1. SUMMARY	163
5.2. CONCLUSIONS	164
5.3. RECOMMENDATIONS	166
REFERENCES	170
APPENDICES	182
APPENDIX A: QUESTIONNAIRE FOR INSTRUCTORS.....	182
APPENDIX B: QUESTIONNAIRE FOR STUDENTS.....	187
APPENDIX C: INTERVIEW GUIDE FOR INSTRUCTORS	191
APPENDIX D: INTERVIEW GUIDE FOR STUDENTS	193
APPENDIX E: SAMPLE TRANSCRIPTION OF EFL INSTRUCTOR INTERVIEW	194
APPENDIX F: SAMPLE TRANSCRIPTION OF THE INTERVIEW CONDUCTED WITH STUDENTS	202
APPENDIX G: SAMPLE COURSE OUTLINES USED	207
APPENDIX H: SAMPLE ASSIGNMENT CORRECTED	208
APPENDIX I: SAMPLE TEST/QUIZ CORRECTED	209
APPENDIX J: SAMPLE MAJOR EXAM CORRECTED	210
APPENDIX K: COURSE OUTLINES ANALYSIS	211

APPENDIX L: INSTRUCTORS’ ASSESSMENT PACKAGES ANALYSIS 212

APPENDIX M: BOX PLOT REPRESENTING INSTRUCTORS’ ASSESSMENT ASSUMPTION DATA
DISTRIBUTION OF NORMALITY 213

APPENDIX N: NORMALITY TEST FOR STUDENTS’ ASSESSMENT ASSUMPTION DATA USING BOX
PLOT 214

LIST OF TABLES

Table 1. Different classifications of classroom assessment methods	24
Table 2 Comparison of philosophical beliefs and theoretical assumptions of alternative and traditional assessment	43
Table 3 Samples and sampling of the participants of the survey questionnaire	65
Table 4. Reliability Coefficient Scores on assessment practice and assumption	76
Table 5. Instructors' demographic characteristics	85
Table 6. Students' characteristics	86
Table 7. Frequency of use of assessment for learning	87
Table 8. Frequency of use of assessment of learning	88
Table 9. Frequency of use of assessment as learning	89
Table 10. Frequency of use of assessment methods	90
Table 11. Sources and modes of assessment feedback	92
Table 12. Verification assessment feedback	93
Table 13. Directive assessment feedback	94
Table 14. Scaffolding assessment feedback	95
Table 15. Descriptive statistics of clustered classroom assessment practices	96
Table 16. Descriptive statistics of instructors' assessment assumptions	99
Table 17. Mean, standard deviation, and one sample t-test results of instructors' clustered assumptions	101
Table 18. Descriptive statistics of students' assessment assumptions	102
Table 19. Mean, STD, and one sample t-test results of students' clustered assumptions	104
Table 20. T-test results of instructors' assessment practice comparison upon academic qualification	105
Table 21. T-test results of instructors' assumptions comparison upon academic qualification .	107
Table 22. ANOVA results of instructors' assessment practice comparison based on experience	109
Table 23. Tukey HSD Post-hoc Multiple Comparisons of Group Means	111
Table 24. ANOVA results of instructors' assumptions comparison based on teaching experience	112
Table 25. Tukey HSD Post-hoc Multiple Comparisons of Group Means	113

LIST OF FIGURES

Figure 1. Traditional assessment pyramid	23
Figure 2. Reconfigured assessment pyramid	23
Figure 3. Conceptual framework of the study	61
Figure 4. A framework of data collection, analysis and integration for the study.....	82

ABSTRACT

This study aimed to explore assessment assumptions, practices and challenges in the enactment of communicative English language skills course in Ethiopian universities. To this end, a mixed methods research design was employed, and data were collected from EFL instructors and first-year undergraduate students through questionnaires, interviews and documents. The quantitative data were analyzed through descriptive and inferential statistics whereas thematic and content analysis approaches were applied for the qualitative data. The findings revealed that EFL instructors conducted assessments mainly for accountability purposes and relied on some particular conventional assessment methods to realize this. In addition, although instructors used both written and oral feedback modes frequently, they were found to be the sole source of feedback. These feedbacks often focused on verifying and directly providing correct answers than on scaffolding learners to generate by themselves. The findings also showed that students were highly traditionalists in their assumptions. On the other hand, although EFL instructors seem to hold mixed assumptions (progressive and traditional), most of their assumptions were those underlying traditional assessment approaches. The few alternativist assumptions held by EFL instructors often did not translate into practice due to various external factors such as the established assessment system, students' language proficiency and overall competence, large class size, and workload. The study suggests the importance of ensuring that the assessment practices in the target course are aligned with the current assessment developments so that students' English language learning is more supported. Changing EFL instructors' and students' assumptions about assessment and working on barriers that hinder interpreting assumptions into practice are also the other implications of the study.

Article Published

Title: University teachers' underlying assumptions about assessment in English as a foreign language context in Ethiopia

Journal: Cogent Education

Publisher: Taylor & Francis

Date of publication: 05 April 2024

VOL. 11, NO. 1, 2335748

Link: <https://doi.org/10.1080/2331186X.2024.2335748>

CHAPTER ONE

INTRODUCTION

1.1. Background of the Study

Relationship between instruction and assessment is integral to effective teaching and learning. Instruction informs assessment by outlining the expected outcomes, and assessment guides instruction by providing students' understanding, strengths and areas needing improvement. So, alignment between the two creates supportive environment that can enhance students' English language skills effectively (Alderson, 2005; Phuong and Nguyen, 2019).

Hence, particularly, classroom assessment has been receiving great attention, as it has a significant influence on learning. Classroom assessment is assessment accurately in the hands of instructors (Black and Wiliam, 1998), in which both instructors and students engage to gather information on teaching and learning used for different purposes. However, classroom assessment is only useful when used properly. Evidences on assessment purpose, method and feedback also indicate that assessment may support or weaken students' motivation and performance depending on the way it is designed, implemented, and used. Accordingly, current reforms encourage, for example, using assessment both for accountability purpose and to facilitate teaching and learning, alternative assessment methods, and feedbacks that are engaging and focus on supporting learning (Shute, 2008; Brown and Abeywickrama, 2010; Guo, 2020).

Recent policies issued by Ethiopian Ministry of Education also outline expectations for classroom assessment across HEIs by proposing various novel assessment approaches. The key concern lies in integrating assessment with instruction so that learning is more enhanced. Hence, a stronger interest now lies in gathering data on students not only to measure their achievement of the required competencies, but also to facilitate instruction and learning through the use of diagnostic, formative, and summative assessments (MoSHE, 2019; EETA, 2023).

As such, employing a variety of assessment methods has become compulsory. For instance, the new road-map, commenting to detach from paper-based-only assessments, envisioned alternative assessments as the way forward (MoE, 2018). EFL instructors are expected to use a variety of methods such as presentation, written assignments, interviews, portfolio writing, speech delivery,

etc. (MoE, 2013; MoSHE, 2019). This does not mean, however, that formal methods should be abandoned when assessing learning outcomes. In line with MoSHE (2019), at least one quiz, one test, a mid and a final examination are conducted to evaluate and grade students' achievement. Further, the use of assessment feedback to enhance teaching and learning has been vital in the policy documents. Particularly, feedback practices that are timely, engaging and that focus on supporting learning are emphasized (MoE, 2013; MoSHE, 2019; EETA, 2023).

Shortly, the currently recommended assessment practices hold considerable promise. According to Price, Light and Pierson (2014), they are essential to teaching and learning at least through four major dimensions. They enable modifying instruction, assessing a broader range of skills and abilities, offering students feedback about their learning and on how they can improve, and giving them new roles in assessment process. This is also supported by empirical evidences that suggest a weak to strong effect on student achievement (Black & Wiliam, 1998; Hattie & Timperley, 2007; Nibret, 2013; Wafubwa & Csíkos, 2022). Hence, successful implementation of such assessments in HEIs can enable equipping EL learners with the necessary EL skills they need for their academic as well as non-academic purposes.

However, despite such potentials, there has been a growing concern regarding effective implementation of these assessment reforms (Buhagiar, 2007; Carless, 2015; Mitana, 2018). This concern doubles in the developing countries where such assessment is not even part of the practitioners' conversations (Price et al., 2014). The limited studies in Ethiopia HE EFL context also depict this reality (e.g., Dagne, 2009; Mebea, 2008; Zelalem & Emily, 2018).

Literature sources indicate that usually reforms in assessment remain unimplemented effectively, and this is due to multifaceted factors in relation to the practitioners' cognitive readiness, dispositions and available resources. Stabile (2014), in particular, attributed the reason to the sole emphasis given to developing the new standards neglecting the practitioners' assumptions. This suggests the role the practitioners' assumption has in implementing quality assessments. In this regard, Wellberg and Evans also stated, "the success of any educational policy meant to improve instruction is heavily dependent upon the assumptions that teachers have" (2022: 3).

Thus, assumptions, which are taken for granted as a basis of argument or action and which constitute the basis of the judgments the practitioners make as part of the assessment process, are

important factors in implementing quality assessments (Delandshere, 2001; Brookfield, 2013; Atkin, 2017). So, the observed failure to implement the assessment reforms can be related to the practitioners' assumptions, as every assessment decision and action they take are reliant on their assumptions. That is, assumptions are determinant factors that can either encourage or deter the practitioners from engaging in real transformation towards the innovative assessments.

Holzweiss, Bustamante and Fuller (2016) stressed that successful shifts towards positive cultures of assessment can be gained by paying attention to practitioners' assumptions. Since changes in assessment behavior often follow changes in its assumptions, clarifying their assumptions has huge implication for PD. It enables unveiling those interfere with the innovative practices for examinations upon which changes in assumptions can be realized, resulting in a positive culture of assessment. Hence, in addition to the assessment practices, the underlying assumption has also become the focus of the present study. Leung (2005) and Taylor (2009) also suggest researchers to focus not only on the practices, but also to delve into the mental world of instructors to better understand the theories they use to define and conduct classroom assessment activities.

This study particularly emphasized on assessment specific to the course communicative English skills. As there has been high need in Ethiopia to improve university students' EL, teaching the language has become essential part of university education, and this compulsory course is delivered to all undergraduate students to alleviate the problem in students English. In this regard, the role of assessment is paramount, for it has direct influence on students' EL ability. In this course, assessment is a valuable tool for providing reliable information about its teaching and learning for instructors and students to make decisions related to aspects of learners' EL ability (Alderson, 2005; Motuma & Chacka, 2023).

This suggests significance of assessment study in the target course. Researchers also note the need for such study to have a clearer picture of how far assessments implemented meet the learning needs of EL learners (Leung, 2005; Motuma & Chacka, 2023). Thus, more needs to be clarified about assessment in communicative English skills course, what the current study intended to do to make sense of how assessments are supporting learners' EL improvement.

Obviously, effective implementation of quality assessment is highly reliant on instructors given their key role as frontline curriculum implementers. However, neglecting the importance of

students' views in such studies is capital mistake, as they are also the front-liners on and by whom assessments are conducted. Increasingly, policies are also holding them accountable for classroom assessment (Lund, 2008) since their involvement enhances achievement by motivating them to feel responsible for their own learning and improvement (Tan, 2012; Wellberg & Evans, 2022). This, therefore, brings forward the significance of the current study's focus on EFL instructors and students to investigate assessment practices and the underlying assumptions in the enactment of communicative English language skills course in Ethiopia higher education.

1.2. Statement of the Problem

In Ethiopia, there have been exertions to reform the education system to meet the needs of the workforce regarding current century skills. Assessment was a cornerstone of the movement. Consequently, emphasis has shifted to the formative aspects of assessment rather than summative, constructive feedback rather than just marks and grades, using various methods than depending on a single method, and student-led assessments rather than relying on assessment by instructors (MoE, 2012; MoE, 2013; MoE, 2018; MoSHE, 2019; Motuma & Chaka, 2023). These shifts lie in the basic assumption that the quality of EL learning can be enhanced by improving the quality of its assessment. Therefore, effective implementation of the assessment reforms in the target course can help realize the intended change in students' English.

However, there is high concern regarding effective implementation of these assessment reforms. The researcher's observation from his teaching experience tend to show that EL assessments are far from the recommended practices, and the traditional approaches are still dominating university assessments. It seems that assessments are still operating much like they were decades ago, being detached from instruction. Unfortunately, enhancing students' learning is hardly possible through the old approaches. For example, the traditional pen and paper assessments are claimed to encourage surface than deep approach to learning that drives students to run for grades rather than learning English for communicative purposes (Hughes, 2003; Brown & Abeywickrama, 2010). Black and Wiliam (1998), based on meta-analysis of over 250 studies, also showed that formative feedback practices enhance teaching effectiveness and produce student academic achievement gains ranging from 0.4 to 0.7 magnitude of effect size. Thus, perhaps, assessment is a factor for the problem in Ethiopian students' EL proficiency.

Although there has been growing interest in assessment to enhance quality of EL education and the recent changes made on the area carries considerable promise in this regard, there are still concerns about students' ability in using the language. Studies also showed as many students, even at their final year, face difficulty to attend their study in the target language (TIIE, 2012; Berhane, 2019). This suggests how deep-rooted the problem is, as it transited and being reflected in students' EL ability that in turn influenced their academic study.

While assessment continues to be a hot topic of discussion in relation to its decisive role in enhancing students' EL, the availability of adequate data on EFL instructors' assessment practices is critical. This helps to understand what is needed, either at a policy or university level or both, to help them adapt their practices in ways that utilize best practices. Additionally, knowledge of EFL instructors' and students' assessment assumptions is essential to know whether they are compatible or not with current theoretical approaches and with the expectations of recent reforms on classroom assessment. This enables, when necessary, to help them develop availing assumptions in current foreign language assessment.

Accordingly, there have been efforts by local researchers to study this prevailing educational phenomenon. For example, Mebea (2008) investigated the practice and attitude of continuous oral assessment in EFL classes. According to data from eleven instructors and 120 students, both sample groups had positive attitude for oral assessment though that was not interpreted into practice by the instructors. Absence of assessment policy and inadequate preparation were identified to hinder effective implementation of continuous oral assessment.

Similarly, Dagne (2009) studied the practice of continuous assessment by EFL instructors at Jimma Teachers' College. Data were gathered from 20 instructors and 141 2nd and 3rd year students through questionnaire, interview and classroom observation. The results indicated that more than 70% of instructors did not use continuous assessment in EFL classrooms. This was attributed to lack of adequate training, large class size and workload. Abiy (2013), who investigated high school EFL teachers' and students' perceptions, attitude, and practice of continuous assessment, similarly came to the conclusion that continuous assessment was not properly practiced by the teachers.

Addressing assessment techniques, purposes and feedback, Geberew (2014) investigated formative assessment practices of EFL teachers. The findings from questionnaire and interview-based data showed that most of the teachers used assessment of learning. Teachers' limited use of assessment feedback to adjust instruction was also reported. It was indicated that formative assessment was not well practiced, and diligent effort is required to help EFL teachers use assessments properly. Zelalem and Emily (2018), who assessed formative assessment practices and beliefs of EFL instructors at Debre Markos University in writing skill context, also had similar findings. Data were collected from 25 instructors through questionnaire and interview. The study revealed instructors' positive beliefs on the importance of formative assessment although they could not practice assessment for formative purposes as the findings from interviews indicated.

More recently, Motuma and Chaka (2023) examined the implementation of alternative assessment in communicative English skills course at three universities. Data gathered from EFL instructors and students revealed dominance of instructor-centered assessment strategy, which even lacked progressiveness and relevance. Most of the instructors also preferred traditional assessment methods to alternative assessment methods and found the former more convenient to use in the target course than the later. Similar results were revealed by Nibret (2013) in his study of the effect of modifying EFL teacher's assessment method on students' integrated approach to learning English. The quasi-experimental study revealed the teacher's dominant use of traditional testing methods although the training intervention helped him to employ alternative assessment methods, which in turn positively affected students' approach to learning.

Through these local studies, portions of the current status of English language assessment could be gleaned. However, some gaps were observed that might have limited our better understanding of the phenomena. For example, there was methodological gap in well addressing the research problem. Most of the studies depicted the evidences in assessment practices by relying on data from instructors, neglecting the views of other frontliners such as students and data from the relevant documents. Although Mebea (2008) and Motuma & Chaka (2023) emphasized on both instructors and students, they failed to incorporate data from the relevant documents, which however, could have enabled a more complete understanding on the practice of assessment. Hence, the present study fills the gap by collecting data from EFL instructors, students and from

existing relevant documents on the area in order to gain a better understanding about the phenomena under study. Such sources of data provide the conditions for systematic data collection on assessment policy implementation (Lund, 2008; Kezar, 2012).

In addition, although certain assessment components were repeatedly addressed in the studies, the purposes EFL instructors' feedbacks focus on, the other essential element of assessment feedback practice, received no attention. However, since instructor feedback is a key factor in improving students' English language (Hattie & Timperley, 2007; Gan et al., 2021), what functions the provided feedback serve worth an investigation (Shute, 2008; Brooks et al., 2019). To this end, this study filled a gap in existing studies by addressing the functions the feedbacks provided by EFL instructors focuses on, a vital issue but which remained a gap.

Holroyd (2003) argues that there would be subject-based differences in assessment practice and thinking. Cumming (2001) study, which revealed that EFL instructors' assessment beliefs and practices varied in line with particular course they teach, also shows similar reality. This suggests desirability of subject/course-specific assessment studies. Despite the fact, none of the studies have studied assessment specific to communicative English language skills course, except the one by Motuma and Chaka (2023), which was even limited to assessment methods. So, more evidence is needed about assessment practices in the target course so as to come to a firmer conclusion. Thus, by delimiting the study of assessment specific to communicative English skills course and by also including other aspects of assessment, the current study is initiated to fill the existing gap.

On top of these, as far as the researcher's knowledge is concerned, studies are lacking in Ethiopia regarding English language assessment assumptions at tertiary level despite the growing recognition that assumptions greatly influence the success of assessment implementation. As a result, the assumptions underlying assessment practices in ELT and in particular in communicative English skills course are not yet clear, and our knowledge of this area is very limited. The area also seems to gain little attention at global level, as the target of most studies in this area was on general education and/or lower educational levels (e.g., Delandshere, 2001; Klenowski & Willis, 2011; Mitana, 2018), remaining a knowledge gap regarding assessment assumptions in HE EFL context. Overall, following a dearth of studies on the area, there has

been a growing world-wide call for more studies (NRC, 2001; Klenowski & Willis, 2011; Holzweiss et al., 2016; Mitana, 2018; Dawson et al., 2019).

This gap is considerably problematic given the power of assumption in determining what is assessed, how it is assessed, how learners are treated, and instructors' positionality. The situation puts the ongoing efforts to redress the assessment problems in Ethiopian universities under question. To foster meaningful change in assessment, EFL practitioners' assessment assumptions should be brought out of the closet and examined in light of current ideas of assessment and learning, either to accept as compatible or reject as unjustified. Hence, this study is initiated to shed light on assessment assumptions and fill the gap in the literature.

Generally, the present study aimed at exploring English language assessment emphasizing on assumptions, practices and challenges in the enactment of communicative English language skills course in Ethiopian universities by answering the following research questions:

- To what extent EFL instructors' assessment practices in communicative English language skills course support students' English language learning?
 - For what purpose/s do EFL instructors primarily conduct assessments in communicative English skills course?
 - To what extent the instructors utilize alternative assessment methods in the target course?
 - How do EFL instructors provide assessment feedbacks to their students in the target course?
- What are the assumptions underlying instructors' and students' assessment practices and views in communicative English language skills course?
- Are there any significant differences regarding assessment practices and assumptions with respect to instructors' characteristics/profile (sex, educational level, teaching experience)?
- If any, what are the factors that influence EFL instructors from implementing their assessment assumptions into practice?

1.3. Objectives of the Study

1.3.1. General Objective

The major objective of this study was to investigate EFL instructors' and students' assessment assumptions, practices, and challenges in the enactment of Communicative English Language Skills course in Ethiopian universities.

1.3.2. Specific Objectives

Specifically, the study attempted to:

- determine the purpose/s for which EFL instructors primary conduct assessments in communicative English skills course;
- examine the extent to which EFL instructors utilize alternative assessment methods in the target course;
- assess how the instructors provide assessment feedback to their students in the target course;
- explore the assumptions underlying EFL instructors' and students' assessment practices and views in communicative English skills course;
- examine if there are any significant differences regarding assessment practices and assumptions with respect to the instructors' demographic characteristics or profile;
- identify factors that might hamper EFL instructors from interpreting their assessment assumptions into practice;

1.4. Significance of the Study

First, insight from this study may contribute to the on-going efforts to improve the teaching of communicative English skills course in general. Given lack of research on assessment in this particular context combined with the increasing emphasis on classroom assessment to enhance students' EL learning, the researcher believed that conducting study on the area does mean contributing to the efforts being undertaken in Ethiopia in response to the deteriorating English language education.

Second, investigation into instructors' assessment practices and assumptions would be helpful for EFL instructors in increasing their awareness and knowledge of assessment. As the findings

of the study sheds light on the areas of the instructors' weaknesses, which echo areas of needs, the outcomes of this study are believed to influence them to elevate their assessment literacy by taking part in different assessment professional networks.

Third, findings from this study will also be important for those who design and deliver professional development programs by providing rich information regarding assessment in communicative English skill course. By clearly indicating where the instructors currently are in their assessment practices and the assumptions they value, the findings allow PD designers and providers for a more effective approach to close any gaps in instructors' pedagogical knowledge, which, in turn, can close gaps in student knowledge.

Finally, the information from this study can also be used as input in developing assessment courses specific to ELT than using the 'one-size-fits-all approach' in which the same course is delivered regardless of fields of study. That is, the findings may serve as an important basis for corresponding department or for the concerned bodies at national level to formulate course on assessment specific to ELT. Finally, findings from this study will also be a vital reference for forth coming researchers on the area.

1.5. Scope of the Study

Since the current study focuses on investigating assessment practices and assumptions underlying in the context of communicative English language skills course, conceptually it is delimited to classroom assessment by following some suggested key assessment process routes including purpose, methods, feedback and communication (Rea-Dickins, 2001; Hill & McNamara, 2012). Accordingly, high-stake or standardized assessments are out of the focus of the present study. The scope of the study was also delimited to EFL instructors who are teaching communicative English language skills course and undergraduate first year students who are taking the course, for they are legitimate for the course under study. Furthermore, the study is delimited to focus on the assessment practices and assumptions as they occur in public universities located in western Ethiopia. As such, universities located in other areas of the country and non-public universities are out of the scope of this study.

1.6. Limitations of the study

This study is not without limitations, and that should be acknowledged. Because of some academic schedule disrupts and overlaps in the universities, the semester when this study was conducted was shorter than regular semester. As a result, interview-based elicitation technique was used to elicit participants' assumptions about assessment. However, additional insights regarding participants' assumptions about assessment, that could not be revealed by the elicitation technique applied in this study, might have been depicted if alternative techniques such as stimulated recall interviews following video recordings and repertory grid techniques were also used.

Another limitation of this dissertation is related to language used in students' questionnaire. Based on the assumption that students at university level are able to understand and respond to texts written at phrasal and sentence levels in English language, English version of questionnaire was used for students. However, examining the pattern of a few students' responses to the survey questions led me to suspect their clear understanding of the items, which in turn thought to raise concern on validity of the data. Had questionnaires translated into students' native language been used, we would have more confidence on the data.

Finally, as the study focused on instructors teaching communicative English skills course in three universities (Jimma, Wallagga, and Metu) located in western Ethiopia, representativeness of the study sample could be limited. Particularly, in relation with elimination of Dambi Dollo University from the sample universities because of security issue following serious political unrest in the area, the potential number of instructor participants was decreased, which has significantly limited representativeness of the study sample.

1.7. Operational Definitions of Basic Terms

Assessment: in this study, the term is defined as the process of assessing students' knowledge and skills through both standardized and non-standardized tools for the purposes of both placement and making inferences about learning.

Classroom assessment: it is used to denote assessments which are different from standardized, high stake and commercial, but which are accurately on the hands of instructors (Black and Wiliam, 1998), in which instructors as well as students use to gather information on the qualities

of students' work and the use of that information by instructors and students for the purposes of teaching, learning, evaluating and reporting course achievement.

Assessment practice: understood as any activity or task instructors and students carry out in communicative English language skills course. In particular, it refers to activities conducted in relation to the why (purpose) and how (method and feedback) of assessments in the target course.

Assumptions: understood as propositions or implicit/explicit statements EFL practitioners make about assessment that are taken for granted as a basis of argument or action. They are conceptualized not as what instructors and students say they believe about assessment but what their justifications behind their assessment behavior (what they say and do) reveal about what they believe about assessment. Shaped by past experience and learning, they are upon which practitioners' assessment theory is constructed, and are fundamental to assessment reasoning, and subsequently to behavior and action.

Alternative assessment: is understood here as any assessment that is different from the traditional assessment paradigm. It is a process that is firmly integrated with instruction and primarily positioned for promoting learning.

CHAPTER TWO

REVIEW OF RELATED LITERATURE

2.1. Introduction

This part begins with discussing the role of university-based English language improvement courses and the major changes they have undergone. Then, it presents the assessment-focused discussions of education reform initiatives that took place on Ethiopian higher education in the past couple of decades. Assessment approaches, methods, concepts and other assessment issues addressed in the documents and which have been guiding the practices will be highlighted so that current assessment orientations and expected practices will be made clear. Then, by presenting and discussing how the current approach to assessment emerged, the chapter explicates the role of classroom assessment in ELT. It also addresses how, why and what type of assessments need to be prioritized in the current Ethiopian universities in its discussions of assessment practices by encapsulating the major stages. Succeeding the conceptualization and influence of assumptions, the literature review also elucidates assumptions underlying educational assessment. Furthermore, the section presents review of studies conducted on the area. Finally, the theoretical and conceptual frameworks of the study will be illustrated.

2.2. The English Language Improvement Courses

Following its role in academics, developing students' English language has become essential part of university education (Alderson, 2005; Manning, 2013). Similarly, the use of the language as a media of instruction in Ethiopian universities has necessitated developing students' English proficiency. Accordingly, universities have been offering English skills improvement courses, and all undergraduate students, irrespective of their fields of study, take these compulsory courses. The purpose of teaching the courses at this level is thus to help students acquire proficiency in using the language mainly for their academic purposes.

Historically, the courses have undergone several changes. For instance, "Freshman English 101A" and "Freshman English 101B" support courses were used to be offered to first year university students before 1996 (Geremew, 1999). The courses, which were criticized for their limited emphasis to language structure and decontextualized nature of teaching the skills (Geremew, 1999), were then replaced by two 'College English' courses, with the introduction of

‘Sophomore English’. The revision was, thus, to shift the emphases of English language teaching and learning from language forms to its communicative aspect. However, the revision was not fruitful, as students’ English remained poor. A survey conducted in 2007 also revealed as there were serious English language problems among graduates (MoE, 2007), which warranted revising the courses.

Thus, in line with the shift in the curriculum from knowledge to competence-based approach, the language improvement courses were revised for the other term. This time, one ‘Communicative English Skills’ course and one ‘Basic Writing Skills’ were introduced under ‘Basic English Skills’ module. The courses were designed with learner-centered approach and the principles of communicative language teaching (MoE, 2013; Bedilu, 2020). Progressive assessment to support students’ learning has also become center of attention (MoE, 2013). It was hoped that the new courses would help alleviate the challenges identified in the former one’s.

According to recent studies, however, students continue to experience difficulty in using English for academic and non-academic functions (TIIE, 2012; Mesfin, 2016; Berhane, 2019). Perhaps, in response to these disconcerting findings, the service courses were reviewed and revised once more in 2019. This time, ‘communicative English skills I’ and ‘communicative English skills II’ courses have been introduced, and the course in writing skills is eliminated from the curriculum. Although not clearly indicated, disregarding the writing skill course seem to base on the assumptions of the equivalent function the language skills serve and holistic view of language. According to some anecdotal evidences from the course instructors, however, there are considerable concerns regarding the quality of these service courses, from their design to contents and approaches. This makes one to question how far the new courses are helping in developing students’ English language proficiency and alleviate the challenges in the area.

The role of assessment in enhancing learners’ English language learning necessitates using quality assessments in the English language improvement courses. That is, if the assessments that contemporary literature recommends are meaningfully executed in the teaching of the courses, there will be sound instruction and learning. As a result, enhanced learners’ English language communication skills is expected. This amplifies the need to examine the existing assessment system in the target courses and have a clearer picture of it. That is also why the

present study intended to do this in order to understand how assessments are contributing in developing learners' English language communication skills.

2.3. Assessment and its reforms in Ethiopian HE

A growing body of research suggests that assessment can greatly enhance teaching and learning if it is well integrated with instruction (Tan, 2012; Price et al., 2014; Al-Bahlani & Ecke, 2023). Perhaps, this notion has influenced Ethiopia education curricula, as integrating assessment to the teaching learning process has been a concern starting from the promulgation of the 1994 ETP, as also argued by (Dagne, 2009). To be specific, the education policy initiatives that took place on HE in the past couple of decades have made assessment one focus area of the movement.

Usually new tenets and assumptions in learning that gain prominence in the time influences policy movements in education. Similarly, in promulgating the 1994 education and training policy, constructivist approach appears to be introduced in the education system, which had been under behaviorist theory. The policy has come with some shifts including in the assessment of learners. Criticizing the assessment approach in the past, in which students were solely tested based on summative assessments, the policy had proposed the use of continuous assessment (MoE, 2002). "Continuous assessment in academic and practical subjects, including aptitude tests will be conducted to ascertain the formation of all-round profile of students at all levels" (MoE, 1994: 18). Continuous assessment became the focus of the 1994 ETP given the limitations of the traditional assessment approaches in evaluating students' performance.

Higher Education Proclamation, which was established basically in response to the demand to lead higher educations by legal directions and guidance particularly on quality and relevance (FDRE, 2009), also came with some insights. Students' assessment was among other issues detailed in the proclamation No. 650/2009. It primarily focuses on assessment, evaluation and grading principles, ethics and fairness. For instance, article 21 suggests aligning classroom assessment with instruction. The document also dictated some modes of assessment including both the formal and informal assessment modes. It was indicated that assessment results, besides other educational quality enhancement strategies, are used as evidences for institutional quality enhancements (FDRE, 2009: 4988-89).

Under article 41.8, the document also promises that in-service trainings on assessment theories and skills shall be provided to instructors. This may suggest the need to continuously equip instructors with the assumptions underlying current assessment approaches since according to Xu & Brown (2016), they must understand the nature and requirements of their assessment tasks in order to support enhanced language learning. Otherwise, if proper support lacks on the new assessment developments, equally, it is unlikely to expect its effective implementation from the instructors. Perhaps, what remains a challenge for educators in Ethiopia is lack of professional support on the new assessment developments. Furthermore, it is also indicated that mastery and successful application of assessment skills are critical pre-requisite to continue as instructor and for promotion in HEIs (FDRE, 2009).

Later, the Ministry of Education has issued the “Harmonized Academic Policy of Ethiopian Public Higher Education Institutions” in 2012. Its main purpose was harmonizing the academic policies to maintain uniform policy across institutions (MoE, 2012). Students’ assessment was one of the six broad themes the policy has emphasized. Article 56, for instance, recommends continuous assessment, which comprises of 50%, and allots the remaining 50% to the final exam. It also suggests the use of alternative assessments such as tests, assignments, reports, presentations, etc through written, practical or oral approaches (MoE, 2012: 87).

Soon, the department of EL under the MoE has designed a new “Nationally Harmonized Modular Curriculum for Undergraduate English Language Program” (MoE, 2013). The knowledge-based approach curriculum, which is indicated to have limitations in satisfying the demand of the market, had led curriculum revision. Thus, the curriculum is designed focusing the competencies the learners in the program need to acquire. Noticing that students shall be assessed both on continuous and final examination basis, the document specifies portfolio writing, classroom presentation, quiz, written assignments, speech delivery, etc. as alternative methods through which EFL instructors cover the 50%/60% continuous assessment (MoE, 2013). Instructors are also suggested to use authentic and performance assessment approaches.

The two education reforms, “The Ethiopian education development roadmap” and “Ethiopian Higher education curriculum framework”, recently introduced have also come with some insights about assessment, which are to guide the assessment practices. The Ethiopian education development roadmap, for instance, highlighting the introduction of continuous assessment in

the past fifteen years, stated as it had little positive impact on the quality of teaching and learning (MoE, 2018). The road map also comments the need to detach from paper-based-only assessments and suggests authentic assessment, rationalizing the assessment of the learners' performances must provide an opportunity to verify that they can apply it to the real-world setting (MoE, 2018). Further, it clearly states instructors/teachers' improper assessment practice: “instructors, *fearing aggressive evaluation from students*, are not stiff enough to properly assess students' performance resulting into inflated grades” (MoE, 2018:52). However, if instructors are really carrying out assessment the way they are doing just because of fear from students' evaluation is open to discussion. Overall, the document has envisioned authentic and alternative assessment approaches as the way forward for Ethiopian HEIs.

Based on the road-map, “Ethiopian Higher education curriculum framework” (MoSHE, 2019) was designed. The inadequacy of the existing curriculum to address national unity, critical thinking, employability skills, communication skills, digital literacy and so on, are the focal points that attracted the design of the new curriculum. Regarding assessment, it underlines that the interdependence among instruction, assessment and curriculum provides optimum condition for learning that results in higher students' achievement. It also emphasizes that assessment should be used both to measure students' achievement of the required competency and facilitate learning (MoSHE, 2019: 17). The need to involve students in the assessment process is also emphasized. The intention seems to shift the assessment approach that solely depends on instructors to that of student-led too, so that learners will be self-directed individuals who take greater responsibility in their own learning.

To summarize, assessment has been one of the critical agendas of the reform initiatives and has been emphasized as a means to ensure quality of education. To that end, the reform movements have introduced some assessment orientations, approaches, techniques and concepts which are expected to be implemented and to inform assessment at higher education institutions. However, the documents have limitations in providing the necessary details on assessment. Besides, leave alone specific to EFL, standalone assessment-based reform and comprehensive assessment standards on general education is still lacking in the country. This, in turn, may pave way for the institutions and instructors to implement assessment according to their own interpretations.

2.4. Assessment and its role in ESL/EFL Teaching

Until recently, assessment in EFL/ESL classrooms, mainly informed by behaviorist theory, was governed by summative evaluation, the focus being on linguistic accuracy and distinct language features, rather than communicative competence (Rea-Dickins, 2001; Hill and McNamara, 2012; Fulcher, 2014). There has been increasing criticism since such approach to assessment and evaluation hardly emphasizes and measure students' use of the target language for communication purpose. Since recent years, however, there has been a greater concentration on the importance of communicative competence in language study, probably, as a result of increased dependency on language as a communicative tool. Therefore, communicative teaching methodology calls for a considerate emphasis on learning to use language in authentic settings and the assessment of language proficiency in realistic terms (Bachman, 2007; Fulcher, 2014).

As a result, there has been a dramatic change in language instruction and assessment (Buhagiar, 2007). As instruction methods change to incorporate new standards and principles, assessment approaches and strategies have been adjusted accordingly. For example, alternative assessment, in the form of self-assessment, peer assessment, and portfolios are incorporated into assessment methods and instruction as they are found to positively impact learning. Such adjustments are also made in line with shifts in assessment purposes. The sole evaluative purpose of assessment has been changed to include the use of assessments to facilitate instruction and learning too.

Currently, classroom assessment is an integral part of instruction in ELT (Brown, 2003; Fulcher, 2010; Latif & Wasim, 2022), and it can establish the current status of learners and what level they have achieved, give them feedback on their learning, diagnosis their needs for further development, and it can enable the planning of curricula, materials and activities (Fulcher, 2010). Unlike the earlier understanding that bestows the responsibility of assessment to experts, today, classroom assessment in higher education EFL/ESL context is viewed as collective responsibility of different agents including the institution, instructor and students, and each part distinctly benefits from the assessment.

HEIs are expected to make important professional decisions with regard to the teaching-learning process. They make different decisions on content, materials, methods and other alternatives in order to enhance quality of education, and more specifically, students' English language knowledge and skills (Norton, 2009; Holzweiss et al., 2016). However, sound decisions require valid information and should not be judgmental, and this vital information can be effectively

provided by quality assessments (Xu & Brown, 2016). Therefore, classroom assessment is vital for higher institutions, as it validates and enhances decision-making processes.

In addition, since assessment can provide the reliable information English language instructors need in the teaching-learning process, the instructors will be able to check if their instructional approaches, methods, and strategies meet English language learners' learning needs using assessment results. Then, using this as a basis, they can adopt effective alternative ways, such as devising instructional activities that better serve students and doing remedial activities like re-teaching (Alderson, 2005; Price et al., 2014). Therefore, as Fulcher (2010) also asserts, EFL instructors can benefit from assessment by obtaining accurate evidences, which is used to provide formative feedback for language learners, scaffold learning, evaluate the quality and pace of learning, inform pedagogy and feed into curriculum design.

On the part of the learners, assessment can offer a learning opportunity by providing them a chance to learn from their mistakes. Alderson (2005) argue that if EL instruction is linked to assessment, that gives students insights into their own thinking and growth, and they gain new perspective on their potential to learn and use the language. Assessment has also the capacity of leading students to deep approach of learning though poorly designed assessments can also encourage surface learning. Particularly, learners in higher education can benefit a lot from assessment if instructors are able to design assessments which require learners to demonstrate higher level of learning that result in adoption of deep approach to learning (Norton, 2009; Siarova et al., 2017).

Generally, classroom assessment has the potential to make instruction effective and learning more meaningful, and as communities of HEIs, instructors are expected to depict the potential of assessment in the teaching-learning process. However, although assessment promises effective instruction and meaningful learning, there is no guarantee that this is so until the implementors are knowledgeable and skillful especially on the newly emerging assessment developments and able to implement them (Scarino, 2013; Xu & Brown, 2016). In other words, every stakeholder has to effectively play his/her role in order to effectively implement quality assessment, for the effective application of assessment in ELT requires efforts of all.

2.5. Assessment practices

Assessments can be conducted in a variety of settings, including language and general education within the higher education and school sector, vocational education, as well as within the workplace. Within school and higher education contexts, assessment typically refers to the process of gathering and organizing evidence of student learning for making inferences about teaching and learning activities (Buhagiar, 2007; Tan, 2012; Hill & McNamara, 2012). Given the current study's situation within higher education, the discussion of each key stage in the assessment process is therefore specific to a classroom-based assessment context.

Assessment process within this setting typically encompasses a number of components such as defining the purposes of the assessment, constructing or selecting assessment methods to collect evidence of learning, correcting and marking, interpreting assessment outcomes collected, giving feedback and communicating, and grading decision making (Hill & McNamara, 2012; Xu and Brown, 2016). The other essential practice in modern assessment is involving students in the assessment process (Xu & Brown, 2016; Sheehan & Munro, 2017). However, this section presents and briefly discusses only some of the key components in the assessment practices process relevant to the research questions. They are also the commonly suggested assessment practices in classroom context to enhance learning.

2.5.1. Purposes of Assessment

While implementing assessment, at the beginning, instructors and others need to take into account the purposes for which assessments are carried out. knowing the clear purpose of assessment means understanding why any type of assessment is being used (Fulcher, 2010; Xu & Brown, 2016). This also entails relating assessment purpose with learning goals and specific content being learned and employing relevant assessment strategies for that specific learning outcome (Ibid). Hence, in higher learning institutions instructors need to identify why they conduct certain types of assessments while assessing English language skills and how to relate different assessment methods with the purpose for which they use.

Assessments are carried out in the classrooms for a variety of functions including: for instructional (adjusting instruction), placement (grouping students in levels), evaluation (determining progress) and accountability (providing information to administrators) (Hughes, 2003; Norton, 2009). Other writers categorize these assessment functions broadly into formative

and summative (Alderson, 2005; Fulcher, 2010; Xu & Brown, 2016) despite disagreements in terms of their relationships. For some, definitions regularly given for these assessment types, which relates to the timing (at the end or during process), has become a cause for distinctions (Brown, 2004; Fulcher, 2010). For others, the distinction concerns the nature of the assessment event and the nature of the assessment purpose, that is, the use to which assessment results are put. For instance, Harlen (2005) asserts that at the heart of the distinction between summative and formative assessment is the use of evidence, who uses it, how and for what purposes rather than the frequency and timing of assessment.

Whether there is continuous or dichotomous relationship between formative and summative assessment is also a topic of debate. Some tend to regard the two as a dichotomy that cannot be reconciled, conflicting at every facet: theoretical underpinnings, functions, approaches and procedures (Teasdale & Leung, 2000; NRC, 2001). In contrast, other groups claim that there is no a sharp discontinuity between the summative and formative assessments and perceive the two as sitting at the two ends of the assessment continuum (Harlen, 2005; Norton, 2009).

Since recent times, the classification of assessments purposes that include: assessment of learning, assessment for learning, and assessment as learning have been receiving attention. Assessment of learning, also called summative assessment, is correlated to the use of usually tests and exams at the end of lessons, a semester/course for the purpose of grading, certification, evaluation of progress, or even for researching the effectiveness of a curriculum (Earl, 2003; Brown, 2004; Norton, 2009; Fulcher, 2010). As it usually emphasizes on accountability purposes, the prevalence of assessment of learning practices in the classroom may result in restricting students' engagement in assessment, through which students self-monitor and self-correct their learning, however. Hence, overemphasizing the practice of assessment of learning particularly in communicative English skills course would be counter to the goal of the course.

On the other hand, assessment for learning also often associated with formative assessment, which emphasizes the integration of assessment to instruction, is practiced mainly to enhance instruction (for the immediate benefits of students' learning) (Earl, 2003; Fulcher, 2010; Brookhart, 2011). In assessment for learning, instructors use assessment as an investigative tool to find out as much as they can about what their students know and can do, and what confusions, preconceptions, or gaps they might have. So, this assessment type is essential in determining

English language learners' areas of weaknesses and strengths and monitoring their progresses, overall resulting in learners' higher achievement (Alderson, 2005; Hill & McNamara, 2012; Dashti, 2019). Similarly, empirical evidences consistently showed how students' achievement boosts when assessments are used as an investigative tool to find out where students are in their learning progress and gaps they might have, upon which the necessary measures are taken in the instruction (e.g., Black & Wiliam, 1998; Hattie & Timperley, 2007).

The final category of assessment purpose, assessment as learning, reinforces and extends the role of assessment for learning, but the role of students magnifies in the assessment as learning. It emerges based on the current view of how learning occurs: learning is not just a matter of transferring ideas, but it is an active process of cognitive restructuring that occurs when students interact with new ideas (Earl, 2003; Stiggins & Chappuis, 2004; Gonzales & Aliponga, 2012). The main objective of such assessments is empowering students so that they can make decisions and solve problems independently being critical assessors of themselves (Earl, 2003). If this is emphasized, particularly in language learning, then the difficulty that students experience in learning English language can be minimized. This is because conducting assessments for this purpose creates a learning environment that develops and supports students' metacognitive skills and helps students learn from the assessment process as well as from the assessment results. Furthermore, it has been indicated that the current century learning skills such as critical thinking, problem solving, reflective learning and creativity are effectively developed through the use of assessment as learning (Stiggins & Chappuis, 2004; Price et al., 2014).

Based on the above discussion, it can be said that employing the three purposes of assessment is a must in ELT classrooms, and their hand-in-hand practices are widely encouraged since each of them have distinct purposes to serve. Hence, the debate over the purposes of assessment is mainly focused on the prioritizing of a certain purpose among the other choices rather than a choice matter of one or the other. In light of these debates and tensions, Shepard (2000) and Fulcher (2010) assert that the extreme favor of one purpose over another is dangerous, and recommend a balanced practice. This will also cater to the interests of the stakeholders involved and the society at large (Stiggins, 2004). However, discussions of some writers such as Black and Wiliam (1998) and Earl (2003) tend to suggest the need to minimize assessment of learning in the classroom justifying that the main goal of assessments in the classroom is to enhance

instruction and improve learning. Currently, assessment for and as learning seem to have gained momentum than assessment of learning in the classrooms. Figure 2.1 and figure 2.2, adopted from Earl (2003) display this reality.

Figure 1. Traditional assessment pyramid

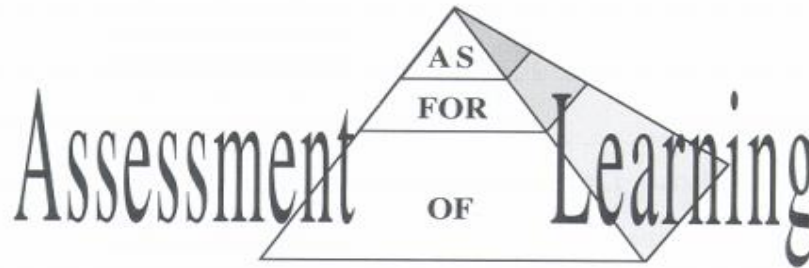
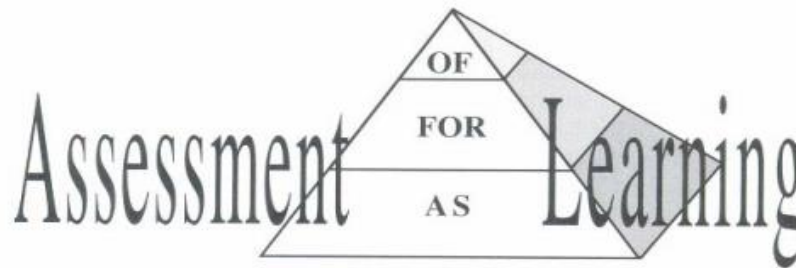


Figure 2. Reconfigured assessment pyramid



In spite of the fact, several previous studies conducted in higher and lower educational levels have consistently reported that assessment of learning is the predominant purpose for which assessments are practiced in EFL classrooms (Gebre, 2014; Saefurrohman & Balinas, 2016; Zelalem and Emily, 2018; Dashti, 2019). In the context of communicative English language skills course, Kitaw (2017) also showed that assessments are used solely for grading purposes and not for improving learning. Yet, there are instances in which assessments are primarily used by university instructors for instructional and student-centered purposes rather than for accountability (Cheng et al., 2004; Shohamy, Inbar-Lourie, & Poehner, 2008).

2.5.2. Assessment Methods

The other point is regarding the use of assessment methods that matches to a particular predefined assessment purpose. Of course, in order for teaching to be more effective, assessment methods should also be aligned with teaching methods, taking into account the learning goals

(Brookhart, 2011). If they are wisely selected, the assessment methods we adopt have an important role in the quality of learning and may be seen as an incentive for study and for improved performance (Cheng et al., 2008; Norton, 2009). That is, assessment methods can either support or diminish learning depending on their usage. Brown (2004) and Struyven *et al.* (2006) argue that the ways in which students look at learning are influenced by the ways in which they perceive assessment tasks. For example, for students who perceive assessment tasks are designed in a way that engage them in selecting one true answer, learning is seen as reproducing facts than meaning making.

There are different categories for assessment methods such as traditional Vs alternative, subjective Vs objective, select type Vs performance, or oral Vs written. For instance, Katz and Gottlieb (2013) classified types of language assessment methods used in classrooms on multi-levels as follows.

Table 1. Different classifications of classroom assessment methods

Selected response		Constructed response format		
Multiple choice	Brief Constructed Response	Performance-Based Assessment		
True/false				
Matching	Gap fill	Product	Performance	process
Same/different	Short answer	focused	focused	focused
Grammatical/ ungrammatical	Cloze	Essay	Oral Presentation	Observation
	Label a visual	Story/play/poem	Dramatic reading	Checklist Rubric
	Sentence completion	Portfolio	Role-Play	Anecdotal notes
		Report	Debate	Reflection
		Video/audiotape	Interview	Journal
		Poster session	Online chat	Learning log
		Project		

Basically, for many writers the assessment methods are broadly categorized as traditional and alternative assessments (Brown, 2003; Hughes, 2003; Cheng et al, 2008; Brown & Abeywickrama, 2010; Letina, 2015). Paper and pencil-based true-false, matching, multiple-choice, gap-filling and short answer methods are regarded as traditional assessments.

These traditional assessment methods dominated the assessment scene until the past two or three decades. Even, they are the overwhelmingly employed methods in the current educational practices of many, particularly, developing countries (Price *et al.*, 2014; Kitaw, 2017; Mitana, 2018). This prevalence of the traditional assessment methods can be related to the ease in which they could be administered in a wide variety of situations (Carless, 2015; Elshaw, et al., 2017; Sara, 2019). Such assessments allow a large number of people to be examined on identical material, at the same time, under similar conditions. They are relatively simple to create, easy to administer and to mark and score. From the psychometric theorists' point of view, the great emphasis paid to the traditional assessment methods may also relate to the belief that they are more valid and reliable (Serafini, 2002; Fulcher, 2010). These reasons might explain why this type of assessing became very popular even in the contemporary world.

However, traditional method of assessing students' learning is criticized for a number of reasons. Perhaps the most valid criticism of traditional assessment is its disregard for the essential connection between learning and assessment (Serafini, 2002; Buhagiar, 2007; Cheng et al., 2008; Brown & Abeywickrama, 2010). For example, tests or exams are often reserved until the end of learning, and they are utilized as a measurement of how students have performed, rather than how they are performing. Second, in many traditional assessment methods, students' views are not respected (Carless, 2015). Since students are often required to choose between answers that are similar in appearance, they are restricted from expressing their fuller views.

Third, traditional assessment methods are criticized for they do not reflect the principles of cognitive thinking. As assessment goals are different today because learning goals are different, assessment practices that emphasize on rote memorization of information have little to do in the current century environment. In this regard, Anderson (1998) suggests that traditional assessment represents the best way for evaluating traditional teaching which is based on facts presentation and their reproduction. Thus, it is not suitable for use in contemporary teaching which is carried out as discovery learning and which promotes diverse learning styles and multiple levels of thinking (Anderson, 1998; Brown & Abeywickrama, 2010; Letina, 2015). Study by Struyven et al. (2006) also showed the low effect of traditional assessment methods, compared to the alternative ones, on students' performance. In short, the traditional methods are criticized since they encourage reproduction, memorization and guessing; measure lower-order skills; promote

surface approach to learning; and they do not prepare students for real-life situations (Brown, 2003; Struyven et al., 2006; Boud & Falchikov, 2007; Katz & Gottlieb, 2013; Sheehan and Munro, 2017).

Hence, alternative assessment has emerged in response to such criticisms to traditional assessment methods (Buhagiar, 2007; Tan, 2012). Literature provides different terms such as performance, authentic, or innovative assessment that are used interchangeably with alternative assessment methods. However, for the purpose of clarity and consistency the term alternative assessment is used throughout this paper. Alternate ways of assessing students consider variation in learner needs, interests, and learning styles, and they have the potential to enrich the relationship between instructors and students, and to enhance learning (Carless, 2015; Sheehan and Munro, 2017). So, alternative forms of assessment are varied and numerous unlike traditional assessment which is identified primarily with written tests. Moreover, according to Janish et al. (2007: 226), alternative assessment “has the potential to reverse the traditional paradigm of student passivity and replace it with student initiative, self-discipline, and choice.”

The most popular examples of alternative assessment methods include self and peer-assessments and performance-based assessments. Self and peer-assessments are one of the commonly suggested methods in higher education context (Boud & Falchikov, 2007; Fulcher, 2010; Filler, 2017). It seems that the emergence of self and peer-assessment is grounded in self-directed learning and student-centered concepts, as independent learning is their core principle. Self-assessment, which requires students to assess their own work, can motivate them into becoming more active in their learning process (Brown, 2003; Katz & Gottlieb, 2013). It can help them locate their own strengths and weaknesses, i.e., they monitor their learning progress, and learners who recognize their own progress will be more motivated to stay on task (Bachman, 2007; Boud & Falchikov, 2007; Carless, 2015). Of all others, most important is that it can help students realize that the ultimate responsibility for learning is theirs. Shortly, employing self-assessment method can help students improve significantly and results in increased learning.

On the other hand, in peer-assessment students assess their classmates' work. Since learning is a social practice, no doubt that it can be more facilitated through peer-assessments. In addition, it is also helpful, the researcher believes, to develop their language performance for English language learners who fear practicing the language in front of their instructors, as they may feel

more comfortable when assessed by their peers. Students who engage in peer-assessment can have increased motivation, the opportunity for authentic communication, grasp in-depth understanding, and experience feeling of commitment and confidence (Brown, 2003; Brown & Abeywickrama, 2010; Katz & Gottlieb, 2013). However, it should be noted that peer-assessment has remained a challenge for instructors, as students' language proficiency, large class-size, time, and so on significantly affect its successful implementation. Overall, self and peer-assessment strategies motivate students to learn, assist them to become autonomous learners, and foster their feeling of ownership for their learning, resulting in enhanced learning achievement (Brown, 2003; Boud & Falchikov, 2007; Fulcher, 2010; Wellberg & Evans, 2022).

The other form of alternative assessment, performance-based assessment, has been gaining much attention from educators. According to Macmillan (2001:196) it "is one in which the instructor observes and makes a judgment about the student's demonstration of a skill or competency in creating a product, construing a response, or making a presentation". The concept of performance-based assessments seems to arise in response to a gap in many students to apply what they have learnt. It is obvious that many students are skilled on how to repeat facts, procedures, and concepts; however, they are consistently criticized for their inability to apply them to real-world problems that require complex thinking and reasoning skills (Janish et al., 2007; Wellberg & Evans, 2022). Hence, performance-based assessment requires students to apply their knowledge, skills and understanding in real-world contexts. It provides a great opportunity to make learning meaningful to students and to encourage them to be creative, innovative, and constructive (Brown, 2003; Tan, 2012; Wafubwa & Csíkos, 2022). The current advocacy behind performance-based assessment lies in its ability to bridge the gap between knowing and doing, as it allows students to apply what is learned in solving real-life problems and to think differently. Thus, it has advantage over the traditional assessment which only requires student to choose rather than creating responses.

Performance-based assessment has varieties of forms that can be generally classified under performance-based tasks, portfolio, journals and diaries, and projects. Performance-based tasks are tasks that require students to use a combination of knowledge and skills they have acquired to produce, perform, or accomplish a given task (Macmillan, 2001; Katz & Gottlieb, 2013). Tasks may include assignments, interviews, problem-solving tasks, communicative pair-work, role

playing, journal writing, debate, and group discussions (Ibid.). Such assessment methods are commonly employed to assess students' knowledge and skills, in which they are required to demonstrate their actual performance. Portfolio assessment, on the other hand, involves a collection of samples of students' work that display their skills, efforts, abilities and achievements during the course (Brown, 2003; Brown & Abeywickrama, 2010).

Student journals and diaries are the other alternative assessment writing activities with the primary goal of maintaining a steady stream of communication between the student and the instructor (Brown, 2003; Fulcher, 2010). If students are encouraged to write journal and diary in the target language, they will be more interested in expressing themselves clearly and correctly. Finally, project refers to prolonged investigation carried out by student into a topic selected in consensus between student and instructor. Thus, project works are a means of empowering students by giving them the responsibility of choosing a topic and managing the development of their work. It also offers students the opportunity to work independently or in groups for group projects.

Generally, based on the purpose for which it is used, it is not uncommon to use any of these assessment variety in EFL classrooms. Norton (2009) argue that each of the types can be used in the classroom interchangeably depending on the purpose for which assessment is to be employed. Brown (2003) and Fulcher (2010) also share the perspective that not a separate but a balanced use of assessment types that direct us to the goal of benefitting from its contribution to teaching and learning. To put in a nut shell, progressive use of a variety of assessment methods, apart from exclusively using the terminal paper and pencil tests, is highly recommended in assessing students' English language skills, particularly to address equity and fairness.

However, studies have frequently showed that while EFL instructors express that they also use assessment to improve students' language skills, and are working under communicative language teaching approach-oriented curriculum, in reality, their assessment practices are limited to the use of traditional assessment methods. This trend is evidenced by study of (Cheng et al., 2008; Dagne, 2009; Nibret, 2013; Kvasova & Kavytska, 2014; Sara, 2019) that reported that university EFL/ESL instructors tend to use tests as these were practical assessment instruments. These studies have also reported the polarized practice and excessive use of objective item questions, which according to Hughes (2003) and Sheehan and Munro (2017), encourages surface rather

than deep approach to learning that leads students to run for grades than learning English language for communicative purposes. Consequently, they came to the conclusion that assessments are practiced in EFL classrooms incongruent to the recommended practices in the recent literature.

2.5.3. Provisions of assessment feedback

In the teaching-learning process, assessment and feedback cannot be separated, and that is why feedback is the other most important aspect in the assessment process that integrates teaching, learning, and assessment (Black & Wiliam, 1998; Shute, 2008; Filer, 2017). It is the most important practice in modern assessment that potentially impact learning. Hattie and Timperley (2007: 81) define feedback as “information provided by an agent (e.g., instructor, peer, book, parent, self, experience) regarding aspects of one’s performance or understanding”. Using assessment feedback as a tool for enhancing teaching and learning has been increasingly advocated by scholars (Black and Wiliam, 1998; Fulcher, 2010; Brooks et al., 2019; Guo, 2020), as it is the most common feature of successful teaching and learning activity. Studies by Hattie and Timperley (2007) and Guo (2020) have also showed how assessment feedback enhance students’ learning and achievement.

Feedback provided to students should involve their learning progression, such as what they currently know, understand, and can do as well as how much progress they have made over a time and what is expected from them in future (Shute, 2008; Fulcher, 2010). This time, learners will also aware of what they have learned and what aspects of their performance should be improved (Fulcher, 2010). Thus, the main purpose of assessment feedback is to reduce the discrepancies between students’ current understandings and performance and their desired learning goals (Hattie and Timperley, 2007; Filer, 2017). That is, by giving feedback about the particular qualities of the learners’ work, with advice on what he or she can do to improve, it is possible to close the gap between their current EL performance and the expected.

This can be realized when the feedbacks are formative than summative (Shute, 2008). Although there has been tendency of equating the activity of feedbacking to enhance learning only to formative assessment practices, since recent times, it has been argued that summative assessments of the students’ language skills can also be used formatively to give constructive

feedback and improve learning (Brown, 2003; Fulcher, 2010). This is also indicated in Irons (2008) study, in which formative feedback is described as any information that accelerates student learning based on comments relating to either formative assessment or summative assessment activities. This suggests that instructors can use summative assessment tasks formatively to provide formative feedback to improve students' learning. As a result of feedback's formative effect on learning, students' knowledge and skills are formed into a more developed stage than they were prior to implementation of the particular feedback intervention (Hargreaves et al., 2000). Based on meta-analysis of over 250 studies, Black and Wiliam (1998) also showed that formative feedback practices enhance teaching effectiveness and bear student academic achievement gains ranging from 0.4 to 0.7 magnitude of effect size.

Generally, previous studies have also indicated how different approaches to feedback have different effects on student motivation and learning (Black & Wiliam, 1998; Boud & Falchikov, 2007; Brooks et al., 2019; Guo, 2020). For instance, according to Guo (2020) study, based on the type used, instructor feedback can be effective or dysfunctional in promoting student learning. Hattie and Timperley (2007), following their literature review, reached on similar conclusion that the type of feedback and the way it is given can be differentially effective. Yet, there is consensus that feedback practices that engage both instructors and students, employ variety of modes, and focus on serving different purposes actually promote students' learning (Shute, 2008; Filler, 2017).

Taking this into consideration and for in-depth comprehension of the concept of feedback, it is important to recognize its various categories and types. Researchers categorize assessment feedback into various forms depending on certain factors including, formality level, timing, frequency, sources, mode, audience, amount, function or focus, and so on. The present study investigates the practice of assessment feedback provisions by emphasizing on sources, modes, and focus/purpose of feedback types, which the following discussion is also limited to.

Sources of feedback: the sources of assessment feedback can be teacher, peer or self (Filer, 2017; Tasdemir & Yalcin, 2018). However, teacher feedback or teacher-led assessment feedback is the most commonly practiced and preferred feedback type by students, as teacher is considered the richest source of information in the classroom. According to Filer (2017), traditional classroom methodologies bestow the provisions of feedback to the teacher in all cases in the

language classroom, and the shift from teacher only feedback to also student-led feedback practices is a recent history in ELT. Although teacher-centered feedback provisions are criticized for a number of reasons, there is still a real need and value to teacher feedback, for many EFL students greatly value teacher feedback than feedback from peers. In addition, research has indicated that, if effectively conducted, teacher feedback is a key factor in improving student learning (Hattie and Timperley, 2007; Guo, 2020).

The other source of assessment feedback is students themselves through both self and peer-feedback. Peer-feedback, in which learners engage in giving each other feedback by commenting on their work, is more often preferred in modern language learning environments (Harmer, 2007; Filer, 2017). It is increasingly advocated, particularly in language learning context, for different reasons. It creates a more interactive environment for students to share their ideas with each other, and it increases learner autonomy as they take more responsibility for their work and are less dependent on the teacher (Brown & Wiliam, 1998; Hattie & Timperley, 2007; Shute, 2008; Brooks et al., 2019). Further, Filer (2017) noted that peer-feedback can make students critical readers, as it engages students in deep thinking and revising others' work with more attention. Self-feedback is also the other type where learner provides feedback on his/her own task. That is, learners themselves can be source of feedback by repairing their productions either written or spoken. This process is known as self-correction or self-repair, and it most commonly been discussed with regard to written feedback (Tasdemir & Yalcin, 2018).

In short, there is a general agreement that taking the responsibility of giving feedback out of the instructor's hands and allowing the students to be involved in the process has decisive role in increasing learning. Scholars on the area also caution the danger of dominantly relying on instructors as a source feedback and underscore that feedback is only helpful and support students' learning when students attend to and act upon it (Black & Wiliam, 1998; Hattie & Timperley, 2007; Filer, 2017). It seems apparent that a combination of instructor feedback coupled with students' feedback is of great value to students. Despite the fact, it has been consistently indicated by studies that the provision of assessment feedback is still under instructor/teacher control and students are rarely seen being used as sources of assessment feedback (Shute, 2008; Tasdemir & Yalcin, 2018; Dashti, 2019; Brooks et al., 2019; Gan, 2020).

Modes of feedback: provisions of feedback can also be classified according to the modes through which they are communicated. Instructors use different approaches to communicate their feedback on students' oral and written performance and competence so that students can learn from their errors and enhance their target language skills. According to Harmer (2007), there are two major forms of giving feedback on students' English language performance or understanding: oral and written forms. Oral feedback is information communicated verbally to students regarding their task or performance in the classroom, and it is usually provided during a lesson. Hence, although it is sometimes undervalued since it is less formal, oral feedback is a very powerful and effective approach to achieve the intended goal, as it is usually provided easily in the teachable moments and in a timely manner (Cheng & Wang, 2007; Harmer, 2007; Saefurrohman & Balinas, 2016).

On the other hand, written feedback, which refers to feedbacks given on students' tasks in a written form, often tends to appear after a task (Harmer, 2007). Similar to oral feedback, it can enhance learning if it is effectively implemented. Effective written feedback is identified as timely, that provides learners with information about what and where they are doing well, what needs improvement and next steps, and that is written in a manner that is understandable to the learner and actionable so that they can make revisions (Cheng et al., 2004; Hattie & Timperley, 2007; Shute, 2008). Although written feedback is frequently used with students' written performance and oral feedback is often employed in oral tasks or performance, the use of both modes in any context is helpful to address students with different needs and learning styles. It is the combination of both oral and written feedback that is more beneficial for students' learning than written or oral feedback only (Cheng & Wang, 2007; Harmer, 2007). Several previous studies have also consistently shown EFL/ESL teachers' use of both modes simultaneously (Cheng et al., 2004; Cheng & Wang, 2007; Saefurrohman & Balinas, 2016).

Focus or purpose of feedback: the final classification of assessment feedback in this research relates to the focus of feedback or the function the provided feedback serves. Researchers have proposed different types of feedback functions. The major ones and which are commonly known and most important in the context of higher education are verification, directive, and scaffolding.

Verification feedback refers to a dichotomous judgment of a student's response or answer by affirming it as correct or incorrect (Shute, 2008; Guo, 2020) and provides binary information that

describes whether or not the response given is correct (Black and Wiliam, 1998). The most common way is using right and wrong signs on the places that are correct, not completely correct or incomplete and assigning scores, numbers or grades on students' assessment papers as judgement of their academic performance (Shute, 2008; Guo, 2020). Hence, as this type of feedback often emphasizes on telling the right and the wrong knowledge, it may limit students' opportunity of making meaning based on their understanding. In line with this, Guo (2020) and Gan et al. (2021) point out that verification feedback aligns with and suits the traditional classroom teaching context, where transmitting information is the primary task of teaching. Since verification feedback carries no additional information about the task other than its state of achievement, it does little in facilitating learning (Black and Wiliam, 1998; Filer, 2017).

Directive feedback, also called corrective feedback, refers to direct written or oral provisions of the correct answers to students (Guo, 2020). Teachers are often engaged in directly providing the correct answer for students' wrong answer, or in supplementing the correct answer for the incomplete one (Shute, 2008). This provision of corrective feedback can take place in different situations, including, while correcting students' assessment papers (by providing the answers on the papers), when students receive back their assessment papers (by writing the correct answer on the board or on separate sheet), or during presentations (by correcting errors on moment). Although it is important to provide correct answer to any given question to help students know the right answer, teacher corrective feedback does not provide students with meaningful information on their learning progress and criticized for encouraging learners to focus on passing assessment than mastering learning goals (Black & Wiliam, 1998; Gan et al., 2021).

So, in order to combat such problems, a move away from purely teacher-led feedback provision is generally suggested, as this is more beneficial for students and can help them improve (Filer, 2017). Hence, the concept of scaffolding feedback, which values and amplifies learners' role in feedback provisions, has emerged. It refers to the use of a series of successive cues, prompts, hints, or partial solutions that consist of breaking down tasks into easier or smaller parts to facilitate students in generating correct answers by themselves (Brooks et al., 2019; Guo, 2020). Similarly, Shute (2008: 157) describes that scaffolding or facilitative feedback "provides comments and suggestions to help guide students in their own revision and conceptualization".

So, this feedback type focuses on directing and guiding students towards independently generating answers to questions and autonomously completing or solving a given task.

Scaffolding feedback is highly advocated in the contemporary classrooms as it also enhances self-directed learning and engages students in a deeper process of understanding and knowledge creation, taking them away from passively receiving information (Hattie & Timperley, 2007; Filer, 2017; Gan et al., 2021). Further, Shute (2008) points out that scaffolding feedback allows learners to perform more advanced activities and to engage in more advanced thinking and problem solving than they could without such assistance. Studies have also shown as the use of scaffolding feedback can effectively result in enhancing learning and achievement by increasing students' motivation (Hattie & Timperley, 2007; Shute, 2008; Gan et al., 2021).

In short, there is general agreed view that effective feedback should include elements of verification, directive, and scaffolding feedback (Shute, 2008). Students' English learning can be more enhanced by using these three feedback types in a balanced approach. However, what is consistently reported by non-local studies is that scaffolding feedback is less frequently employed in the classrooms compared to verification and directive feedback which are frequently practiced by teachers (Hattie & Timperley, 2007; Shute, 2008; Brooks et al., 2019; Dashti, 2019; Guo, 2020). While Shute (2008) and Dashti (2019) findings revealed more focus of teachers' feedback on justifying marks and grades and their limited practices of self-regulatory level feedback to build deeper and conceptual understanding, Guo (2020), from students' perspectives, reported that teachers frequently practiced directive feedback than the other types. On the other hand, Gan et al. (2021) study reported frequent use of scaffolding feedback among instructors of two universities.

2.5.4. Involving Students in the Assessment Process

Although engagement of students in the assessment process has been discussed in each of the above sections of the assessment practice, this section more broadly and inclusively presents students' role in the current assessment system. The concept of learning by doing and student-centered approach in teaching-learning process make it compulsory to involve students actively in the assessment process. Learners can be involved in both types of assessment approaches, but these has been lacking for centuries in the conventional assessment. Stiggins and Chappuis

(2004) argue that in traditional forms of assessment, learners are observed to be passive partners while they are active learners and become the master of their own learning in formative assessment practices.

Hence, another very essential practice in modern assessment is involving students in the assessment process in various ways starting from its planning to the use of assessment information. Learners should be provided opportunities for self-progress, assessing their own work, assessing peers' work and collaborate their teachers in developing standards and norms for their tasks (Stiggins & Chappuis, 2004; Brookhart, 2011; Xu & Brown, 2016). Of course, students must be aware of the learning intentions and what, why, and how they will be assessed at the beginning of instruction, and they must be provided opportunities to take part in the assessment processes starting from the beginning. This also includes activities such as consulting students about the type of assessment they would like to have and communicating them transparently with regard to criteria, time and content of the assessment (Stiggins & Chappuis, 2004). Furthermore, the authors elucidated that learners are engaged in the assessment process means they are also able to use information from assessments to manage and monitor their own learning (Ibid.).

Traditionally, the use of assessment information to impact teaching-learning process is equated with teachers; however, in the contemporary view of learning, students also need to use information from their own assessment intelligently in order to improve their learning outcomes (Lund, 2008). This notion is also described in Popham's definition of formative assessment which states: "a process in which assessment-elicited evidence is used by teachers to adjust their ongoing instructional activities, or *by students to adjust the ways they are trying to learn something*" (Popham, 2009: 5, emphasis added). In short, students can also use assessment-led evidences to inform their learning, and by actively engaging learners in the assessment process, it is possible to enhance achievement since it motivates them to feel responsible for their own learning and improvement (Brookhart, 2011; Xu & Brown, 2016).

2.6. Assumptions about assessment

2.6.1. Conceptualizing assumptions

The definition that literature provides for the term assumption is often not clearly articulated particularly in situations where it is intricately connected to concepts such as beliefs and knowledge. Of course, the definitions provided for assumption and the way it is thought about in the literature differs from writer to writer depending on their particular focus (Atkin, 2017). For example, Brookfield (1992, 2013) suggests that there are three ways to think about assumptions: as taken for granted beliefs, as rules of thumb that guide action and finally, as common-sense beliefs and conventional wisdom that we invoke. Then, with the final way of thinking, he stated that whatever we say to justify our actions or reasoning counts as an assumption.

Many contend that assumption is thought as part of one's belief system comprising things that are largely and generally held to be true (Brookfield, 1995; Reis, 2012; Ergun, 2013; Nkwake, 2013; Stabile, 2014; Atkins, 2017; Boone & Willment, 2018). For Nkwake (2013), assumption is the taken-for-granted beliefs about the world, and they form the belief that we hold. Similarly, while Reis (2012: 12) defines assumptions as "thoughts we take for granted and believe them to be true", for Yanchar & Slife (2004) assumption is an idea or proposition which is perceived to be true or taken for granted.

This taken-for-granted world of assumption is one of the key characteristics that are recognized by writers. NRC (2001), in their seminal report, was one of those to devote attention to the implicit nature of assumptions, identifying their importance as the taken-for-granted basis for developing an argument. Indeed, this is also reflected in the definition given by Atkin (2017) which describes assumptions as something that is certain to happen without proof. Assumptions, therefore, are so implicit and embedded in our thinking that they generally remain hidden and difficult to identify (Nkwake, 2013; Stabile, 2014), but also may be acted upon without attention being paid to their contextual validity and accuracy (Brookfield, 1992). Indeed, in his recent work, Brookfield (2013) also states that using the terms right or wrong assumptions is unhelpful since they are so contextually dependent, rather, they are merely more or less accurate depending on the information that has been available or used at the time. They are more than just a response to a situation and are informed by beliefs and values that shape thinking and actions in response to the situation and context (Nkwake, 2013; Brookfield, 2013; Atkin, 2017).

Therefore, it can be argued that assumptions are things which are accepted without proof and may form part of our belief system but are based on our experiences and learning. In the context

of health care, after reviewing the available literatures on the area, Atkin (2017: 46) tends to give the definition considered the broader one:

A taken for granted interpretation of the world, which is held with a sense of certainty until challenged through interaction with a different perspective and evaluated through a process of critical reflexivity. They are founded on past experience and are often tacitly held, influenced by both emotion and rational thought and are fundamental to reasoning, sense making and subsequently to behavior and action.

As mentioned above, the terms assumptions and beliefs are often used interchangeably. The premise ‘accepting something as true without proof’ characterizes definitions underlying both. However, based on definitions provided, it can be understood that beliefs are firmly held than assumptions (Oxford Dictionary, 2015; Atkin, 2017), which specify existence of differences in the degree of strength and certainty, implying assumptions are more open to change. So, assumptions are defined here as the experience and learning informed propositions or implicit/explicit statements instructors and students make about assessment that are taken for granted as a basis of argument or action, and are susceptible to change through examinations and critical reflections (Delandshere, 2001; Brookfield, 2013; Ergun, 2013).

They are conceptualized not as what instructors and students say they believe about assessment but what their justifications behind their assessment behavior (what they say and do) reveal about what they believe about assessment. Their assumptions about the form and nature of classroom assessment contents (how and by whom knowledge is defined), how it is assessed and learning evidence is interpreted (in an objective standardized or in a subjective manner), and the nature of the relationships between the instructors and students involved in the assessment process were emphasized. This is also based on Delandshere’s (2001) framework of how current assessment addresses epistemological and sociological questions.

Assumptions are the foci for educational theories, i.e., every educational theory is established based on a set of assumptions (Schraw & Olafson, 2008; Nkwake, 2013). Accordingly, every assessment is predicated on the assumptions the education theories make (NRC, 2001; Elton & Johnston, 2002; Pellegrino & Chudowsky, 2003). Hence, people’s assumptions about assessment can be identified and described based on the assumptions these theories make. For instance,

based on assumptions upon which the traditional and alternative assessment approaches are constructed, it is possible to indicate where the stakeholders' assumptions about assessment lean.

Brookfield (1995) has identified three different categories of assumptions that people make in the educational processes and need attention: causal, prescriptive, and paradigmatic. Causal assumptions, which he stated the easiest to identify, are stated in predictive terms of identifying consequences from actions. They describe cause-effect relations, reflecting the implication that an action taken will result in certain outcome. E.g., employing peer-assessment enhances learning. Prescriptive assumptions describe what we think should happen in a situation, how instructors should behave, what good educational processes look like and obligations that students and instructors have to each other. E.g., instructors should engage students in the assessment process. On the other hand, Brookfield (1995) describes paradigmatic assumptions as the most difficult to articulate and reveal, prompting us to reflect on the 'common-sense categories' we construct about our learners. They are our most deep-seated beliefs and things that we know to be true and is considered to be the most difficult to uncover. E.g., university students are self-directed learners.

2.6.2. Influences of Assumptions

Literature indicates that assumptions can influence situations both positively and negatively. For example, Adair (2009) asserts that assumptions that are made consciously actually encourage exploration and creative thinking. This is on the basis that assumptions that are held "deliberately and temporarily" create the opportunity to use the imagination to consider what might be possible (Adair, 2009: 61). He distinguishes these from assumptions that are unconscious, which he argues hinder innovative thought.

Assumptions can have a negative influence, particularly in relation to education. They can be used to stereotype learners in terms of their race, sex, class participation, competence, disability etc. (Nkwake, 2013), and they are harmful when not examined and lead to professional intrenchment when not probed (Adair, 2009; Reis, 2012; Atkin, 2017). Their influence is thought to be so detrimental to effective relationship between learners and instructors. Iezzoni (2006), to solve this, suggests that staffs should start with the premise 'make no assumptions' (p.212).

Indeed, tacit assumptions that go unknown or are ill considered can have unhelpful effects. However, the solution Iezzoni's (2006) suggests that assumptions should be suspended tends to imply that people can almost erase personally held assumptions, which does not seem practicable. In order to suspend assumptions, so that their negative effects are not felt by others, they need to be made explicit, either internally through reflection or externally by explicitly uttering them (Ergun, 2013; Stabile, 2014; Atkin, 2017). If assumptions remain hidden, they are not consciously accessible and reflection upon them would be impossible, and thus, suspending them is not helpful. Hence, whether valid or not, assumptions should be made explicit, and it is not wise to deny our assumptions, as also asserted, "to deny a proposition is not the same as to confirm its denial" (Nkwake, 2013: 131, referring to Russell, 1904).

Instructors and students make assumptions about different aspects of classroom assessment, i.e., purposes, contents, methods, processes, or quality issues that influence the practices of assessment in and outside the classroom. Scholars also persuasively argued that assumptions inevitably underlie the way people see things and perform their work (Brookfield, 1995; Stabile, 2014; Wellberg & Evans, 2022), even when people are not aware of these background beliefs (Yanchar & Slife, 2004). In this sense, assumptions are enabling, rather than limiting; they are what actually allow for the possibility of activity, decisions and reasoning. Upon this perspective, it is impossible to theorize or to practice in the field of assessment without assumptions to guide one's work since they shape an individual's assessment activity at every step.

Assumptions about assessment held by the staffs and students can positively influence their assessment reasoning and have constructive effect on their relationships, and generally how assessment is practiced. An instructor who assumes there are multiple realities, and thus, assumes that inviting students to demonstrate different perspectives on an activity enhances understanding, may also design open-ended questions and conduct its scoring with similar perspective. It has been argued that one of the central features of educational development programs is the importance they place on the underlying assumptions faculty bring to their work (Booke & Willment, 2018), as "the success of any educational policy meant to improve instruction is heavily dependent upon the assumptions that teachers have" (Wellberg and Evans, 2022: 3).

Overall, assumptions are perceived to influence one's assessment behavior and action both positively and negatively, given the power of assumptions in determining what is assessed, how it is assessed, how learners are perceived and treated, and instructors' positionality. To this end, Holzweiss, *et al.* (2016) suggest researchers to capture more effectively the underlying assumptions that various stakeholders in higher education (e.g., administrators, faculty, students) hold about assessment in order to promote and establish a true and effective culture of assessment. Therefore, the first step to benefit from the assumptions made in the field of assessment is to explore them in detail so that examining them can be possible.

As practitioners' assessment assumptions form the core of their assessment behavior and changes in their behavior often follow changes in their assumptions or beliefs (Stabile, 2014; Latif & Wasim, 2022), unpacking EFL practitioners' assumptions about assessment as they are used to inform practice has huge implications, particularly for professional development. Once assumptions underlying the current practices are made conscious, assumptions interfering with reform-oriented assessment practices can be identified and paid attention to when developing professional learning opportunities. Clarifying existing assumptions thus provides opportunities to unveil those interfering with the innovative practices for examinations (for critical reflection) through which shifts in assumptions can be realized, resulting in a positive culture of assessment.

In line with this, Delandshere (2001: 113) also argues that without clarity on assumptions behind assessment, there will not be theoretical and philosophical debates, which in turn may "result in practices that tend to reproduce themselves in a vacuum, resist change, and are disconnected from relevant issues of knowledge, power and social organization in general". Thus, in order to examine their influence on assessment-related decisions and practices, EFL assessment practitioners' assumptions should be brought out of the closet. Atkin (2017) made that very point by stating, it is apparent that the influence of assumptions can be explored and understood when they are made explicit; therefore, finding ways to illuminate assumptions is important in order to determine their validity. Leung (2005) and Taylor (2009) also suggest researchers to focus not only the practices, but also to delve into the mental world of instructors to better understand the theories they use to define and conduct classroom assessment activities.

2.6.3. Exploring Assumptions

The exploration of assumptions is one aspect of critical thinking, and it is widely exercised in the critical reflection practices (Brookfield, 1995; Gabbitas, 2009). Hence, in different fields including health care, program evaluation, and education, it is used as a gear for reflection practices. Exploring assumptions is the first step in reflecting on one's own or others' decisions and practices in order to make explicit the assumptions a person or a group of people holds, and then, to bring to critique and examination in line with the evidences and contexts in which they are held. Thus, such process of illuminating assumptions includes not only uncovering but also analyzing its implications and reflecting on its origins, nature and consequences (Brookfield, 1995). It further includes exploring new and alternative ideas in order to revise or replace old assumptions (Brookfield, 1995; Atkin, 2017).

Gabbitas (2009) suggested how and what one can go through while illuminating assumptions. These include, possessing an awareness of basic assumptions, examining motives, discovering assumptions, examining their implications, evaluating assumptions and implications, and finally considering alternatives. Of course, the first three can be used in the current study since this study is aimed to explore and identify the assumptions that underly EFL assessment practices rather than reflecting on the origin, implications or evaluation of them, which are focal in critical reflection process.

The first step and an important starting place to explore people's underlying assumptions is having an awareness of those basic or general assumptions. Gabbitas (2009) argues that the good knowledge of some philosophical issues is so fundamental in exploring assumptions in educational practices including how people view learning and teaching. They are also basics in exploring and understanding assumptions in specific educational affairs like in assessment. Thus, having basic understanding in relation to some theoretical and philosophical issues such as epistemology, ontology, the place of context, learners' agency and power, and others are important in exploring and understanding those assumptions underpinning assessment.

It is argued that examining a person's motives underlying his/her work can lead to discovering assumptions (Gabbitas, 2009). This view is also expressed in Booke and Willment (2018) article that explores instructors' underlying assumptions of teaching through answers to the question of what explains why the instructors are thinking and deciding the way they are. Flynn and Schachter (2017) have also used the same strategy to explore instructors' assumptions of how

students learn through what they call ‘pedagogical reasoning’. In the same vein, Brookfield (1995) commented that responses to questions such as ‘why did you do this?’ and ‘why do you think that?’ can guide the exploration of assumptions. Adding, he asserted that whatever we say to justify our actions or reasoning counts as an assumption.

For instance, is a instructor using performance assessment is because he/she thinks that students develop their English language skills more by performing than by selecting from what is provided? Follow up questions such as ‘Why would this method be better than some other methods?’ can also help in explicating the assumptions more in detail. According to Gabbitas (2009), while answering such questions the respondent may make practical (e.g., because it was requested by a manager) or theoretical (e.g., because this best addresses how I believe learners process knowledge) motives. Examining such motives may also lead to more questions and the discovery of further assumptions.

As described, while examining motives underlying decisions or practices, further questioning and examining can direct to recognition of more assumptions. The next stage, thus, is analyzing how those reasons can be connected to theoretical, philosophical or sociological assumptions. This is what Gabbitas (2009) call discovering assumptions. At this point, the question might be asked is ‘given that assessment is practiced this way, what does that say about knowledge, learning, context, agency, power relationship, or value?’ based on (Gabbitas, 2009: 50). In this manner, the enquirer more clearly sees how decisions and practices are connected to assumptions about the world and human experience.

With regard to locating and connecting the assumptions into their corresponding theories, the common practice is discussing the assumptions under contrasting theories. Much of the discussion has focused more on constructivist and relativist assumptions in contrast to cognitivist and behaviorist views (Gabbitas, 2009). Some researchers also use subjectivist vis a vis objectivist/ positivist category of assumptions in the context of teaching and learning (Carson, 2005), or on the basis of some (two or more) handpicked theories. With regard to assessment, similarly literature presents varying categorizations of theoretical and philosophical assumptions underlying assessment (Anderson, 1998; Serafini, 2001; Elton & Johnston, 2002; Mitana, 2018). For instance, Anderson (1998) saw these as two ends of a continuum and labelled traditional assumptions and alternative assumptions. He identifies, contrasts, and describes assumptions

upon which traditional and alternative assessment approaches are constructed on several interconnected factors. His typology presents comprehensive categorization system for each of the factors and provides a useful framework to explore people's assumptions about assessment.

Table 2 Comparison of philosophical beliefs and theoretical assumptions of alternative and traditional assessment

Traditional assessment	Alternative assessment
Assumes knowledge has a single (universally) consensual meaning	Assumes knowledge has multiple meanings
Treats learning as a passive process (the emphasis is on learning about something, rather than on learning how to do something)	Treats learning as an active process (the emphasis is on learning how to do something)
Separates the learning process from the final product (evaluate only the final product)	Emphasizes the learning process and the final product of learning (taking into account what, why and how students learn)
The focus is on the use of pieces of information (using lower levels of reviews)	The focus is on research, i.e., developing the ability to solve real problems.
Assumes the purpose of assessment is to document and monitor student learning and to classify students by their scores.	Assumes the purpose of assessment is to facilitate learning.
Students' cognitive, affective and conative abilities are separate (emphasis on the cognitive dimension)	Recognizes a connection between students' cognitive, affective, and conative abilities.
Embraces hierarchical model of power and control (students do not participate in decision-making)	Embraces a shared model of power and control (students participate in decision-making)
Perceives learning as an individual enterprise (student should independently solve a given task).	Perceives learning as a collaborative process (instructor and student are classmates).

Adapted from Anderson (1998)

2.7. Assumptions underlying traditional and alternative assessments

By examining literature relating to educational assessment, one can detect the existence of two different forces that influence the manner of assessment in education. It presents an extreme polarization of assessment that presupposes varying theoretical assumptions. On one extreme side of the pole, assessment tends to be systematic, psychometric, rigid, standardized, objective, quantitative and dogmatic in approach with a number of traditional rules and procedures. The purpose of such assessment is usually selection, grading, and certification. On the other side of the pole are assessments that are social, informal, subjective, context dependent, qualitative and less systematic. In this approach, the purpose of assessment is basically to support teaching and

learning. This section discusses the basic assumptions underlying traditional assessment and alternative assessment approaches.

2.7.1. Assumptions of Traditional Assessment

Traditional curriculum operates based on the idea of objectively and scientifically identified standards (Schraw & Olafson, 2008). This standards model is grounded in the philosophies and theories that assume individuals need to have fixed dispositions to behave in certain ways across situations. That is, there exists fixed, objective and measurable- “a gold standard” based on the assumptions, “knowledge is monolithic, static and universal” (Delandshere, 2001:127). An EFL instructor operating primarily from this approach, for example, is more likely to endorse a belief in a universal curriculum, and that a course must present a predefined body of knowledge that consists of facts, principles, procedures, and theories. Learning, thus, entails the effective transmission of this body of knowledge from the knower to learner and is conceived as a step-by-step acquisition of specific discrete skills (Biggs 1995; Russ et al., 2008; Letina, 2015). Recognized experts have to control what and how students learn, choose what is essential to learn, the sequence, and the learning activities (Delandshere, 2001; Buhagiar, 2007; Schraw & Olafson, 2008).

Traditionally, assessment is viewed as a mechanism of defining student ability by quantifying the degree of attainment of contents prescribed in the curriculum (Biggs, 1995; Serafini, 2002; Moss, 2003; Letina, 2015). The purpose of assessments is to create hierarchies of value- such as separating who know from who do not know (Anderson, 1998; Elton and Johnston, 2002). Hence, the genre assumes that assessment is means of measuring the amount of knowledge acquired by a student over a period of time (Biggs 1995; Russ et al., 2008). This implies that knowledge exists separate from the learner, and students work to accumulate rather than to construct knowledge. This, in turn, indicates that traditional assessment focuses on correctness of knowledge: ideas are good to the extent they align with what is presented in textbook or curriculum (Russ et al., 2008). In this regard, the best method to assess students’ ideas is to compare their responses to accepted acknowledged knowledge as represented in the textbook or curriculum (Ibid.).

Thus, based on its monolithic, static and universal assumptions of knowledge, traditional assessment advances the notion that there is certain specified knowledge that needs to be transmitted to learners, which are later expected to be reproduced without error through exams (Anderson, 1998). In addition, in order to effectively transmit and assess its efficient transfer, the genre also assumes that complex knowledge should be broken into its components. Traditional assessment relies on a fundamental assumption that an entity to be measured can be broken down into small and independent measurable pieces as components for objective observation and scoring (Serafini, 2001; Buhagiar, 2007; Klenowski & Willis, 2011). Accordingly, students' language competence is believed to be well measured in additive assessments of writing, grammar, reading, vocabulary, and other discrete points of the language.

Hence, assessment focuses on evaluating effective transmission of the homogeneous, essential and discrete knowledge. Since assessment focuses on measuring the universal truths, the dimension being tested is also assumed to be stable (Delandshere, 2001; O'Donovan, 2017). That is, content of assessment and assessment measures are stable across settings and students, the point that carries with it the concepts of consistency (all students are expected to take assessment under the same condition to obtain the most objective data) and comparability (observations should yield the same results) (Anderson, 1998; Delandshere, 2001; Buhagiar, 2007; Klenowski & Willis, 2011).

In line with the notion of universal education, traditional approach advocate for standards model of assessment which is viewed as a form of scientific enquiry that discovers the fixed student's ability objectively and reliably (Delandshere, 2001; Serafini, 2001; Moss, 2003). Judgement and interpretation of students' learning outcome is made by following 'scientific' procedures and method of inquiry such as by calculating validity and reliability indexes (Serafini, 2002; Klenowski & Willis, 2011). So, assessment is too much inclined to the quantification and numerical description of students' learning outcomes. This is grounded in the premise to achieve objective, rigorous and accurate interpretations and to easily identify and track their achievement. In this regard, Biggs (1995:3) argues that such model of assessment is grounded in the assumptions that "public standards can be set; ... and objective, fair and consistent judgments are possible to determine whether the standards have been met or not". In other words, traditional assessment assumes that objective and consistent judgments are possible to determine

whether the prescribed standards have been met or not (Serafini, 2002; Klenowski & Willis, 2011), implying the assessment results are accurate indicators of students' learning. Assessment in the traditional approach is, thus, grounded on certain principles that are universal-standardized for all students.

Assessment issues are under the control of classroom instructor, the expert. Instructor determines what and how students are assessed, choose what is essential to assess, when to assess, and the assessment tasks and activities (Buhagiar, 2007; Schraw & Olafson, 2008; Tan, 2012). Implied in this is also the existence of hierarchical power relationships between the instructor and students. Regarding who should participate in the acts of assessment, Delandshere (2001) states that assessment is typically thought of as a judgement made by the assessor (knower) about the examinee. That is, instructors are empowered to make judgement about students' knowledge. In line with this, Short & Burke (1991 in Anderson, 1998: 6) state that assessment is imposed on students because of different assumptions- "that the questions which drive the curriculum must be supplied by recognized experts, that the vast majority of what is to be learned is already known, digested, and organized, and that there are acknowledged correct responses to the curricular questions which are to be asked".

Congruent with their hierarchical model of power and control, thus, there must be some kind of distance between the instructor and students in assessment. Traditionally, instructor controls power in assessment, as it is assumed that only experts (instructors) know about assessment and involving students in the process may lead to invalid results since they lack knowledge and expertise that can allow them to take part in assessment and worth for consideration (Anderson, 1998; Delandshere, 2001; Serafini, 2002; Tan, 2012; Lau, 2013).

2.7.2. Assumptions of Alternative Assessment

Alternative assessment, which was introduced primarily in response to the current meaning-making process view of learning (Buhagiar, 2007; Al-Bahlani & Ecke, 2023), is understood here as a paradigm shift from traditional assessment and is a process that is firmly integrated with instruction and positioned for promoting learning (Tan, 2012). So, it is based on a totally different set of philosophical beliefs and theoretical assumptions (Bintz, 1991, cited in Anderson, 1998). As Anderson (1998) noticed, it is not what students answer, but why and how is central in

the assessment of students work. In this sense, alternative assessment focuses more on process than product of learning, as its primary purpose is assumed to support learning (Tan, 2012).

Alternative assessments come from the subjectivist point of view, which looks at knowledge as individualized, depending on individual perceptions, and localized because people are different (Buhagiar, 2007). That is, knowledge is seen as temporary, developmental and relative in this assessment tradition. It is conceived as something that the students construct, not something passed from omniscient authority to the naive passive one (Anderson, 1998; Janisch *et al.*, 2007). Alternative assessment, thus, gives value for the locally, socially and culturally mediated construction of knowledge. One of the fundamental assumptions that underly is that assessment should encourage knowledge construction ability of students, unlike the traditional assessment that advances the view of a common core of knowledge that needs be directly transmitted from the instructor to students. In line with this, Janish et al. (2007) asserted that “the theoretical framework for using alternative assessment in the classroom includes considering learners as constructors of knowledge ...” (p.221). Nasab (2015) also added that alternative assessment grants students to generate knowledge and consider them as agents of their learning.

Relating to this, alternative assessment holds the position that there is not one single correct way towards learning and one correct way of solving a problem or responding to a question (Anderson, 1998; Black & Wiliam, 2009). Students are allowed the freedom to construct the reality they are satisfied with using the method that satisfies them by engaging themselves in self and peer-learning activities (Serafini, 2001; Carson, 2005). In the context where varied outlooks are welcomed to a particular question, there is no hard and fast rule, and students are allowed to demonstrate their varied and justified perspectives in a task. Thus, assessment is seen as a tool to seek for students justified answer rather than that looks for correct prescribed canonical knowledge as presented in the textbook or curriculum (Russ et al., 2008; Schraw & Olafson, 2008). Embedded in such approach is the assumption that there could be possible alternative answers and ways of responding to a particular content assessed. Janisch *et al.* (2007) also states that the focus of assessment in this approach is on the ‘reasonableness’ of the student’s approach in solving problems and performing tasks, not just to look for the correct response. Thus, students are made to engage in the higher order thinking skills demanding tasks upon which different and multiple views emerge.

Further, in relation to its view of knowledge as holistic rather than compartmental, alternative assessment places greater emphasis on a more holistic approach to language acquisition and assessment (Bachman, 2007). Particularly, influenced by the communicative approach to language teaching, which calls for a considerate emphasis on learning to use the target language in natural settings, alternative assessment emphasizes on assessment practices that depict communication function of language (Rea-Dickins, 2001; Hill & McNamara, 2012; Fulcher, 2014). Hence, based on the understanding that this cannot be realized through separate teaching of the skills and assessment of distinct language features, integrative assessment of the language skills is highly advocated. This grounds in the assumption that integrated skills assessments that accounted for a holistic observation of students' use of the target language is effective, which implicitly suggests holistic view of knowledge.

Another issue emphasized in this approach is the contextual nature of assessment. Assessment issues should take into consideration the context within which knowledge is constructed. It is perceived that assessments and what they measure are contingent upon the social context in which they are designed and used (Janisch *et al.*, 2007). Unlike the traditional approach, which views social, cultural and political factors as threat to assessment validity, for the alternative approach, they are “integral and meaningful components, without which the assessment effort would be sterile, useless and meaningless” (Huot, 1996 in Elton & Johnston, 2002: 43). So, alternative assessment rejects the notion of objectivity and value-free assessment, rather it acknowledges subjectivity in assessment, which it considers also important (Anderson, 1998; NRC, 2001). Assessment in this paradigm is viewed as a social, contextually specific, and interpretive activity; it is not viewed as an objective measurement process, intended for comparisons and prescriptions; rather, it is seen as a human interaction involving the human as the primary assessment instrument (Serafini, 2001).

Similarly, individual learner's subjectivity is the foundation of assessment- an assessment that is tailored to individual learner's experience (Janish *et al.*, 2007; Black & Wiliam, 2009). Since knowledge can be learned in different ways, it should also be assessed in different ways, and variations in assessment is viewed as legitimate, and thereby differences in assessment standards, methods and marking criteria (Gipps, 1999; McNamara, 2001; Serafini, 2002; Buhagiar, 2007). A variety of assessment tasks need to be used so that students with diverse learning needs, styles,

backgrounds or preferences can be addressed and demonstrate their learning (Elton & Johnston, 2002; Janisch, *et al.*, 2007). Hence, the approach assumes that an adapted assessment that recognizes individual differences and environmental variations that may appear among institutions and individual learners safeguard the issue of equity and fairness. Universal and standardized principles are, thus, has no place in alternative assessment.

Based on its view of social construction of knowledge, there should be negotiation among students, with their instructors, and other stakeholders on every classroom assessment issues (Anderson, 1998; Shepard 2000; Black & Wiliam, 2009). Particularly, students along with instructors decide on what and how to assess, seek and interpret evidence for use to decide where they are in their learning, where they need to go, and how best to get there (Serafini, 2002; Wellberg & Evans, 2022). This of course presupposes to create trusting relationships and the sharing of power. The proponents of alternative assessment insist that students should be given power in assessment since they are also effective instructional decision makers that contribute to their learning (Buhagiar, 2007; Tan, 2012; Nasab, 2015). “The theoretical framework for using alternative assessment in the classroom includes ... empowering learners.” (Janish et al., 2007: 221). So, one of its assumptions is that instructors should share power with students to allow them determine how well they are learning (Anderson, 1998; Tan, 2012; Letina, 2015). Generally, alternative assessment shares theoretical perspectives that are fundamental in constructivism (Buhagiar, 2007).

2.8. Factors that influence EFL instructors’ assessment assumptions and practices

Instructors’ assessment assumptions and practices are the result of multi-faceted factors. It is argued that instructors may well articulate their assumptions/beliefs, but their practices in the classroom are influenced by several factors. In line with this, Xu and Brown (2016) stressed that instructors cannot implement their assessment-related assumptions and beliefs in a way they feel right since “they are employed within an immediate workplace community and larger social, political, and cultural contexts” (p.157). These factors could be internal, contextual or external.

According to Izci (2016) and Cheng et al. (2004), internal factors are personal factors related to instructors’ assessment assumptions/beliefs and knowledge (e.g., experience, training, and

knowledge in assessment). Other factors could be external related to the context and the external policy (Izci, 2016). Contextual factors are those related to the institution's assessment culture, students' beliefs and expectations about learning, students' motivation and proficiency levels, classroom management and routines, and context (Taylor, 2009; Scarino, 2013). Contextual factors may also include teaching/learning environment, such as class size (Cheng et al., 2004).

Instructors' experience and education could have a major role in shaping their assumptions and practices. Borg (2006) argued that school experiences have a significant impact on instructors' assumptions. In this regard, teaching experience is important factor that could influence instructors' assumptions and practices, and some studies affirmed classroom experience to have a great influence on instructors' practical knowledge and thus their activities and decisions (Borg, 2006). It is commonly reported that having more years of teaching experience positively influence instructors' beliefs and practices (Cheng et al., 2004). However, Rodriquez and Mckay (2010) pointed out that experienced instructors were resistant to professional development and held negative attitudes toward changing their assessment practices. Such experienced instructors' resistance and thoughts could result in rejecting assessment reforms if they think their current practices are right.

Education and professional development play critical role in shaping instructors assessment assumptions and practices. Through ongoing training and workshops, instructors gain a deeper understanding of effective assessment strategies, learn to recognize diverse learning needs, and adopt evidence-based practices that enhance student achievement (Brown, 2004; Borg, 2006; Oo et al., 2023). Professional development encourages reflection on current assessment approaches, challenging instructors to critically evaluate their assumptions about assessment that results in acquiring assumptions compatible with innovative practices (Ergun, 2013; Brookfield, 2013). Through further education, instructors engage with new assessment theories and methodologies, and they become more adept at creating fair, inclusive, and dynamic assessment environments that not only measure student learning but also support it.

Several studies showed factors relating to context might influence instructors' beliefs and practices (Cheng et al., 2004; Taylor, 2009; Dashti, 2019). They have examined one or more of these contextual factors that may have an influence on instructors'/teachers' practices. The context in which assessment occurs (e.g., the country, cultural, or institutional context) has been

a common topic of investigation for many researchers concerned with instructors' assessment beliefs and practices (Cheng et al., 2008; Fulmer et al., 2015; Latif and Wasim, 2022). According to Cheng et al. (2008), instructors' practices come from a combination of knowledge, beliefs, values, and experience, and these practices are influenced by the interaction of such attributes within the "instructional context in which they teach" (p. 25).

For example, Cheng et al. (2004), by emphasizing the role of context, experience, high-stakes mandates and assessment knowledge, conducted a comparative study to examine instructors' assessment practices in three different contexts. Using questionnaires, instructors from Hong Kong, Canada, and China self-reported their decision-making practices on the assessment of students' language skills and abilities. Cheng et al. (2004) showed the complex assessment practices present in ESL/EFL classrooms. They attributed the complex nature of such practices to several internal and external factors, such as the nature of the courses, the instructors' teaching experience and assessment knowledge, the teaching and learning environments (e.g., class size), and the influence of external tests on instruction and learning.

For example, instructors in Hong Kong used fewer objective assessment scores and fewer assessment purposes than did instructors in Canada and China. Further, instructors in Hong Kong did not experience stress because of the externally mandated tests, which are usually believed to dominate teaching in other contexts. In Canada, ESL/EFL students are required to pass the Test of English as a Foreign Language (TOEFL) to gain admission to a university degree program. In China, every student is required to pass mandated tests to obtain a bachelor's degree. Despite the differences in practices, purposes, and assessment methods, all instructors spent a similar amount of time on assessment, and all had similar course loads. However, instructors from Hong Kong and Canada spent more time judging and scoring than did their counterparts in China. Cheng et al. (2004) attributed this difference to instructors' obedience to the mandated testing policy and class sizes in China.

In their later study, Rogers, Cheng, and Hu (2007) obtained similar findings but chose to approach their research from the perspective of instructors' beliefs. The researchers used a questionnaire with 44 instructors in Hong Kong, 95 instructors in Canada, and 124 instructors in China and found these instructors held similar beliefs on the importance of assessment in teaching and learning, although they demonstrated different practices from instructors in the

previous study. However, the instructors' expressions of their beliefs were uncertain, mixed, and contradictory, particularly concerning the time required for assessment, the use of paper-and-pencil performance assessments, and their understanding of and preparation for assessment. They also demonstrated differences in their confidence levels in terms of applying what they had learned about assessment. However, Rogers et al.'s (2007) findings showed more similarities than differences, and the authors attributed this to several factors: whether the criteria were set by university or departmental policies or by experienced teachers, students' expectations, and the requirements for demonstrating linguistic proficiency.

More recently, Latif and Wasim (2022) investigated tertiary EFL instructors' assessment related personal theories, conceptions and beliefs. Twelve instructors from three universities participated in the qualitative study. The study revealed diversity, complexity and uncertainties in instructors' beliefs and personal theories related to various aspects of assessment. It also revealed that while tertiary EFL instructors held contemporary theories and beliefs about assessment, they were not reflected in their classroom practices because of contextual and institutional dynamics at micro and macro levels that affected their decision-making process in their assessment practices. These factors include assessment policy, assessment knowledge, and cultural, academic and professional backgrounds. Generally, the existing studies show the complex nature of assessment influenced by distinct institutional, cultural and educational policy dynamics of the social context, which affect instructors' assessment assumptions, beliefs, and practices (McNamara, 2001; Remesal, 2011; Ergun, 2013; Fulmer et al., 2015; Latif and Wasim, 2022).

2.9. Previous Studies on English Language Assessment

In addition to other educational aspects, quality classroom assessments are also proven to improve instructors' instruction and learners' English language learning (Alderson, 2005; Batchman, 2007; Brown & Abeywickrama, 2010). Thus, the relation between quality assessments and quality English language education might well bring the importance of practitioners' assessment practices and assumptions in the context of EFL to the fore. Although research in this area particularly at higher education is limited and deserves much more attention (Price et al., 2014; Nguyen & Walker, 2016; Dawson, et al., 2019), research studies to date have yielded significant results and implications regarding several issues in assessment.

Some studies emphasized the practice of continuous assessment in EFL classrooms in different contexts. One of the major findings common to these studies is that EFL instructors do not carry out continuous assessment appropriately. For instance, Mebea (2008) carried out his study in order to find out EFL instructors' and students' practice and attitude of continuous oral assessment. According to data from ten instructors and 75 students, both sample groups had positive attitude for continuous oral assessments though that was not interpreted into practice by the instructors. Students' linguistic background, absence of assessment policy and inadequate preparation were factors identified to hinder effective implementation of continuous oral assessment.

Similarly, Dagne (2009) explored the practice of continuous assessment by EFL instructors at Jimma Teachers' College. Data was gathered from twenty instructors and 141 2nd and 3rd year students through questionnaire, interview and classroom observation. The results indicated that more than 70% of teachers did not use continuous assessment to assess students' English language performance, and formal tests and exams are more considered. Lack of adequate training, large class size and workload were among the major factors hindering the proper practices of continuous assessment in EFL classrooms. The researcher finally suggested the need for further studies.

In secondary school context, Abiy (2013) also studied EFL teachers' and students' perceptions, attitudes, and actual practices of continuous assessment. Questionnaires, interviews and content analyses were used to gather data. The findings indicated that continuous assessment was exclusively used for developing students' intellectual skills and assignments were used to evaluate students' achievement of the learning objectives. The researcher concluded that continuous assessment was not properly implemented and its function to diagnose student's problems and modify the teaching was found to be very limited and even absent in most cases.

Studies were also conducted in relation to assessment methods used in EFL classrooms. Nibret (2013), for example, studied the effect of modifying EFL teacher's assessment method on students integrated approach to learning English. An EFL teacher and students in a section were participated. Checklist was used to compare the nature of the teacher's assessment before and after the training, and questionnaire was used to compare the before and during intervention trends of students' global-analytical learning dimension. The results showed that although the

teacher had relied on traditional testing methods, he began to use alternative assessment methods because of training intervention, which in turn positively affected students' approach to learning. More studies to get deeper insight about current principles of assessment was suggested.

More recently, Motuma and Chaka (2023) also assessed implementation of alternative assessment methods in communicative English language skills course at three universities. Questionnaire-based data gathered from 128 EFL instructors and 241 students revealed dominance of instructor-centered strategy in assessing students' communicative English skills. The assessment strategies employed were also reported to lack progressiveness and relevance. The study also showed that most of EFL instructors preferred traditional assessment methods to alternative assessment methods and found the former more convenient to use in the target course than the latter.

Addressing aspects such as methods, purposes and feedback, Geberew (2014) investigated secondary school EFL teachers' formative assessment practices. Based on data gathered through questionnaire and interview, the findings revealed that most of the teachers used assessment for the purpose of learning. On the other hand, teachers' limited use of assessment feedback to adjust instruction was reported. Similar to others, the researcher concluded as diligent effort is required to help teachers use assessments properly.

Specific to writing skills, Zelalem and Emily (2018) also assessed formative assessment practices and beliefs of EFL instructors at Debre Markos University. The study was also aimed to determine the relationship between the instructors' beliefs and practices. The quantitative data generated from 25 instructors revealed their positive beliefs on the importance of formative assessment, and that they practiced formative assessment averagely, leading to moderate and significant correlation. However, findings from the qualitative data, which showed that instructors were unable to practice assessment for formative purposes in writing classrooms, revealed mismatch between their beliefs and practices.

Through these local studies, portions of the current status of English language assessment could be gleaned. For instance, we partially understood that current English language assessment practices are not in line with best practices. However, these studies depicted the evidences in assessment practices by highly relying on data from instructors, neglecting the relevant

documents and views of other frontliners such as students despite their pivotal roles. They all placed instructors or teachers at the centre of attention except Mebea (2008) and Motuma & Chaka (2023) studies, which also incorporated students' views. Yet, they both failed to incorporate data from the relevant documents on the area, which however, could have enabled a better and more complete understanding about the phenomena under study.

In addition, even though certain assessment components are repeatedly addressed in the studies, assessment feedback and particularly focuses of the feedbacks teachers or instructors provide, the other essential element of the assessment practice, received little or no attention. However, as most important in classroom assessment practice is using feedback to enhance learning (Hattie & Timperley, 2007; Fulcher, 2010), what the provided feedbacks emphasized worth an investigation (Shute, 2008; Brooks, Carroll, Gillies & Hattie, 2019; Gan et al., 2021).

Furthermore, while subject specific assessment studies are encouraged, most of the studies have emphasized on general ELT. In particular, none of them investigated assessment issues specific to communicative English language skills course except Motuma and Chaka (2023) study, which was even limited to only the method aspect of assessment practice. In addition, Yitayal and Amare (2023) also conducted their assessment study specific to the target course although their focus was on instructors' assessment literacy. Holroyd (2003) argued that there would be subject-based differences in assessment practice and thinking. Cumming (2001) study also revealed that, based on specific purpose of the courses, EFL university teachers' assessment beliefs and practices varied in line with particular course they teach. Shrestha (2014), in this regard, argues that the assessment practices of communicative English skills course are different from the assessment practices in most specific language courses, including productive and receptive language skills. This suggests the need to conduct assessment study specific to course.

The studies also seem to have given less attention to the relation instructors' assessment has with their background characteristics or profile. Literature shows that instructors' assessment cognition and behavior are the reflections of their contexts and experiences, and these cognition and behavior are influenced by the interaction of such attributes (Scarino, 2013; Xu & Brown, 2016). Particularly, instructors' demographic characteristics such as sex, experience, and academic level play significant role with regard to their assessment practices and assumptions (Gipps, 1999; Ergun, 2013; Xu & Brown, 2016; Gan et al., 2021). Generally, it can be argued

that there is limited study to suggest which individual differences are more often found in instructors, especially, that possesses traditionalist and alternativist assessment assumptions. Hence, it is essential to carefully diagnose the factors in relation to assessment practices and assumptions, as understanding this would be instrumental in tailoring assessment instruction for the on-the-job EFL instructors.

On top of these, as far as the researcher's knowledge is concerned, studies are lacking regarding assumptions underlying current assessment practices in higher education EFL context, particularly in communicative English language skills course despite the growing recognition that assumptions greatly influence the success of assessment implementation. As a result, the assumptions underlying assessment practices in the course are not yet clear, and our knowledge of this area is very limited. This area also seems to gain little attention at global level, as the target of most studies was on general education and/or lower educational levels (e.g., Delandshere, 2001; Klenowski & Willis, 2011; Peterson & Bainbridge, 2012; Mitana, 2018), remaining a knowledge gap regarding assessment assumptions in HE EFL context.

This gap is considerably problematic given the power of assumption in determining what is assessed, how it is assessed, how learners are perceived and treated, and instructors' positionality. For example, it is more difficult and frustrating for English language instructor to adapt learners-involved assessment practice when he/she assumes that assessment is a mechanism for making learners accountable for their learning, as this new method does not fit within the instructor's current world view. So, the instructor is more likely to interpret the method as extra work rather than as effective way to enhance learning.

Thus, the situation renders futile the ongoing efforts to redress the EL assessment problems in Ethiopian universities. This is because EFL practitioners' assessment assumptions should be brought out of the closet and aired for debate in order to examine their validity in the light of current ideas of assessment and learning either to accept as compatible or reject as unjustified: the process that results in fostering meaningful change in assessment enabling critical reflection.

Four research gaps are thus identified from the review. First, there is a need to incorporate multiple sources of knowledge in studying assessment practices to gain a better and more complete understanding. Such sources provide the conditions for systematic data collection on

assessment policy implementation (Lund, 2008; Kezar, 2012). The present study, thus, fills the gap in the literature by incorporating voices of EFL instructors and students as well as data from existing relevant documents. Second, what instructors' assessment feedback focuses on worth an attention, as it is most important in classroom assessment practice to enhance learning. This vital issue, yet which remained a gap in the studies, is also emphasized in the present study. Third, there is a need to delimit the study of assessment to specific course rather than collectively treating assessment in ELT, as subject-specific assessment study seems desirable for the validity matters. Thus, by investigating participants' assessment assumptions specific to communicative English skills course, the current study is initiated to fill the existing gap. Besides, more evidence is needed about the assessment practices in the target course so as to come to a firmer conclusion. Finally, the present study is also initiated to shed light on the assumptions underlying the practices of assessment in the target course and fill the gap in the literature.

2.10. Theoretical Orientation

The current study positions itself within social constructivism. According to Shepard (2000), the constructivist model should hold sociocultural theory, cognitive theory and constructivism. He, thus, expounded the constructivist model to the theoretical framework of social constructivism to include the cognitive, sociocultural and constructivist models. Accordingly, based on the social constructivist model, he states that learning is “an active process of mental construction and sense making” (Shepard, 2000: 6), and that both the learner's cognitive ability and the context should be emphasized. In this study, thus, classroom assessment is seen from a social constructivism theory that involves the social and cultural contexts of assessment as well as the cognitive aspects.

According to social constructivists, instructors mediate students learning through assessment (McNamara, 2001; Rust, O'Donovan & price, 2005). That is, assessment is an integral part of and central to teaching and learning, and this is the main focus of classroom assessment, which is placed primarily to improve learning. The purpose is not just to measure how much knowledge or skill is remembered but rather to assess how information is structured and how much the level of students' knowledge or skill has changed throughout the learning process (Lund, 2008). Hence, traditional examinations at the end of chapters or courses are less valued since they do not allow capturing the students' knowledge changes throughout the learning process. As a

result, assessment is recognized as ongoing process that helps the students continue to learn (Shepard, 2000; James, 2006).

The social constructivist approach encourages EFL instructors to assess learners using a broader range of assessment methods, as this allows capturing all the important learning goals and aligning assessment to instruction, which is impossible for limited assessment types (Shrestha & Coffin, 2012). This suggests using multisource method that promotes learners' language learning and enables understanding the actual progress. Alternative assessment methods in the form of presentations, debates, projects, portfolios, creative expressive activities and so on, which allows observing and demonstrating what one can do with the knowledge and skills he/she has acquired, represent this type of assessment (Buhagiar, 2007; Katz & Gottlieb, 2013)

The theory also emphasizes the importance of feedback, as it views assessment as formative process that encourages students to learn and relearn continuously. Based on its key principle that learning results from our reflections on feedback from environmental interactions, feedback is central in assessments informed by social constructivism (Rust et al., 2005; Shepard, 2006). Thus, it promotes assessment practices that value revisions and encourage students to continuously construct and reconstruct their knowledge to evolve and change their understanding in response to feedback (Black & Wiliam, 2009). Such practices enhance students understanding by involving them in reflective activities where they consider the strengths and weaknesses of their learning and make plans for subsequent actions (Shepard, 2000).

Classroom assessment also comprises social, collective and contextual dimensions. The social constructivist model, which requires a complicated assessment process to assess the deep understanding and learning of students, involves the engagement of both students and instructors (Gipps, 1999; Rust et al., 2005; Lund, 2008). Classroom assessment is a collective and collaborative activity in which both instructors and students take part and have essential roles in the process (Shepard, 2000; Lund, 2008). Hence, these collective and collaborative nature of assessment underpinned in social constructivism may allow me in understanding the current study's research problem by emphasizing not only on EFL instructors but also their students.

Assessment in social constructivism engages learners and places them at the center of learning by actively involving them in the assessment process (Rust et al., 2005). It should also be tailored to the individual's needs and should be constructed in a way that recognizes the social

context in which the assessment is used (Lund, 2008). That is, assessments that enable learners to construct knowledge based on their experiences and interactions with others and that provides opportunities for multiple perspectives should be designed, rather than that merely assess whether students have a set of pre-determined knowledge. Shepard (2000) also illustrated that the social constructivist approach to assessment enables learners to construct knowledge by working together.

The importance of social and cultural contexts in the conceptualization of practitioners' assessment assumptions, beliefs and practices has also been resonated in recent documents (Scarino 2013; Stabile, 2014; Xu & Brown, 2016; Latif & Wasim, 2022). They argue that practitioners' assessment related traits are the result of their interaction with the social and cannot be separated from the context in which they occur, suggesting their importance in assessment studies. That is, practitioners' assessment assumption and practice are both social and contextual, shaped and influenced by several factors including background characteristics and institutional and contextual dynamics. This theoretical insight, thus, serves as a rationale for this study, first, to situate the research problem into Ethiopian HE EFL context and specifically into communicative English skills course context. It also enables me to understand the participants' assessment assumptions and practices in relation to their background characteristics, and to understand the connection between instructors' assumptions, practices and contextual factors.

2.11. Conceptual Framework of the Study

The present study's focus on classroom assessment in EFL context lies in the premise that quality assessments have decisive role in improving students' EL skills. That is, if the current research-led assessment developments and thinking are well considered in the assessments of particularly EL improvement courses, there will be sound instruction and learning. As a result, enhanced learners' English language communication skills is expected. That is why significant shifts and changes in both theory and practice of assessment have been made.

However, little is known whether the existing assessment practices align with the new assessment developments and the assumptions have shifted with current perspectives in assessment and learning, as efforts to assess how assessment operates in higher education, particularly in the communicative English skills course, seem very limited. This may cripple the

ongoing efforts to improve the quality of assessments. Hence, this is the driving force behind the present research, which tries to shed light on the practices and assumptions of assessment in the enactment of communicative English language skills course to understand how assessments are contributing to students' EL learning.

The conceptual framework of the study comprises four dimensions. To begin with, assessment practice is conceptualized as any activity or task practitioners carry out to gather information on learners' learning in order to both facilitate learning and measure progress and performance. Although it embraces a broad spectrum of activities, its component in this framework denotes how assessments are carried out in terms of the purpose, methods, and feedback, which are also suggested to be the key components in the classroom assessment process routes (Rea-Dickins, 2001; Hill & McNamara, 2012). As discussed earlier, considering the substantial role of effective assessment practices in enhancing learners' language achievement, an attempt to understand the practice of assessment in EFL classrooms is significant.

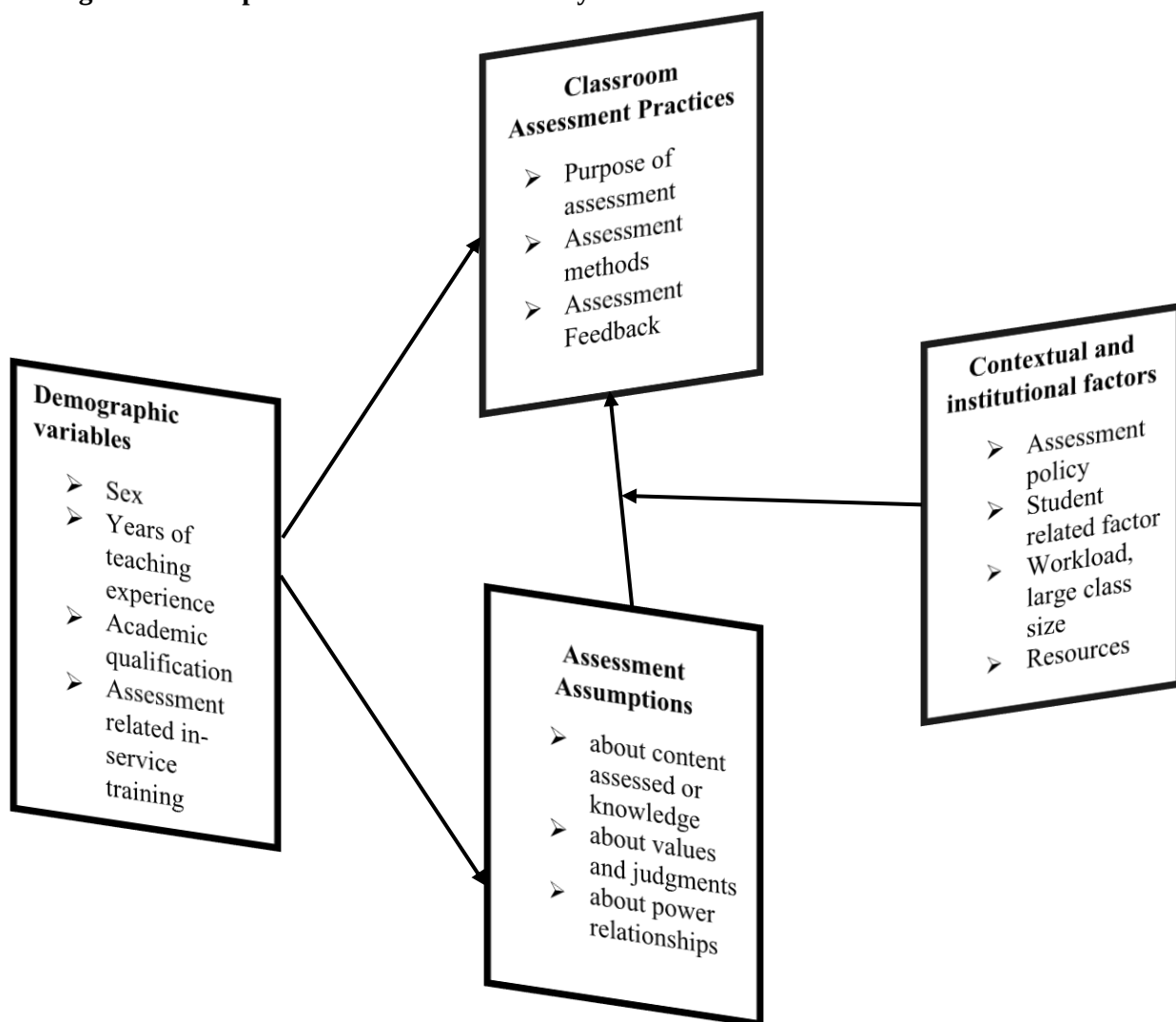
However, implementation of quality assessments is highly dependent on the practitioners' assumptions since their assessment related decisions and practices are usually shaped by their assumptions (Ergun, 2013; Wellberg & Evans, 2022). That is why assumption became the other dimension in the conceptual framework. Assumptions are propositions practitioners make about assessment that are taken for granted as a basis of argument or action, molded by their past experience and learning. Practitioners hold diverse assumptions about assessment that can be aggregated into two broad categories: traditionalist and alternativist. EFL assessment practitioners are expected to possess assumptions that keep pace with current language learning and assessment theories than those rooted in the traditional assessment culture so that the intended changes in students' English is attained by improving the quality of its assessments. Thus, exploring the assumptions upon which the existing assessment practices in the target course rest on is substantial.

The final dimension in the framework, contextual factors, withholds the view that assessment is context dependent. The current conceptualization of classroom assessment also emphasizes the need for a broader view of assessment within sociocultural, political, and historical frameworks to better understand why assessments are perceived and practiced the way they are (Gipps, 1999; Scarino, 2013; Xu & Brown, 2016). Particularly, teachers' demographic factors are often

reported to shape their assessment cognition and action (Dashti, 2019; Gan et al., 2021). This suggests the need to carefully diagnose these factors in relation to assessment. In this study, the relation between instructors' some selective background profile such as sex, teaching experience, and academic qualification and their assessment assumptions are examined.

This framework has also considered factors mediating between instructors' assumptions and practices to have a clear understanding regarding the underlying reasons for any discrepancy between instructors' assessment assumptions and practices.

Figure 3. Conceptual framework of the study



CHAPTER THREE

THE RESEARCH METHODOLOGY

3.1. The Research Paradigm

Pragmatic paradigm was adopted for the current study since the researcher believed that the phenomenon under study is complex that needs multipronged approaches. Pragmatic paradigm allows using any available research approaches and methods as far as they work best so that the study is fully addressed. Thus, the paradigm deemed appropriate for this study since it is suitable for researchers whose studies can neither be satisfied through positivists nor interpretivists.

The use of pragmatic paradigm was also to bring more valid and reliable findings by drawing from the strengths and minimizing the weaknesses. Applied linguists also acknowledge the use of different research methods since insight into complementary realities is gainful and so that the scope of research in applied linguistics is developed (Dornyei, 2007; Tsushima, 2015). Likewise, the nature of the current study by itself necessitated identification of complementary realities, which was considered positive for the reliability and validity of the interpretations. Hence, by utilizing objective approach such as survey questionnaire, the study sought objective reality on assessment practices and assumptions, the researcher being independent and separate. Side-by-side, respondents were also invited to explain their views in order to allow emergence of different subjective realities based on their lived experiences. Adopting both qualitative and quantitative research approaches, thus, appeared to be an appropriate methodological approach for this study.

To put in a nut shell, there are some aspects of pragmatism which appealed to the researcher that are believed suit the nature of the current study. First, the fact that both quantitative and qualitative approaches and methods can be used in single research fascinated the researcher to adopt the paradigm. This was to consider the full spectrum of approaches and methods available in order to maximize the research validity and since a focus on either approach cannot address the research questions in full breadth and depth. Second, the paradigm was adopted since the researcher gave value for both subjective and objective knowledge and, finally, since it allows abductive reasoning to move back and forth between deduction and induction.

3.2. Research Design

Mixed methods research design was adopted for the current study bearing in mind its suitability to the philosophical framework adopted and considering the research questions of the study. The design was also preferred since it allows completing the study with more depth and breadth by combining procedures with qualitative and quantitative research (Cohen et al., 2007; Tsushima, 2015). Therefore, within this design, the study employed quantitative and qualitative research approaches to obtain a better understanding and provide alternative perspective on assessment practices and assumptions underlying in higher education EFL context.

Three factors should be taken into consideration when designing mixed method studies: priority, implementation, and integration (Onwuegbuzie & Collins, 2007; Creswell, 2014). In the current study, priority or more weight is given to the qualitative approach than the quantitative. As to the implementation, the collection and analysis of quantitative and qualitative data were made concurrently, at approximately the same time. Finally, regarding integration, the qualitative and quantitative data were combined at the phase of interpretation.

Accordingly, a design ‘QUAL + quant’ was used, where the relative dominance of the qualitative data is represented by the capital letters and the plus sign shows the concurrent collection of both the qualitative and quantitative data. That is, the data collection for both approaches is made at approximately the same point in time, and as such, the study employed concurrent research strategy. This strategy enables more complete understanding of the research problem since one data collection form supplies strengths to offset the weaknesses of the other. It also allows comparing data on the same phenomenon, or validating, confirming or corroborating results of one form with the other. In other words, to end up the study with more valid and well-substantiated conclusions about assessment practices and the underlying assumptions in communicative English skills course, adopting concurrent strategy within mixed methods research design was gainful.

3.3. Population of the Study

In this study, it was assumed that English language assessment is collective responsibility of different stakeholders. Thus, unlike the common research practices that confine assessment

studies to instructors or teachers, the current study honoured and incorporated voices of both EFL instructors and students.

Accordingly, the population of this study comprised EFL instructors and students in the public universities located in western cluster, where the study was conducted. The area of the study was selected based on convenience. That is, the universities geographical proximity and accessibility to the researcher was considered in choosing the area. There were five universities in the western cluster, namely: Asosa University, Jimma University, Metu University, Mizan-Tepi University and Wallagga University. However, the later emerged Dambi Dollo and Gambella universities are also located in the study venue; this made the number of universities in this cluster to reach seven. Since the study was to understand the practices and assumptions of assessment in communicative English skills course, all EFL instructors and students teaching and learning the target course in the universities located in the study venue were population of this study.

3.4. Sample and Sampling Techniques

In the case of the current study, sample is conceived as the number of individual participants selected systematically from the population. It also referred to the number of universities chosen from which individual participants of the study are selected. Generally, four universities, 84 EFL instructors and 389 students were samples selected for the current study.

First, the four universities from which participants are selected were identified systematically. According to ministry of education, the seven public universities in the study venue are categorized into first, second, third and fourth generations. While one university (Jimma) was categorized under first generation, two universities were categorized under each second generation (Mizan-Tepi and Wallagga), third generation (Asosa and Metu), and forth generation (Dambi Dollo and Gambella). This categorization was taken into account in selecting participant universities to ensure the representation of universities of different generations in to the study. Thus, one university from each generation was included into the study using simple random sampling technique except for the university in the first generation that was automatically included, as it was the only in the category. Accordingly, Dambi Dollo, Jimma, Metu, and Wallagga universities were identified. However, due to different problems, particularly the

security issue and political unrest during data collection on the area, the researcher could not reach Dambi Dollo University.

The next step was selecting individual participants. Since this study operated within mixed methods design, sampling techniques and process used to select the participants for the quantitative and qualitative data were made distinctly as follow. Regarding sampling of instructors, there were only 84 EFL instructors teaching communicative English language skills course in 2022 GC academic year in the three universities participated in the study. Selecting a sample from such a small target population was a bit unimpressive for conducting sound quantitative analysis. Hence, all the instructors were invited to participate in the study using comprehensive sampling technique. To this end, while 39 EFL instructors were from Jimma University, 26 and 19 EFL instructors were from Wallagga University and Metu University, respectively.

On the other hand, multi-stage sampling technique was used to select the student participants for the survey questionnaire. This technique was used as it was not possible to get accurate number and complete list of first year students taking communicative English skills course in the academic year in the sampled universities. First, based on prior categorization of the students into natural and social science strands in the universities, sample of sections were selected from each strand in each university. There were 24, 16, and 21 natural science sections and 15, 8, and 11 social science sections in Jimma, Wallagga and Metu universities, respectively. Accordingly, 25% of the sections in each strand were selected from each university using simple random sampling technique. This resulted in identifying 15 natural science and 9 social science sections in general. Next, sample students were selected from the 24 sampled sections. In this regard, 35% of students in each section were selected from each university again using simple random sampling technique. The use of simple random sampling technique to select both sample sections and students was to ensure that each section and student has an equal and independent chance of being selected.

Table 3 Samples and sampling of the participants of the survey questionnaire

University	Number of Instructors	Number of students					
		Natural			Social		
		Total section	Sampled section	Sampled students	Total section	Sampled section	Sampled students
Jimma	39	24	6	86	15	4	54
Wallagga	26	16	4	78	8	2	38
Metu	19	21	5	81	11	3	52
Total	84	61	15	245	34	9	144

As to the number of participants selected for qualitative data, basing different criteria, several alternatives, which sometimes contradict each other, are presented in the literature. For instance, regarding the number of interviewees to be included, Onwuegbuzie & Collins (2007) suggest ≥ 3 interviewees per sub-groups and ≥ 12 interviewees in general. Dornyei (2007) commented 6-10 might work well; and for Creswell (2012), it can range from one to multiple cases depending on the type and design of the study. This makes determining the number of participants for interview schedule a bit difficult. However, it is generally accepted that sample for qualitative data should not be too small as to not prevent data saturation and too large not to render data analysis and interpretation complex and result in superficial perspectives. Hence, basing these descriptions, nine EFL instructors (three from each university) and twelve students (four from each university) were sampled for the interview schedule. This number can be acceptable since the depth of insights about the phenomenon of the study was needed in the qualitative study and the number was manageable for postgraduate solo researcher.

The main goal of sampling in qualitative study is to find individuals who can provide rich and varied insights into a particularly selected phenomenon so as to maximize what we can learn, and this is best achieved through purposive sampling (Denzin & Lincoln, 2005). Accordingly, purposive sampling technique deemed appropriate for this study and used to select a predetermined number of participants using predefined selection criteria for the interviews. Thus, particularly using maximum variation sampling technique, individuals with diverse range of backgrounds and that are all relevant to the phenomenon under study were targeted to ensure the diversity of the different groups. The technique was also chosen since the researcher wanted to gain greater insight into the phenomenon under study by looking at it from all angles.

Thus, mixed sex groups, a range of teaching experience, academic qualifications and so on were taken into consideration for the selection of the instructor participants. As there were few female instructors in the universities, the majority were males (N= 7) and females (N= 2) with teaching experience ranging from 4 to 19 years. The majority of them were master's degree holders (N= 6) while Ph.D. holders were (N= 3). TEFL graduates comprised the majority (N= 7) whereas the rest were non-TEFL graduates (N=2). Similarly, students' varied backgrounds that include mixed sex groups, different age ranges and field of study were considered while selecting. Accordingly, almost equal number of females (N= 5) and males (N= 7) found in the age between 17-22 years and attending natural science classes (N= 6) and social science classes (N= 6) were selected.

3.5. Data Collection Methods

In accordance with the study design, interview, questionnaire and document analysis methods were used to collect the qualitative and quantitative data for the study.

3.5.1. Interviews

Interviews were used as the main data collection instrument in this study. The reason for this relates to lack of validated standardized questionnaire for exploring assumptions about assessment in EFL context. In such circumstances, emphasizing on qualitative approach is necessitated (Denzin & Lincoln, 2005). The reason also relates to the study's aim to gain an in-depth understanding of the phenomenon under investigation. Hence, open-ended interviews were used to allow in-depth interpretations and better understanding of the research problem. Specifically, in-depth semi-structured interviews were chosen considering their advantages, the type of data they gather, and the researcher's current knowledge and skills.

Semi-structured in-depth interview offers a good opportunity for the interviewer to manage the interview schedule easily, help exploring at a deeper level, and allow flexibility to raise probing questions that enable gaining more insights into the phenomena (Dornyei, 2007; Denscombe, 2007). Although guiding questions were prepared in advance, the researcher could ask additional probing questions to gain a deeper understanding of assessment practices and assumptions. In this regard, the use of open-ended questions was essential to fully and richly collect data from the participants by allowing the interviewees to freely express their experience of assessment and

reasoning behind assessment practices. Overall, to leverage these advantages, in-depth semi-structured interviews were employed in this study.

The interviews were divided into two sub-themes: assessment practices and assumptions. The first set of interview questions directed to eliciting more and deep information about assessment practices in communicative English skills course. In line with the research questions, the questions were developed based on the purposes for which instructors conduct assessment, methods used to assess students, and assessment feedback provisions.

The second set of interview questions was prepared to elicit the participants' reasoning behind their assessment views and practices in order to uncover their underlying assumptions. According to Brookfield (1992, 2013) and Gabbitas (2009), peoples' assumptions about assessment can be discovered by asking questions such as 'Why did you do this?' and 'why do you think that?'. Accordingly, the question schedule with which the interviews were guided was organized into two sections, the first being focused upon specific assessment set and used by the participants, and the second upon general issues related to assessment.

The first interview section targets only instructors, as it focused on typical assessment practices they selected and used. Instructors were asked to provide reasons for the assessments they used. This is also to mean that the questions prepared were presented in integration with the assessment practice questions; they were not designed separately. Thus, by focusing on typical assessment practices related to methods, techniques, feedback, purpose, and so on the instructors utilized, the interview questions asked them why they practiced the way they did in order to elicit descriptions of thinking that informed their assessment practices. The purpose of this line of questioning was to elicit assumptions grounded in practices. The second section of the interview that covered questions that focus on general assessment issues was developed for both instructors and students. The questions ask their views in terms of students' engagement and roles, factors that determine what to/not to assess, good and bad students' work, universal and sociocultural reflections of assessment and good and bad assessment experiences to uncover their underlying assumptions.

The interviews were conducted in face-to-face approach, as this enables to better understand the responses of the interviewees through both verbal and non-verbal communications (Patton,

2002). All the interviews with instructors were conducted in English because these EL instructors can communicate their views using the language and to use the advantage of overcoming the reliability and validity problems that come with translation. However, for students, the interviews were conducted in their native language so that they can easily and fully express their views. The interviews took from 43 to 66 minutes and all were audio recorded with the permissions granted.

3.5.2. Questionnaire

Questionnaires were also used to collect information on participants' assessment assumptions and practices because they are efficient in measuring factual and behavioral assessment information (Dornyei, 2007). The use of questionnaire was essential to generalize the study from the sample to the wider population and to make inferences about EFL instructors' and students' assessment assumptions and practices in the target course. The questionnaire had three parts: assumptions about assessment, assessment practices and demographic information.

Assumption about assessment questionnaire was designed to explore assumptions underlying assessment practices in communicative English skills course. As it was not possible to find appropriate instrument for this purpose, the researcher developed a questionnaire used to assess EFL instructors' and students' assumptions about assessment. First of all, an extensive literature review was conducted to identify the dimensions and subdimensions to be included. Accordingly, Anderson (1998) categorization of assessment assumptions into traditional and alternative assumptions was considered. He identifies, contrasts, and describes assumptions upon which traditional and alternative assessment theories are constructed on some interconnected factors. His typology provides a useful framework to explore assessment assumptions. Adopting this framework, factors widely and commonly discussed in the relevant literature such as content assessed/knowledge, value and judgement, and power and control were emphasized in developing the scale.

Next, the initial item pool was created by collecting statements reflecting assumptions that underly both traditional and alternative assessments from relevant literatures (e.g., Biggs, 1995; Anderson, 1998; Teasdale & Leung, 2000; Delandshere, 2001; Buhagiar, 2007; Klenowski & Willis, 2011; Fulcher, 2014; Wellberg & Evans, 2022). Multi-item scales, a cluster of several differently worded items that emphasized on the same target, were developed to increase the

stable component the items share and to minimize the extraneous influences unique to individual items. The item pool created was critically examined several times by the researcher, and a number of changes and modifications were made including the reduction in number of items. After that, by posing the scale to educational psychologists, EFL instructors, TEFL PhD student, and finally, to the advisory board, all items were subjected to content and face validity. After rephrasing, replacing, adding or removing items, finally 20 items (ten reflecting assumptions underlying traditional assessment and the other ten reflecting assumptions underlying alternative assessment) were retained.

The second part of the questionnaire, assessment practice, was adopted from various studies. The questionnaires adopted were chosen from the existing based on some quality criteria. For example, questionnaires specifically designed in EFL university context and those developed with multiple items for each of the factors were prioritized to have instrument with greater validity. In addition, validated instruments having greater alpha coefficient (0.7 or above) were selected while adopting.

The assessment practice questionnaire has three sections. The first was concerned with the purposes for which assessments are conducted. This was adopted from Gonzales and Aliponga (2012) questionnaire that comprises five factors. For this study, three factors namely: assessment for learning, of learning, and as learning were adopted. ‘Assessing to inform’ factor, whose purpose is to provide information to parents about the performance of their children, was excluded since this is not applicable to university. In addition, since both assessment for learning and assessment for instruction are all about facilitating instruction, they were merged into assessment for learning factor. The determination was also based on how most literature on the area generally categorizes assessment purposes (e.g., Earl, 2003; Saefurrohman & Balinas, 2016).

The second section focused on assessment methods. A questionnaire designed by Ali, Khaled, and Mohd (2022) to measure university EFL instructors’ frequency of use of various assessment methods was adopted. It was made up of close-ended items reflecting both traditional and alternative assessment methods. The final section of the assessment practice questionnaire was concerned with assessment feedback provisions. The questionnaire developed by Guo (2020) and Gan et al. (2021) to measure how frequently EFL instructors provided various types of feedback

was reviewed. For this study, verification, directive, and scaffolding scales were adopted with minor modifications. In addition, another scale with four items that related to sources and modes of feedback was developed to measure how frequently instructors and students participated in the provisions of assessment feedback and how frequently oral and written modes are employed. The final part of the questionnaire was intended to collect information about participants' demographic data including sex, years of teaching experience, academic qualification, and number of students in a class for the instructors, and sex, age, and field of study were emphasized as to the students.

Description of the questionnaire

The first part, assumptions about assessment questionnaire measures the participants' assessment assumptions extent of agreement on a five-point scale (from 1 being strongly disagree to 5 being strongly agree). On the other hand, in the assessment practice questionnaire, the respondents had to specify the frequency of use of various assessment practices related to purpose, methods, and feedback using a five-point Likert scale, ranging from Never (1) to Always (5).

3.5.3. Documents

Document analysis is the other data gathering technique used in the current study to understand how assessments are practiced in communicative English language skills course. Accordingly, course outlines or course plans and assessment packages (quizzes/tests, midterm and final exams, assignments) were included in the analysis.

Course outlines of communicative English language skills used in the semester of data collection period were gathered and inspected to examine instructors' practices of assessment in terms of the types of methods they employed. A total of eight course outlines used in the three universities during the semester were collected from the interviewed instructors. On the other hand, assessment papers over one semester were collected from the interviewed instructors to investigate their practices of assessment feedback provisions. Accordingly, one assignment, one quiz/test, and one major exam were gathered from each interviewed instructors, except for three instructors who submitted only two items each. As such, a total of twenty-four assessment papers, which were administered and corrected, were collected.

Documents can be analyzed as a complement to other qualitative data sources as a means of triangulation in the study of the same phenomenon (Denzin & Lincoln, 2005). In this study, likewise, data from documents had a subsidiary role to the interviews and were used to corroborate and augment data from instructors' interviews, thus make the findings more trustworthy.

3.6. Validity and Reliability of the Instruments

Different procedures have been employed to ensure evidences of these validity and reliability qualities in the instruments used in this study. In the case of semi-structured interviews, the issue of validity can be addressed by trying to ensure that the contents of the interview questions are directly related to the research objectives (Gray, 2004). Likewise, the interview questions prepared on the basis of the relevant reviewed literature were further validated by educational psychologists, EFL instructors, TEFL PhD students, and the advisory board. Validation form, which emphasized on 1) relevance of the content of the questions in fully answering the research questions, 2) whether the questions are clear and easy to listen and to understand, and 3) if the instructions are unambiguous, was developed for this purpose. Subsequently, some modifications were made to the interview questions developed. Developing the interview schedule with some participants of the study also allows strengthening the validity of the questions. Accordingly, two EFL instructor colleagues who have familiarity with the topic were invited to the preparation of the interview schedule and refinements of the questions.

Conscious or unconscious bias the interviewer may induce in the interview procedures can also endanger the reliability of interviews. Differential treatments of subjects with respect to, for instance, the length of time given and asking questions out of sequence can lead to this problem. In this regard, based on Gray (2004) suggestion, the same written interview guide in which questions were asked in the same sequence were prepared for all interviewees of this study. They all were also given sufficient time to allow them express their views adequately and freely. Further, the interview schedules were pilot tested prior to using them for the main study. Accordingly, some adjustments were made to the procedures of conducting interviews based on the insights gained.

Validity of the questionnaires was also ascertained using different strategies. In the case of the developed assumptions about assessment questionnaire, appropriate and adequate numbers of items were included for each of the factors. So, the instrument was made to comprehensively cover and represent the content of interest. As to the assessment practice questionnaire, there were certain issues considered as principles while searching and adapting the appropriate questionnaire in order to have instrument with greater validity. For example, questionnaires specifically designed in EFL university context and those developed with multiple items for each of the factors were prioritized. This was to establish the content validity so that the data collected appropriately and fully answer the research questions. Face validity of the questionnaires was also established by paying attention to the layout such as, fonts, spacing, and quality of the paper, that significantly affect the quality and quantity of data collected. Further, the questionnaires were validated by different people, and useful insights were gained regarding their layout, content validity, and clarity of the items and instructions. Hence, the questionnaires were revised and potential problems were illuminated based on the reviewers' comments prior to using for the pilot study.

To make questionnaire items generate the data needed for the study, it was also necessary to obtain some evidence of reliability. In this regard, the pilot study conducted allowed the researcher to estimate the internal consistency of the items used in assumptions about assessment questionnaires. Internal consistency of the items was found to be 0.71 for the instructors and 0.69 for the students, suggesting that the items of the questionnaires had acceptable Cronbach Alpha coefficient level. As to the assessment practice questionnaire, instruments having Cronbach Alpha coefficient of 0.7 by rule of thumb or higher were adopted in order to maintain reliable questionnaire for the present study. The reliability of the assessment practice questionnaire was further checked on the pilot study participants, which indicated that all components of the questionnaires had acceptable internal consistency. These reliability coefficients are reported in detail in the pilot study section.

3.7. Pilot Study

3.7.1. The rationale for conducting the pilot study

As problems may occur when collecting data using unchecked instruments, researchers are recommended to test the instruments they use to forestall potential problems (Dörnyei, 2007;

Creswell, 2014). Dörnyei (2007) also states that carrying out a pilot study provide researchers advance warning about pitfalls and helps calibrate research protocols and the proposed methods or instruments. Hence, the main purpose of the pilot study in the current study was to try out the methods of data collection and analysis procedures. That is, the pilot was conducted to examine the appropriateness of the instruments prepared and overall procedures, and to identify problems or constraints, if any, and make the necessary revisions before they were used in the main study. Further evidences regarding the appropriateness of instruments and procedures of data collection and analysis were gained.

3.7.2. Method

The pilot study was conducted at two brunch campuses of Wallagga University, Gimbi and Shambo campuses, from December 23 to 29, 2022. There were only twelve EFL instructors in both campuses, and they all were included in the pilot study using convenient sampling technique. On the other hand, 48 students were randomly selected and taken as sample for the study. As to data gathering tools, questionnaire and interview were used to gather data for the pilot study. Both instruments were intended to gather data on instructors' and students' assumptions and practices of assessment in the course communicative English language skills.

After rephrasing, replacing, adding or removing items based on feedback from various categories of reviewers, generally a questionnaire with sixty-nine items reflecting assumptions about assessment and assessment practices were retained and pilot tested. There was a high need to pilot assumptions about assessment questionnaire, as it was newly developed by the researcher. With regard to assessment practice questionnaire, as mentioned earlier, some items were modified and developed, and it was applied to a different context. As such, it was necessary to trial the entire survey. So, a twenty-three-item assumption about assessment questionnaire and a forty-six-item assessment practice questionnaire were administered to both EFL instructors and students. Except for some minor differences, both questionnaires prepared for instructors and students were almost identical and designed to collect similar data. This assisted the researcher to easily compare if the responses of instructors and students matched or not. Apart from asking to fill in the questionnaire, the participants were also requested to comment on any difficulties they might experience in terms of understanding the instruction and questions, terminologies used, and any opinion they might have about lay-out presentations.

Since interview was the main data source of the study, it was very essential to trial the interview schedule. Therefore, as also suggested by Dörnyei (2007), a trial run was conducted to identify some potential problems before conducting the actual interviews. The semi-structured interviews were intended to gather data on instructors' and students' assumptions and practices of assessment specific to the target course. To pilot the interview schedules, three EFL instructors and four undergraduate first year university student volunteers were interviewed. These participants were chosen because they were instructors and learners of communicative English skills course; so, they had similar characteristics with the participants of the main study.

Once the data were collected, the next step was analyzing them. While the participants' data collected through interviews were analyzed qualitatively, the data collected through questionnaires were analyzed quantitatively. The qualitative data elicited from interviews, which were to depict assessment practices and the underlying assumptions, were transcribed and read several times until information under the predetermined codes of the two major categories were codified and clearly recorded. The qualitative data used to decide on EFL instructors' assessment practices were organized and clustered focusing on the predefined themes (assessment purpose, method, and feedback) and their respected sub-themes. On the other hand, the data that reveals participants' assessment assumptions were clustered according to traditionalist and alternativist themes. Regarding the analysis of the quantitative data gathered through questionnaire, first, the data were collated, tabulated and entered into computer. SPSS software was used to generate results that describe participants' assessment assumptions and practices on each category. The data were then analyzed using both descriptive and inferential statistics.

3.7.3. Lessons learned from the pilot study

The main aim of the pilot study was to find out, among others, the weaknesses observed on the data gathering tools and their strengths so that the necessary measures would be taken in improving the weaknesses. As such, measuring the reliability of the questionnaires used was one of the ultimate tasks of this pilot study. With regard to the assessment assumption questionnaire, the reliability coefficient, which was found to be 0.71 for the teachers and 0.69 for the students after removing some items, suggests that the items of the questionnaire had acceptable internal consistency. This also implies that the data collection instrument was appropriate and effective to investigate assumptions EFL instructors and students hold about assessment.

The reliability coefficient for a forty-six item classroom assessment practices questionnaire also showed acceptable internal consistency. Generally, the reliability coefficient which was calculated to be greater than 0.7 for all components of the assessment practice questionnaire for both participant groups indicated that this questionnaire is appropriate for data collection and has the capacity to provide reliable data. The tabulated summary of reliability coefficient of the assessment assumption and practice questionnaires are illustrated in the following table.

Table 4. Reliability Coefficient Scores on assessment practice and assumption

		Cronbach's Alpha		
	Components	N of items	Teachers	Students
Assessment practice	Purpose	14	.957	.809
	Method	14	.813	.858
	Feedback	18	.752	.710
Assessment Assumption	Assumption	20	.706	.690

This pilot study was also to ensure that there remained no ambiguity or misunderstanding in the questionnaire. So, as discussed earlier, the participants were asked to provide their comment on the questionnaire alongside to fill in. Accordingly, a few clarifications related to the instructions and the wording of some items were needed, and these were dealt with after taking the participants' suggestions into account. For instance, expression of gratitude for the participants for their support, which the pilot questionnaire did not include, was made to be considered to the questionnaire used in the main study. Based on their suggestions, the word "gender" was also replaced by "sex" taking Ethiopia's socio-cultural context into account.

Replacement on few items was also needed. In the first part of the questionnaire, for instance, an item included to elicit participants' assumptions on content assessed theme was found to elicit information different from this theme. Hence, the item was replaced by another item which has direct relation with the theme.

The pilot further revealed that requesting instructors to fill in and return the questionnaire on the day of its dispatch was not feasible and workable since they all demanded more time to complete

the questionnaire. Thus, it was realized that urging instructors to fill and return the questionnaire on spot was not good idea since this might result in superficial and missing responses, and the main study participants would be given few days to complete and return the questionnaire. Generally, the experience of piloting the questionnaires was useful in that it allowed the researcher to test for comprehensibility and internal consistency of the questionnaire items, and to make sure that responses were relevant in relation to the research questions.

Moreover, as a result of piloting the interview schedule, it could be realized that few of the questions were redundant, as the instructors' responses were observed to be repetitive on certain questions. It was also realized that the ordering of the questions had to be restructured in some more logical order. Hence, based on the interviewees' suggestions and insights gained, some amendments were made to the instructors' interview guide. Apart from this, the results show that the instructors' interview questions were clear and understandable, and that the technique (face-to-face interviewing) turned out to be viable and reliable.

In terms of students' interview, the pilot also revealed that some of the questions had to be reworded since the students were repeatedly requesting the interviewer to restate these questions, suggesting their confusion to the questions. It was also realized that some probing questions had to be added since the researcher observed responses which were very general and brief rather than direct and detailed from respondents.

Further, following translation of the interview schedule to Afaan Oromoo, the term "qormaata" became a challenge during interviews. Although the term was used to represent "assessment" (understood in this study as a general term that includes the full range of procedures used to gain information about student learning), students were observed while conceiving the term narrowly as something that represents formal exams. It was difficult to find Afaan Oromoo term that has exact meaning with the term. Hence, it was decided to use the original term "assessment" in the main study interview, realizing that it is a commonly used and understandable term among the students. Generally, piloting the interview guide helped the researcher fine tune the questions which were not clear to the interviewees, merge redundant questions, rearrange, and add and develop probes to enable him gain richer responses from the participants. Besides, more confidence and interviewing skills were gained from conducting the pilot interviews.

3.8. Procedures of Data Collection

The research data inquiry activities commenced as soon as the data gathering instruments were validated, piloted, and approval from the advisory committee was granted. Since the study utilized concurrent design, the qualitative and quantitative data were collected at approximately the same time. Hence, in order to conduct the interviews and administer the questionnaires, the researcher moved to the sampled universities and first contacted heads of the English language department or program leaders and made a request for cooperation with a letter received from Haramaya University. Meanwhile, lists of on duty EFL instructors were received from the three universities that helped easily contact the sample participants.

Next, participant instructors for the interview schedule were identified with the assistance of heads of English language department at each university. As to student interview participants, they were contacted through their instructors, who were assigned for the course communicative English skills at the sampled sections. Once the researcher got access to the instructors and students, all the required participants were consulted for the interview sessions. As a result, consent was granted and appropriate time and place for the interview schedules were arranged with the participants. This time, EFL instructors were also requested to submit copies of documents such as students' assessment papers and course outlines needed for the qualitative part of the study. Then, the individual interviews were conducted by informing the purpose of the study to each participant.

There were some useful strategies that the researcher, who was the interviewer, used to more benefit from the interview processes. For example, to show his dedication to the study and then get good initiation from the participants, the researcher had been at the interview place ahead of the interviewees. In addition, the researcher tried to create a positive atmosphere of trust, friendliness and openness starting from the moment of meeting the participants. This was to allow the participants reflect their opinions on the issue under investigation openly and without any refrain (Patton, 2002). Moreover, all the interview sessions, with the permissions granted from the interviewees, were audio recorded. The researcher also tried to take quick notes while conducting the interviews. The documents needed for this study were also collected during the interview sessions. Finally, as suggested by Patton (2002) and Creswell (2014), the interview data collected were immediately saved safely on computers.

With regard to administration of the survey questionnaires, since the number of EFL instructors teaching communicative English skills course was manageable, the researcher contacted every instructor via the technique described above. Then, he distributed the questionnaires to every instructor after consulting them and getting consent. The researcher convinced and requested heads of the department to take responsibility to collect and submit the questionnaires filled by the instructors. Accordingly, they collected and returned the questionnaires filled by the survey participants in a week period. On the other hand, the researcher administered the survey questionnaire to the student participants being at their classrooms by heading to each sample section. The administration was assisted by the communicative English skills course instructors of the sampled sections. Self-administration of the survey questionnaires was to minimize attrition, which Dornyei (2007) argued that it helps researchers to secure relatively high return rate compared to that administered through third party.

3.9. Data Analysis Procedures

The data collected through questionnaire, interview and documents were analyzed quantitatively and qualitatively. The analysis was made distinct for the quantitative data from questionnaires and the qualitative data from interviews and documents.

3.9.1. The Quantitative Data Analysis

The data which were collected through questionnaires were analyzed quantitatively using statistical software package. SPSS v.23 was used for this purpose since the researcher had access to and was familiar with. This was also for the reason that the package is sufficient for any of the statistical analysis required for this study. In preparing the questionnaire data for the analysis, the data for EFL instructors and students was separately coded. Both EFL instructors' and students' responses to Likert type items were categorized and coded in terms of the different themes pertinent to answering the research questions.

The data were analyzed using Likert-type item analysis, and the numerical values to responses and the numerical representation (the coding) of the items of the Likert Scale data were conducted as Strongly Agree (SA) = 5; Agree (A) = 4; Undecided (U) = 3; Disagree (D) = 2 and Strongly Disagree (SD) = 1 for the items of the assumptions about assessment questionnaire. Similarly, items of the assessment practice questionnaire were coded as Always (A) =5; Often

(O) = 4; Sometimes (ST) = 3; Rarely (R) = 2 and Never (N) = 1. Then, the questionnaire-based data on the participants' assessment assumptions and practices scored and coded this way were entered into a computer, screened and cleaned, and analyzed through both descriptive and inferential statistics.

The researcher conducted analysis of descriptive statistics including means and standard deviations to explain the different features of assessment practices in communicative English skills course. These descriptive statistics were also employed to determine participants' inclination either to traditionalist or alternativist assessment assumptions. Further, independent samples t-test and analysis of variance (ANOVA) were calculated in order to determine differences in assessment assumptions between and among the different groups of EFL instructors. To do so, however, the data were first checked for outliers and parametric assumptions prior to using the tests, and it was proofed that they were appropriate for the statistics.

3.9.2. The Qualitative Data Analysis

The qualitative information collected through the interviews and reviews of relevant documents were analyzed qualitatively. Even though there is no single agreed-upon method for analyzing qualitative data, researchers must decide what they want from the data based on their study purpose. Accordingly, this study adopted thematic and content analysis approaches.

Thematic analysis, which involved the generation and application of codes to data and the identification, analysis, and categorization of patterns into themes, allows researchers to analyze data flexibly, either simultaneously with data collection or after the entire data collection (Ezzy, 2002; Braun & Clarke, 2006). Thus, it permitted the researcher to begin the analysis once the data was collected, as its transcription was also by the researcher. That is, the data collection and transcription were made simultaneously, and this, because of time, allowed no room for its analysis during its collection. For this reason, the analysis took place after the entire data were collected, and thematic analysis is appropriate for such strategy unlike the analysis in grounded theory.

Thematic analysis is also appropriate where the general issue of interest is identified prior to data analysis, still allowing the emergence of themes which are not anticipated (Braun & Clarke,

2006; Saldana, 2009), the other rational to adopt. In this study, since themes which can describe assessment assumptions and practices were identified from the literature, they were not totally left to emerge from the data. Thus, the analysis was done by approaching the data mainly with predetermined categories and allowing incorporation of those which emerged from the data.

As to the analysis process, although there is no one clear-cut way in qualitative data analysis, some guidelines suggest how to run the analysis critically, reflectively and systematically. This study adopted Braun and Clarke six step thematic analysis framework, which the researcher thought a clear guide and that Maguire and Delahunt (2017) argued the most influential framework in the social science.

First, all the audio-recorded interviews were transcribed verbatim. This was done by the researcher himself than hiring transcriber for the sake of familiarizing himself with the data. This is what is encouraged than outsourcing the transcription to others, as it “provides an opportunity to get immersed in the data, an experience that usually generates emergent insights” (Patton, 2002: 441). Once the transcription was completed, the whole transcribed data were read and re-read several times to get sense of the overall data and to determine emerging themes as well, if any. Then, the coding process began by carefully reading the data word by word and assigning codes to a segment of text in order to generate the initial codes. This was followed by grouping and rearranging codes of similar topics into the existing and potential categories based on relationships among the codes. Next, the categories were grouped under major themes.

As the coding process in nature is cyclic and iterative, rather than a one-time approach, the researcher continuously reviewed, revised and refined codes, categories, and themes through constant back-and-forth movement between the entire data set. During analysis, instructors were assigned with names of Instructor 1 to Instructor 9 and students with Student 1 to Student 12 in light of their turn of interviews. Responses of these participants were organized based on major themes and sub-themes of the assessment practices and assumptions. The final analysis was done and reported incorporating the evidence extracts of the themes.

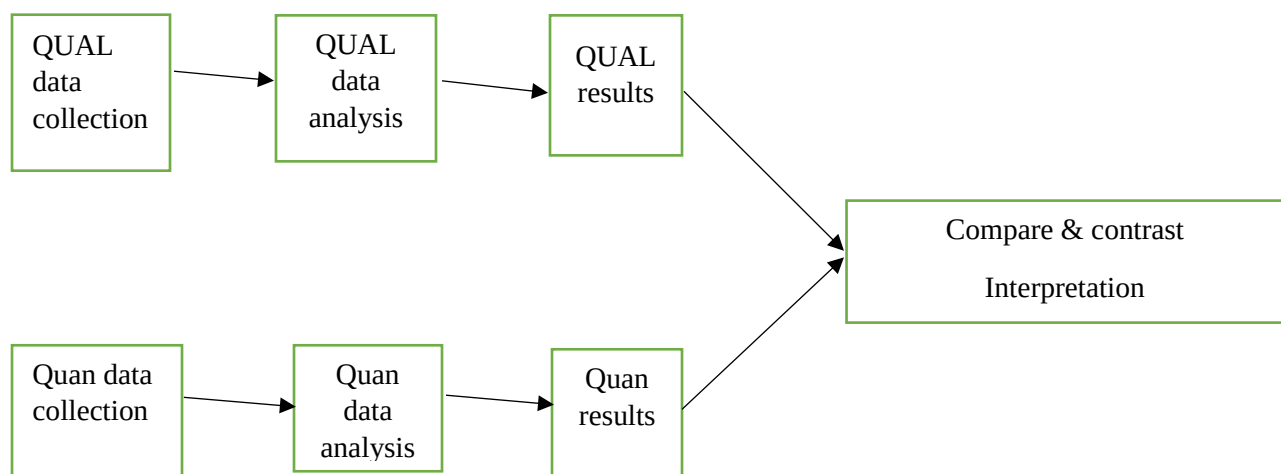
Content analysis was also used to determine the proportion of responses that fell within the themes and sub-themes. Regarding assessment assumptions, each participant’s reason was coded into traditionalist and alternativist categories based on previous framework. In doing this, the

researcher focused on responses that he could confidently code as either traditionalist or alternativist. The qualitative information on assessment practices collected through interviews and documents were also coded for themes and sub-themes based on categories derived from previous works. Then, frequency counts were determined for each theme. To obtain the percentage for each frequency, the total number of codes associated with each sub-theme under each theme was divided by the total number of codes included and multiplied by hundred. Particularly, more information is found in Appendixes K and L in the charts that displayed the combined results of the document analysis.

3.9.3. Integrating the Qualitative and Quantitative Results

The integration of the quantitative and qualitative findings of the study was made in line with the design of the study. In concurrent design the researcher collects and analyzes both qualitative and quantitative data independently and in parallel way. This means the integration of the two data sets does not take place at the stages of collection and analysis but at the stage of interpretation of the data. In the same vein, as already mentioned, data from questionnaires, interviews and documents were collected and analyzed simultaneously and separately, and results from the two data sets were integrated at the stage of data interpretation. Then, the data were really merged by integrating or comparing the results from both questionnaire and interview and documents side-by-side in a discussion. This design is useful to broaden the research perspective by examining how results from different findings complement or corroborate each other (Creswell, 2009; Dornyei, 2007).

Figure 4. A framework of data collection, analysis and integration for the study



3.10. Ethical Considerations

In the current study, though the nature of the study is not that pose any special ethical problems on the participants, some issues were considered to prevent any possible complications. As a researcher, the first thing to do was getting permission from the institutions (the target universities) in order to have access to them. Cohen *et al.* (2007) stress the importance of securing permissions to access the institution where the research is conducted before embarking on data collection. Accordingly, by informing them the purpose of the study, the researcher first secured permission from the universities' principals to conduct the study in their institutions. This was accompanied by submitting a request letter for cooperation received from hosting university (Haramaya). Once the permission was granted, English language department heads and freshman college representatives were approached using the same manner upon which EFL instructors and students participated in the study were accessed.

Studies are more principled if they are conducted through informed consent. Thus, prior to data collection, the sample EFL instructors and students were informed by the researcher that their participation in the study is voluntary-based and requested without any compulsion. The informed consent included details of purpose of the investigation, participants' role in the study and the significance of their contribution. Then, the issues of data confidentiality and privacy of participants were ensured after securing their willingness to participate in the study.

In the case of confidentiality, the participants of the study were assured that the information they provide will be kept secured to protect any threat on them because of the information. Accordingly, the researcher stored the data collected securely, did not allow anyone to access the data, and the information gathered were not utilized in any way that may harm them. Further, to protect the participants' privacy, the questionnaires were made to anonymously returned, and during data transcription and analysis, alphanumeric was applied for all interview participants in order to protect their identity. Finally, the participants were informed that they have the right to be unrecorded at all or on certain parts of the interview, to withdraw from the study at any time, and to choose time and place appropriate for them.

CHAPTER FOUR

DATA ANALYSIS, FINDINGS AND DISCUSSIONS

4.1. Introduction

This study aimed to explore assessment practices and the underlying assumptions in implementing a communicative English language skills course at universities in western Ethiopia. In this chapter, findings from the data analysis process and discussions of the findings will be presented in separate sections. Accordingly, the first section focuses on the analysis and the findings of the study, which begins with the assessment practice followed by the underlying assumptions. For both the assessment practice and assumption themes, first, the analysis and results of questionnaires are presented followed by the analysis and findings of the qualitative data. The second section of this chapter presents discussions of the findings by integrating findings of the qualitative data and the questionnaire.

4.2. Analysis and Results of the Quantitative Data

Prior to discussing the results of the survey questionnaire, it is essential to describe response rate of the questionnaires distributed to 84 EFL instructors and 389 first year students. Despite the diligent effort of the researcher to achieve the highest possible response rate of the questionnaires, there were some unreturned questionnaires from the respondents. Additionally, some were excluded particularly from the students' returned questionnaires since they were not filled properly. Yet, the response rates from both participants are at an acceptable level (Cresswell, 2014). Among the 389 questionnaires distributed to students, 48 of them were not returned at all, and among the 341 returned questionnaires, 32 of them were discarded because of their incomplete responses. So, only 309 (79.4% response rate) questionnaires were properly filled and returned. On the other hand, among 84 questionnaires distributed to EFL instructors, only 65 of them were returned, which counted for 77.4% return rate. Therefore, questionnaire-based data gathered from 65 EFL instructors and 309 students were analyzed in accordance with the research questions of the study as follows preceded by analysis of respondents' characteristics.

4.2.1. Characteristics of participants

It is essential to discuss background characteristics of the participant respondents in order to obtain gateway/general images about them. The data is presented and analyzed using frequencies and percentages as follows.

Table 5. Instructors' demographic characteristics

Characteristics		Frequency	Percentage
Sex	Male	62	95.4
	Female	3	4.6
Academic qualification level	MA/M.Ed.	49	75.4
	Ph.D.	16	24.6
Field of specialization	TEFL	48	73.8
	Non-TEFL	17	26.2
Years of teaching experience	≤5	11	16.9
	6-10	17	26.2
	11-20	27	41.5
	>20	10	15.4
Average number of students in a class	≤40	2	3.1
	41-60	44	67.6
	61-80	17	26.2
	>80	2	3.1

Out of the 65 instructors, while 62 (95.4%) of them are males, the rest 3 (4.6%) comprises female groups, which shows very low proportion of female EFL instructors in the selected universities. Regarding their academic qualification, most or 49 (75.4%) of them were M.A/M.Ed. holders, and 16 (24.6%) were Ph.D. holders. However, instructors with bachelor's degree were not observed in the sample participants in the selected universities. This indicates that a qualification mixes of EFL instructors who were offering communicative English language skills course in the universities under study seemed to harmony with the minimum qualification standard set by ministry of education. In relation with the specialization of the participants,

majority (73.8%) of them were TEFL graduates whereas non-TEFL such as linguistics and literature graduates comprises 26.2% of the respondents.

Unlike the demographic variables discussed above, there are variations in the teaching experience of the respondents. 16.9% of the respondents held 5 years of experience or below; 26.2% of the instructors had teaching experience between 6 and 10 years; instructors with experience between 11-20 years comprises 41.5% of the respondents; and those instructors who had greater than 20 years of teaching experience were 10 (15.4%). Most of these instructors (67.6%) held class size between 41-60 students, and about 26.2% of them reported there were averagely 61-80 students in a class. On the other hand, groups having less than or equal to 40 and greater than 80 average number of students in a class share the same figure 2 (3.1%) each.

Table 6. Students' characteristics

Characteristics		Frequency	Percentage
Sex	Male	181	58.6
	Female	128	41.4
Field of study	Natural science	197	63.8
	Social science	112	36.2
Age	17-19	243	78.6
	20-22	57	18.5
	≥ 23	9	2.9

As demonstrated in Table 6 regarding student participants, there is variation in the sex of the respondents. From 309 students who properly filled in and returned the questionnaires of the study, more than half of them are males (N=181), and the rest (N=128) are females. There also existed unproportional number of participants in terms of their fields of study. While 197 (63.8%) of students were from natural science stream, 112 (36.2%) comprised social science groups. Regarding age group of the participants, majority (N=243 or 78.6%) of them were found between 17 and 19 years, and those ranging between 20 to 22 and greater than or equal to 23 years were 57 (18.5%) and nine (2.9%), respectively.

4.2.2. Assessment Practices

This section deals with the first part of the research questions, which investigated the practices of assessment in communicative English language skills. Conducting assessment includes a number of tasks that practitioners implement to achieve their assessment goals. This includes identifying the purpose of assessment, choosing assessment methods and feedback provision, among the others. Emphasizing on these key components in the classroom assessment process routes, this section presents the analysis and results of data from the participants on instructors' practices of assessment.

4.2.2.1. Purposes of assessment

Respondents of this study were asked to rate for the frequency of use of various classroom assessment purposes (assessment for learning, of learning and as learning). Instructors and students indicated their responses to these classroom assessment purposes in the context of communicative English language skills course.

Table 7. Frequency of use of assessment for learning

Items	Respondents	N	Mean	Std. Dev.
to plan instruction for the course	Instructors	65	3.26	1.189
	Students	309	2.86	1.161
to obtain information on students' progress	Instructors	65	3.58	.998
	Students	309	2.97	.999
to group students for course instruction purposes	Instructors	65	3.25	.985
	Students	308	2.84	1.020
to diagnose strengths and weaknesses in instructor's teaching of the course	Instructors	65	3.80	1.034
	Students	308	2.87	1.017
to provide feedback to learners as they progress through the course	Instructors	65	3.66	1.050
	Students	308	2.77	.984

As shown in Table 7, the instructors' mean scores to three of the five items describing assessment for learning purpose (M= 3.58, SD= .998; M= 3.80, SD= 1.034 and M= 3.66, SD=

1.05) showed that they frequently practiced assessment to facilitate the teaching of communicative English skills course. For instance, they reported that they frequently used assessments to diagnose their own strengths and weaknesses in the course and to provide progressive feedback to the learners. However, their mean scores to two items ($M= 3.26$, $SD= 1.189$ and $M= 3.25$, $SD= .985$) showed that this genre of assessment purpose was occasionally practiced. That is, the instructors sometimes employed assessment for some instructional activities such as grouping students for instructional purposes and planning instruction for the course.

Results from students' responses to the instructors' use of assessment for learning purpose, however, showed that instructors did not frequently use assessments to enhance instruction. The students' mean scores to the items loaded under assessment for learning, which ranged between $M= 2.77$, $SD= .984$ and $M= 2.97$, $SD= .999$, indicated that instructors occasionally used assessments for the particular purpose. That is, they sometimes conducted assessments, for instance, to obtain information on students' communicative English language skills learning progress and to provide them feedback as they progress through the course.

Table 8. Frequency of use of assessment of learning

Items	Respondents	N	Mean	Std. Dev.
to measure extent of learning at the end	Instructors	65	4.59	1.226
	Students	309	3.64	.829
to rank students based on their class performance	Instructors	65	3.86	1.102
	Students	308	3.34	.810
to determine the final grades for students	Instructors	65	4.17	.911
	Students	308	3.89	.790
to provide information to the central administration (e.g., depart., college)	Instructors	65	3.66	1.278
	Students	309	3.41	.853

As demonstrated in Table 8, all the items are reported by instructors to have been frequently practiced. Their mean scores to the four items comprising assessment of learning purpose (4.59, 3.86, 4.17 and 3.66) depicted that they frequently used them. That is, the instructors commonly

practiced assessments in communicative English language skills course for purposes such as determining the course final grade of the students and ranking students based on their course performance.

Similar results were also reported by the student participants. The students' mean scores to all, except one, items (3.89, 3.64 and 3.41) showed frequent practice of assessment of learning purpose. In other words, the mean scores for three of the four items, which can be rounded to value assigned for "often", indicated that instructors usually use assessment of learning. However, students also responded that instructors occasionally used assessments to rank students based on their class performance, an item with mean score of 3.34. Generally, the respondents' mean responses to almost all the items indicated frequent practices of assessment of learning. That is, instructors regularly used assessments to confirm whether or not students have met curriculum outcomes in communicative English language skills course.

Table 9. Frequency of use of assessment as learning

Items	Respondents	N	Mean	Std. Dev.
to create an environment where it is helpful for students to complete a given task on their own	Instructors	65	3.95	.959
	Students	309	2.90	.886
to facilitate students to become independent learners	Instructors	65	3.91	.980
	Students	307	2.81	.919
to help students in recognizing what aspects of their own work in the course need to improve	Instructors	65	3.74	.940
	Students	307	2.70	.887
to help students to examine their own English language learning	Instructors	65	3.46	.920
	Students	308	2.77	.852
to assist students to assess their own performance in class	Instructors	65	3.42	.983
	Students	307	2.60	.920

The above table describes EFL instructors' assessment as learning practices, the third cluster in assessment purpose. The mean value calculated from the instructors' responses for all the items (3.95, 3.91, 3.74, 3.46, 3.42) indicated that this category of assessment is frequently practiced.

The instructors, for example, usually conducted the course assessments to support students' independence in learning and help them examine their communicative English language skills learning. Shortly, the means calculated from instructors' responses for all the items, which can be rounded to the value 'often', indicated that they usually conducted assessments to promote learners' metacognitions.

On the other hand, the students' responses to all the items in this category showed that their instructors did not emphasize on the practices of assessment as learning purposes in communicative English language skills classrooms. The students' mean scores to majority of the items (2.9, 2.81, 2.7 and 2.77) depicted occasional practices of this assessment purpose genre. Even the students showed that the instructors hardly used assessments to assist students to assess their own performance in class, which was represented by mean score of 2.60.

4.2.2.2. Methods of assessing

The second part of the assessment practice survey dealt with the assessment methods employed while assessing students' communicative English language skills. The respondents were asked to rate for the frequency of practices of various methods of assessing students, which can be clustered into alternative and traditional assessment methods. Accordingly, responses from instructors and students' questionnaire results are analyzed and presented.

Table 10. Frequency of use of assessment methods

Methods	Respondents	N	Mean	SD
Asking oral questions (informal)	Instructors	65	3.86	1.088
	Students	308	3.33	1.089
Student self-Assessment	Instructors	65	2.88	.820
	Students	309	2.81	.980
Peer Assessment	Instructors	65	2.78	.944
	Students	307	2.60	.915
Written Assignment	Instructors	65	3.14	.722
	Students	309	3.22	.896
Portfolio	Instructors	65	2.46	1.133

	Students	309	2.44	.971
Project work	Instructors	65	2.95	1.096
	Students	309	2.55	.965
Oral Presentation	Instructors	65	3.32	1.047
	Students	307	2.80	.974
Role Play	Instructors	65	2.88	1.038
	Students	308	2.57	.968
Interview	Instructors	65	2.68	1.105
	Students	308	2.52	.932
Debate	Instructors	65	2.92	.989
	Students	307	2.73	.856
Observation	Instructors	65	3.52	1.120
	Students	309	3.28	.937
Quizzes/tests	Instructors	65	4.17	.821
	Students	309	3.68	.784
Midterm exam	Instructors	65	4.34	.871
	Students	309	3.92	.786
Final exam	Instructors	65	4.54	.812
	Students	309	4.16	.760

As demonstrated in Table 10, the instructors' mean scores to majority of the items (e.g., 2.88, 2.78 and 2.68) showed that they occasionally employed these assessment forms in communicative English language skills classes. They also indicated that they usually used oral questioning and observation assessment methods (Mean = 3.86, SD= 1.088 and M= 3.52, SD= 1.12). However, their mean score computed for the fourth item (2.46) showed that the instructors rarely used portfolio as assessment method in the target course. On the other hand, instructors' mean scores to written tests/quizzes, mid-term and final exams were respectively 4.17, 4.34 and 4.54. This showed that these assessment methods were often used by the course instructors. Generally, responses of the instructors indicated that the conventional forms were the most

commonly used assessment methods in their classrooms while portfolios, interviews and peer-assessment were the methods least used.

Results from students' responses are not far from the instructors. The mean scores to the five of the eleven alternative assessment methods (e.g., 2.60, 2.44 and 2.55) indicated that they were hardly employed by the instructors though their mean scores to the rest methods (e.g., 3.33, 2.81 and 2.80) showed that they were occasionally practiced. They further reported that the customary paper and pencil methods were frequently practiced (Mean = 3.68, 3.92 and 4.16). Overall, data from students indicated that tests, mid-term and final exams were the three most common assessment methods used, whereas portfolio and interview were the two least used methods while assessing their communicative English skills.

4.2.2.3. Provisions of assessment feedback

The final section of the assessment practice survey was intended to gather data on how assessment feedback is practiced. It deals about modes, sources, and focus/purposes of feedback. Accordingly, analysis and results from the respondents' ratings are presented as follows.

Table 11. Sources and modes of assessment feedback

Items	Respondents	N	Mean	Std. Dev.
Instructor provides written comments	Instructors	65	3.54	.903
	Students	309	3.28	1.082
Instructor provides oral comments	Instructors	65	3.78	.673
	Students	308	3.39	.874
Peer-feedback: students commenting on each other's work	Instructors	65	3.55	1.000
	Students	309	2.65	.909
Self-feedback: own evaluation of his/her work	Instructors	65	3.20	1.162
	Students	308	2.55	.916

Regarding source of feedback (who should feedback be given by), the instructors, through their mean scores of 3.54 and 3.78, indicated that they frequently practiced the two instructor-led feedback types. That is, they usually employed both written and oral assessment feedbacks. Data

from instructors further showed that they used both written and oral modes of delivering assessment feedbacks almost equally in the target course. As to student-led feedback provisions, although instructors reported that peer-feedback was used frequently ($M= 3.55$), their mean score to self-feedback (3.20) indicated their occasional use. This means that although the instructors usually allowed students to comment on each other's work, they sometimes provided opportunity to student to evaluate his/her own work.

From students' perspectives, their mean values to the feedback provided by instructors (3.28, 3.39) showed that instructors sometimes provided both written and oral assessment feedbacks. That is, although both instructor-centered modes of feedback provisions are employed by the instructors, they were occasionally employed. When it comes to the feedback practice that involves students, the students' mean scores to peer-feedback (2.65) and self-feedback (2.55) revealed their instructors' infrequent practices. In other words, instructors occasionally and hardly allowed students for peer and self-feedbacks practices, respectively.

Focuses of instructors' assessment feedback

Table 12. Verification assessment feedback

Items	Respondents	N	Mean	Std. Dev.
instructor provides a score/grade on students' assessments	Instructors	65	3.86	.808
	Students	309	3.77	1.101
instructor indicates if a student's answers in the class are correct	Instructors	65	3.94	.788
	Students	308	3.68	1.062
instructor points out where a student is right and wrong in his/her work	Instructors	65	4.03	.749
	Students	309	3.27	.964
instructor points out the places that are not completely correct or incomplete	Instructors	65	3.91	.931
	Students	309	3.50	.914
instructor tells a student what knowledge points he/she has mastered	Instructors	65	3.75	.985
	Students	308	3.55	.979

As specified in Table 12 above concerning the use of assessment feedback for verification purpose, the instructors' mean scores to all the items loaded under this sub-theme (e.g., 4.03, 3.94 and 3.91) indicated their frequent practices of the feedback. The instructors reported to have used verification type of assessment feedback often times in communicative English language skills course. In particular, instructors frequently ensured that they fairly pointed out where a student is right and wrong in his/her work and whether student's answer is correct or not.

Results from students' responses are not quite far from the instructors' responses in terms of frequency of use of verification feedback. Except to the third, the students' mean scores to all the items under this sub-dimension (3.77, 3.68, 3.50 and 3.55) confirmed that instructors' feedback frequently focused on verifying or justifying students' answers. However, students' mean score computed for an item (3.27) showed that instructors only occasionally pointed out where a student is right and wrong in his/her work.

Table 13. Directive assessment feedback

Items	Respondents	N	Mean	Std. Dev.
instructor directly tells a student the correct answer when he/she gets an answer wrong in the class	Instructors	65	3.54	1.174
	Students	308	3.59	.973
when a student gives an incomplete answer, instructor directly supplements it for him/her	Instructors	65	3.54	1.047
	Students	308	3.38	.963
when a student makes mistakes, instructor directly corrects it for him/her	Instructors	65	3.25	1.000
	Students	307	3.44	.966
instructor directly tells a student how to solve difficult problems in his/her work	Instructors	65	3.48	.831
	Students	308	3.45	.979
instructor directly tells students the correct answers to the test questions	Instructors	65	3.29	1.057
	Students	308	3.51	1.045

Regarding the directive type of feedback, instructors' mean values for the three of the five items (3.54, 3.54 and 3.48) depicted that they frequently practiced this genre. For example, directly telling a student the correct answer when he/she gets an answer wrong in the classroom and

directly supplementing to students' incomplete answers were reported by instructors to be most frequently used practices. However, through their 3.25 and 3.29 mean scores, instructors indicated that they occasionally directly corrected when students make mistakes and told them the correct answers to the test questions.

Students' responses to instructors' use of directive feedback also confirmed that instructors frequently used the genre. Except for one, the students' mean responses to all the items loaded under directive feedback (3.59, 3.44, 3.45 and 3.51) showed their instructors' usual practices of the items. However, the students also reported instructors' occasional practice for the item which reads, when a student gives an incomplete answer, instructor directly supplement it for him/her, for which 3.38 mean was computed.

Table 14. Scaffolding assessment feedback

Items	Respondents	N	Mean	Std. Dev.
instructor helps a student solve problems by offering some hints or cues	Instructors	65	3.03	.829
	Students	307	3.01	.907
instructor help students to complete difficult problems by reducing the difficulty	Instructors	65	2.69	.910
	Students	309	2.70	.945
to help students do their homework, instructor reminds them of some learning methods	Instructors	65	3.39	.694
	Students	308	2.78	.886
to help do their homework, instructor reminds students the task requirements/evaluation criteria	Instructors	65	3.25	.717
	Students	308	2.77	1.009

As demonstrated in Table 14, instructor respondents mean scores to all the items loaded under the scaffolding feedback practices (3.03, 2.69, 3.39 and 3.25) depicted that they practiced the genre infrequently. In other words, instructors occasionally employed scaffolding assessment feedback in communicative English language skills course. For instance, in order to help learners do their homework, assignment or project, instructors only sometimes offered students hints/cues and reminded them the task requirements or evaluation criteria.

Interestingly, results from students' responses were compatible with the responses from their instructors regarding the scaffolding feedback sub-dimension. The students' mean scores for all the items loaded under this genre (3.01, 2.70, 2.78 and 2.77) indicated instructors' infrequent practices of scaffolding feedback. This is to means that the instructors moderately employed assessment feedback to scaffold students' learning.

Clustered classroom assessment practices

This section presents the analysis and results from the instructors and students summated mean responses to the practices of assessment purpose clustered into assessment for, of and as learning; methods into alternative and traditional; and the feedback practices clustered into instructor-centered, student-centered, verification, directive and scaffolding sub-dimensions.

Table 15. Descriptive statistics of clustered classroom assessment practices

Components		Respondents	M	SD
Purposes	for learning	Instructors	3.51	.807
		Students	2.86	.739
	of learning	Instructors	4.23	.899
		Students	3.48	.641
	as learning	Instructors	3.70	.802
		Students	2.77	.618
Method	Alternative	Instructors	3.03	.696
		Students	2.76	.573
	Traditional	Instructors	4.35	.746
		Students	3.92	.583
Feedback	Instructor-centered	Instructors	3.66	.638
		Students	3.18	.760
	Student-centered	Instructors	3.38	1.00
		Students	2.60	.711

verification feedback	Instructors	3.90	.503
	Students	3.37	.676
directive feedback	Instructors	3.42	.693
	Students	3.48	.696
scaffolding feedback	Instructors	2.91	.445
	Students	2.81	.687

To begin with the assessment purpose, as shown in Table 15, the instructors' reported summated mean scores were $M= 3.51$, $SD= .807$ for assessment for learning, $M= 4.23$, $SD= .899$ for assessment of learning, and $M= 3.70$, $SD= .802$ for assessment as learning. That is, the instructors reported that they frequently employed all the three assessment purposes in communicative English language skills course. Despite the fact, the instructors' summated mean score to the practice of assessment of learning is found to be higher than that of their summated mean scores to practices of the two others.

On the other hand, the students' summated mean scores to the instructors' practice of assessment for, of and as learning were ($M= 2.86$, $SD= .739$; $M= 3.48$, $SD= .641$; $M= 2.77$, $SD= .618$), respectively. Students reported that instructors infrequently practiced assessment for and as learning but frequently employed assessment of learning. In other words, consistent with the instructors' responses, students reported that instructors used assessment of learning more frequently than assessment for and as learning.

In terms of the assessment methods used, summated mean score of the instructors to the practice of traditional assessment method ($M= 4.35$, $SD= .746$) was higher than that of their summated mean score to the practice of alternative assessment methods ($M= 3.03$, $SD= .696$). That is, the instructors rated that they always employed paper and pencil-based assessment methods in contrast to the innovative methods, which is reported to be moderately practiced.

The instructors' responses of frequently practicing the traditional assessment methods did correspond to their students' responses. The students' summated mean for traditional assessment methods ($M= 3.92$, $SD= .583$) was higher than their summated mean score for the alternative

methods ($M = 2.76$, $SD = .573$). That is to say that the conventional methods were the methods that students reported their instructors used more frequently than the alternative assessment methods. These results from students indicated that although instructors regularly utilized the traditional assessment methods, they moderately used the innovative methods in the context of communicative English language skills course.

Regarding whom feedback is given by, the instructors' summated mean scores to the practice of instructor-led and student-led feedback were ($M = 3.66$, $SD = .638$; $M = 3.38$, $SD = 1.00$), respectively. That is, although both instructors and students took part in the feedback provisions, instructors were observed to be sources of assessment feedback more than students. In other words, instructors provided assessment feedback to students' more often than they allow students to comment on their own and/or their peers assessments.

Correspondingly, students summated mean score to the instructors' utilization of instructor-led feedback ($M = 3.18$, $SD = .760$) was observed to be higher than their summated mean score ($M = 2.60$, $SD = .711$) to the instructors' utilization of student-led feedback. Students reported that assessment feedbacks were occasionally provided by instructors, and their use of learners as a source of feedback was rare. That is, students were hardly involved in the self and peer-feedback practices.

Finally, with regard to the focus of instructors' assessment feedback, their summated mean scores to the practices of verification, directive, and scaffolding feedback types were $M = 3.90$, $SD = .503$; $M = 3.42$, $SD = .693$; $M = 2.91$, $SD = .445$, respectively. In other words, instructors provided verification and directive feedbacks with a high frequency, and provided scaffolding feedback with a moderate frequency. In this regard, while verification feedback was found to be the most frequently practiced, scaffolding feedback was the least frequently practiced type.

Coinciding with the instructors' responses, results from students summated mean scores indicated that verification and directive feedback ($M = 3.37$, $SD = .676$; $M = 3.48$, $SD = .696$) were more frequently practiced than scaffolding feedback ($M = 2.81$, $SD = .687$). Although instructors usually practiced feedback that focuses on directly pointing out the right and wrong answers, they occasionally employed feedback types that focus on verifying and scaffolding. That is,

directive feedback was the most frequently used while scaffolding feedback was found to be the least employed, similar to responses from the instructors.

4.2.3. Assessment assumptions

In order to understand assumptions underlying the participants' assessment practices and views in communicative English skills course, they rated their agreement to items comprised of some factors. Items in each factor were further categorized into two broad dimensions (those reflecting assumptions of traditional assessment and that of alternative assessment). Accordingly, first, results obtained from instructors' questionnaire responses are summarized and analyzed, and this is followed by students' questionnaire responses.

4.2.3.1. Instructors' assessment assumptions

Table 16. Descriptive statistics of instructors' assessment assumptions

Items		N	M	SD	Total	
Content of assessment/ Knowledge factor					M	SD
Traditional	In order to effectively assess, language needs to be broken into its components/skills.	65	3.20	.795	3.29	.559
	The focus of assessment should be on learner's ability to correctly reproduce what he/she has been taught.	65	3.37	.840		
	There are acknowledged correct responses to the course questions to be asked.	65	3.34	.906		
	Assessment needs to focus on certain specified contents similar across contexts.	65	3.25	1.016		
Alternative/modern	Assessment should be designed by integrating EL skills than focusing on isolated skills.	65	3.95	1.082	3.85	.764
	Assessment should require students to apply their own understanding of concepts, theories, etc. to real world problems.	65	4.05	1.052		
	Assessment should engage students in the knowledge creation process than in repeating the established one.	65	3.74	1.149		
	A question can have multiple justifiable correct answers.	65	3.66	1.163		
Values and judgments factor						
Traditional	Students' English language ability should be assessed objectively.	65	3.03	1.045	3.25	.771
	Consistency in assessment measures is the most important element of assessment practices in the course.	65	3.63	1.054		
	The assessments are accurate indicators of students' English language learning.	65	3.08	1.122		

Alternative/ modern	Assessments should be differentiated across student populations and contexts.	65	4.17	.977	3.83	.734
	Assessments only give a rough indication of students' English language learning.	65	3.77	1.115		
	Assessment which results in verbal data is illuminating than which result in numeric data.	65	3.54	.849		
Power and control factor						
Traditional	Much of the power in assessment needs to remain with a course instructor.	65	3.72	.696	3.20	.694
	The course instructor is sole authority to decide on assessment.	65	2.83	1.084		
	Instructors should control assessment since involving students makes the results inherently invalid.	65	3.05	1.007		
Alternative/ modern	Assessment is a joint process between instructor and students.	65	3.35	.891	3.02	.801
	Students should have prior knowledge of what is going to be in the assessment.	65	2.66	.940		
	Students should be powered to decide on their assessments of the course.	65	3.05	1.037		

As can be seen from Table 16, EFL instructors were not consistent in reporting their assumptions about assessment in communicative English skills course. Their summative mean score obtained from the alternative assumptions dimension ($M= 3.85$, $SD= .764$) was higher than the summative mean score obtained from the traditional assumptions dimension ($M= 3.29$, $SD= .559$) regarding content assessed. Similarly, their summative mean score to the alternative assumptions ($M= 3.83$, $SD= .734$) was higher than the summative mean score obtained from the traditional assumptions dimension ($M= 3.25$, $SD= .771$) regarding value and judgment factor. However, it was inversely scored for the power and control factor. Their summative mean score to traditionalist assumptions ($M= 3.20$, $SD= .694$) was observed to be slightly higher than the summative mean score to alternativist assumptions ($M= 3.02$, $SD= .801$).

Upon closer examination of the above table, it appeared that the instructors' mean scores to eight of the ten items (e.g., 3.20, 3.03 and 2.66) that loaded under traditionalist assumptions revealed indecision. That is, they seemed to be skeptical to majority of the statements that reflected assumptions of traditional assessment. For instance, they did not express their agreement or disagreement to the statements that emphasized instructors' sole authority in assessment and assessments are accurate indicators of students' English language learning. On the other hand, the surveyed instructors showed their positive opinion to statements that reflected the importance

of consistency in assessment measures (3.63) and the need for much power to remain with instructors in assessments (3.72). These assumption statements, as compared to their mean scores to other statements, could not reflect the instructors' indecision of the traditional assessment.

In contrast, the instructors' mean scores to all the items, except those under power and control scale, loaded under alternative assumptions dimension (e.g., 3.95, 4.05, 4.17) revealed their agreement ranging from slight to strong. That is, they expressed their positive opinion to all the assumptions under knowledge/content assessed and value and judgment underlying current assessment approaches. For instance, they were positive for integrated skills assessments and if assessments engage students in the knowledge construction process rather than in repeating the established one. Particularly, they clearly indicated their agreement with highest rate to the assumption that assessments should be differentiated across student populations and contexts. That being the case, there were three items to which the instructors observed indifferent. Their mean scores to all the power and control factor items loaded under alternative assumption dimension (3.35, 3.05 and 2.66) indicated indecision. They did not express their agreement or disagreement to all the assumptions that emphasized students' power in assessment.

Instructors' assessment assumptions inclination (traditionalist Vs alternativist)

In order to identify the instructors' general assessment assumption inclination, the items of the questionnaire were categorized into assumptions of traditional assessment and assumptions of alternative assessment broad dimensions. The comparison made between the two clustered assumptions using the instructors' summated mean scores, standard deviation and one sample t-test are presented as follows.

Table 17. Mean, standard deviation, and one sample t-test results of instructors' clustered assumptions

	N	M	SD	Df	T-test	P-value
Traditionalist assumptions	65	3.25	.523	64	3.853	.000
Alternativist assumptions	65	3.59	.547	64	8.755	.000

While comparing the two summated scores, as one can see from the table, the instructors' summated mean score to the traditional assessment focused assumptions ($M = 3.25$, $SD = .523$) was found to be lower than that of their summated mean score to the alternative assessment focused assumptions ($M = 3.59$, $SD = .547$). The instructors indicated higher level of agreement to assumptions underlying alternative assessment approach than they did for assumptions underlying traditional assessment approach. In other words, the results revealed communicative English language skills course instructors' inclination towards alternativist assessment assumptions than traditionalist-oriented assessment assumptions.

In addition, the instructors' mean score for both traditionalist assumptions and alternativist assumptions were above the hypothesized mean of 3 and statistically significant. One sample t-test ($t = 3.853$, $df = 64$, $p = .000$) revealed that instructors reported higher levels of traditionalist assumptions than the hypothesized mean of 3. Similarly, it ($t = 8.755$, $df = 64$, $p = .000$) indicated that instructors' score on alternativist dimension was significantly higher than the average point.

4.2.3.2. Students' assessment assumptions

Table 18. Descriptive statistics of students' assessment assumptions

Dimension	Items	N	M	SD	Total	
	Content assessed/ Knowledge factor				M	SD
Traditional	In order to effectively assess, language needs to be broken into its components/skills.	308	3.65	.861	3.49	.578
	The focus of assessment should be on learner's ability to correctly reproduce what he/she has been taught.	308	3.48	.956		
	There are acknowledged correct responses to the course questions to be asked.	308	3.37	.959		
	Assessment needs to focus on certain specified contents which are similar across contexts.	307	3.46	.954		
Alternative/modern	Assessment needs to be designed by integrating EL skills than focusing on isolated skills.	307	3.51	.975	3.02	.747
	Assessment should require students to apply their own understanding of concepts, theories, etc. to real world problems.	309	2.88	1.120		
	Assessment should engage students in the knowledge creation process than in repeating the established one.	307	2.83	1.161		
	A question can have multiple justifiable correct answers	309	2.86	1.148		
	Values and judgements factor				3.50	.559
Traditional	Your English language ability should be assessed	308	3.53	.848		

	objectively.					
	Consistency in assessment measures is the most important element of assessment practice.	309	3.40	.830		
	Assessments are accurate indicators of your English language ability.	309	3.55	.827		
Alternative/modern	Assessment should be differentiated across student populations and contexts.	309	2.54	.950	2.75	.774
	Assessments only give a rough indication of students' English language ability.	309	2.90	1.114		
	Assessment which results in verbal data is illuminating than which result in numeric data.	307	2.79	1.187		
	Power and control factor					
Traditional	Much of the power in assessment needs to remain with a course instructor.	308	3.55	1.005	3.60	.790
	The course instructor is sole authority to decide on assessment.	308	3.42	1.016		
	Instructors should control assessment since involving students makes the results invalid.	307	3.81	.978		
Alternative/modern	Assessment is a joint process between instructor and students.	307	3.08	1.183	3.18	.840
	Students should have prior knowledge of what is going to be in the assessment.	308	3.03	1.122		
	Students should be powered to decide on their assessments of the course.	309	3.42	1.098		

To begin with the cumulative results of each factor, as shown in Table 18, students were consistent in reporting their assumptions about assessment across the factors. Their summative mean scores obtained from traditional assumptions dimension (M= 3.49, SD= .578; M= 3.50, SD= .559; M= 3.60, SD= .790) were higher than their summative mean scores obtained from alternative assumptions dimension (M= 3.02, SD= .747; M= 2.75, SD= .774; M= 3.18, SD= .840) on content assessed, value and judgment, and power and control factors, respectively. The results indicated that students had more positive opinion for the assumptions that underlie traditional assessment approach than for those that underlie alternative assessment approach.

Closely examining the table, all the items loaded under traditional assessment assumption dimension received agreement scale from students (e.g., 3.81, 3.65, 3.48 and 3.55). In other words, students expressed their positive opinion to all statements constituting the three factors that describe assumptions underlying traditional assessment approach. Particularly, students

indicated their agreement with highest rate (3.81) to the assumption that instructors should control assessment since students' involvement incurs invalid results.

However, students were indecisive to the majority of assumptions that underlie alternative assessment approach. Their mean scores to seven of the ten items loaded under alternative assessment assumptions dimension (e.g., 2.88, 2.90, 3.08) revealed that they neither agree nor disagree to the items. The surveyed students were skeptical, for instance, to the statements that acknowledge variations of assessment measures and joint assessment processing. Despite the fact, the mean scores computed for two statements that emphasized integrated skills assessment (3.51) and the need for students to share power (3.42) indicated students' agreement with these particular items. Yet, with lowest rate of 2.54 mean score, students showed their disagreement with an alternative assumption that assessments should be differentiated across student populations and contexts.

Students' assessment assumptions inclination (traditionalist Vs alternativist)

Students' general inclination towards either traditional or alternative assessment assumptions was determined by comparing their mean scores from the two broadly clustered dimensions. The analysis conducted using the students' summated mean scores, standard deviations, and one sample t-test are presented as follows.

Table 19. Mean, STD, and one sample t-test results of students' clustered assumptions

	N	M	SD	Df	T-test	P-value
Traditional assumptions	309	3.46	.453	308	17.744	.000
Alternative assumptions	309	2.98	.584	308	-.509	.611

The surveyed students summated mean score obtained from traditional assumptions dimension (M= 3.46, SD= 0.453) was observed to be greater than that of their summated mean score to the alternative assessment assumptions dimension (M= 2.98, SD= .584). The students indicated higher level of agreement to assumptions underlying traditional assessment approach than they

did for assumptions underlying alternative assessment approach. Put it differently, communicative English language skills learners held more traditionalist assumptions than alternativist assumptions about assessment.

The students' mean score to the traditionalist assumptions was observed to be above the hypothesized mean of 3 and statistically significant. One sample t-test ($t = 17.744$, $df = 308$, $p = .000$) revealed that students reported significantly higher levels of traditionalist assumptions than the hypothesized mean of 3. On the other hand, their mean score to the alternative assumptions was observed to be lower than the hypothesized mean of 3 although one sample t-test did not prove that the difference is statistically significant. It ($t = -.509$, $df = 308$, $p = .611$) showed that students mean score to alternativist dimension was not significantly lower than the average point.

4.2.4. Differences in practices and assumptions according to instructors' demographic characteristics

This section deals about differences in instructors' assessment practices and assumptions in terms of some selected background characteristics. Independent samples t-test and analysis of variance (ANOVA) were used to examine the statistical differences in practices and assumptions across the different groups of instructors. However, since the number of female respondents to the survey questionnaire was only three, the variable 'sex' was not examined.

Differences in terms of instructors' levels of education

This demographic group refers to academic qualification levels of the EFL instructors in the three universities. While the group with master's degree (M.A./M.Ed.) comprises 49 instructors, the rest 16 constituted the group with a doctoral degree. Independent samples t-test was conducted to determine if a difference in assessment practices and assumption was present between the two sub-groups.

Table 20. T-test results of instructors' assessment practice comparison upon academic qualification

Practices	Dimension	Groups	N	M	SD	Mean difference	t-test	df	P
Purpose	of learning	M.A/M.Ed.	49	3.80	.879	-.00096	-.004	63	.997
		Ph.D.	16	3.80	.988				
	for learning	M.A/M.Ed.	49	3.52	.791	.04745	.203	63	.840
		Ph.D.	16	3.47	.879				
	as learning	M.A/M.Ed.	49	3.62	.831	-.30459	-1.327	63	.189
		Ph.D.	16	3.92	.677				
Methods	Traditional	M.A/M.Ed.	49	4.44	.618	.37968	1.407	18.72 3	.176
		Ph.D.	16	4.06	1.01				
	Alternative	M.A/M.Ed.	49	3.06	.676	.15077	.749	63	.457
		Ph.D.	16	2.91	.767				
Feedback	verification	M.A/M.Ed.	49	3.88	.537	-.06837	-.469	63	.640
		Ph.D.	16	3.95	.390				
	directive	M.A/M.Ed.	49	3.39	.670	-.12474	-.623	63	.536
		Ph.D.	16	3.51	.773				
	scaffolding	M.A/M.Ed.	49	2.89	.465	-.06027	-.467	63	.642
		Ph.D.	16	2.95	.390				

The first analysis was conducted to identify if differences exist in assessment purpose according to instructors' academic qualification levels. Regarding the assessment of learning dimension, as shown in Table 20, the mean score of instructors with master's degree ($M= 3.80$, $SD= .879$) and the mean score of instructors with doctoral degree ($M= 3.80$, $SD= .988$) were observed to be similar. On the other hand, the observed mean score of master's degree instructors ($M= 3.52$, $SD= .791$) was slightly higher than the observed mean score of instructors with doctoral degree ($M= 3.47$, $SD= .879$) to the assessment for learning dimension. However, this difference was not statistically significant ($t = -.722$, $df = 63$, $p= .473$). Regarding assessment as learning, the mean

score of MA participants ($M = 3.62$, $SD = .831$) was lower than the mean score of groups with PhD ($M = 3.92$, $SD = .677$) although independent sample t-test ($t = -1.327$, $df = 63$, $p = .189$) again showed that the difference was not statistically significant. Shortly, there is no difference between EFL instructors with MA and PhD degree regarding the practices of assessment of, for, and as learning purposes.

With regard to assessment methods, the mean score of instructors with master's degree ($M = 3.44$, $SD = .618$) was observed to be higher than the mean score of instructors with doctoral degree ($M = 4.06$, $SD = 1.01$) to the use of traditional assessment methods. Similarly, instructors with doctoral degree rated lower than their counter to the use of alternative assessment methods. The observed mean score of master's degree instructors ($M = 3.06$, $SD = .676$) was higher than the observed mean score of instructors with doctoral degree ($M = 2.91$, $SD = .767$). However, independent sample t-tests ($t = 1.407$, $df = 18.723$, $p = .176$) and ($t = .749$, $df = 63$, $p = .457$) indicated that the difference between the two groups is not statistically significant concerning the use of traditional assessment and alternative assessment methods, respectively.

Finally, regarding differences in the practices of assessment feedback, the mean score obtained from PhD instructors ($M = 3.95$, $SD = .390$) was higher than the mean score from MA degree instructors ($M = 3.88$, $SD = .537$) to the verification dimension. Similarly, PhD holders mean scores ($M = 3.51$, $SD = .773$) and ($M = 2.95$, $SD = .390$) were found to be higher than the mean scores of groups with MA degree ($M = 3.39$, $SD = .670$) and ($M = 2.89$, $SD = .465$) to the directive and scaffolding dimensions, respectively. So, the doctoral instructors showed higher rate than instructors with master's degree to all the verification, directive, and scaffolding assessment feedback dimensions. However, independent sample t-test ($t = -.469$, $df = 63$, $p = .640$), ($t = -.623$, $df = 63$, $p = .536$) and ($t = -.467$, $df = 63$, $p = .642$) respectively proved that the differences between the two groups were not statistically significant for all the three feedback dimensions.

Table 21. T-test results of instructors' assumptions comparison upon academic qualification

	Groups	N	M	SD	Mean difference	t-test	Df	P-value
Traditional	M.A/M.Ed.	49	3.22	.508	-.10880	-.722	63	.473
	Ph.D.	16	3.33	.570				
Alternative	M.A/M.Ed.	49	3.50	.497	-.39783	-2.642	63	.010
	Ph.D.	16	3.89	.598				

As to the traditional assessment assumptions dimension, as shown in Table 21, the mean score of master's degree instructors (M= 3.22, SD= .508) was lower than the mean score of instructors with doctoral degree (M= 3.33, SD= .570). That is, the doctoral instructors' expressed high agreement than master's degree instructors to the traditional assumptions. Despite the fact, independent sample t-test ($t = -.722$, $df = 63$, $p = .473$) revealed that the difference between the two groups is not statistically significant. The results indicated that both master's degree and doctoral degree instructors had similar level of assumptions underlying traditional assessment.

In terms of the groups' alternative assessment assumptions, the observed mean score of master's degree instructors (M= 3.50, SD= .497) was lower than the mean score of instructors with PhD degree (M= 3.89, SD= .598). Consistently, independent samples T-test ($t = -2.642$, $df = 63$, $p = .010$) proofed that the difference between the two demographic groups is statistically significant. The results indicated that instructors with doctoral degree scored significantly higher than those with master's degree on the alternative assessment assumption dimension. In other words, PhD degree holders held much more assumptions underlying alternative assessment approach than master's degree instructors.

Differences in terms of instructors' years of teaching experiences

Years of teaching experience demographic groups include ≤ 5 years group, 6-10 years group, 11-20 years group, and >20 years group. One way ANOVA was conducted to examine differences in practices and assumptions between the groups. As shown in the table below, first, the analysis for the assessment practice was presented followed by the assumptions.

Table 22. ANOVA results of instructors' assessment practice comparison based on experience

Practice	dimension	Group	M	SD	F	Df	P
Purpose	Of learning	≤ 5	3.75	.942	.127	3	.944
		6-10	3.81	.850			
		11-20	3.75	1.01			
		>20	3.95	.715			
	For learning	≤ 5	3.56	.768	.521	3	.670
		6-10	3.31	.649			
		11-20	3.56	.909			
		>20	3.66	.849			
	As learning	≤ 5	3.93	.546	.833	3	.481
		6-10	3.77	.818			
		11-20	3.52	.917			
		>20	3.78	.663			
Method	Traditional	≤ 5	4.58	.368	4.777	3	.005
		6-10	4.06	.959			
		11-20	4.63	.474			
		>20	3.83	.878			
	Alternative	≤ 5	2.94	.763	.518	3	.671
		6-10	2.88	.659			
		11-20	3.11	.620			
		>20	3.15	.614			
Feedback	Verification	≤ 5	3.98	.374	2.147	3	.104
		6-10	4.12	.347			
		11-20	3.73	.940			
		>20	3.90	.271			
	Directive	≤ 5	3.44	.528	1.951	3	.131
		6-10	3.69	.609			
		11-20	3.20	.813			
		>20	3.52	.483			
	Scaffolding	≤ 5	2.96	.430	1.132	3	.343
		6-10	2.75	.424			
		11-20	2.94	.459			
		>20	3.05	.453			

The mean scores to compare the responses of the four teaching experience groups to the purpose of assessment of learning dimension were (M= 3.75, SD= .942), (M= 3.81, SD= .850), (M= 3.75, SD= 1.01), and (M= 3.95, SD= .715) for instructors with ≤5, 6-10, 11-20, and >20 years,

respectively. When the results were compared within the four groups, it was observed that the groups were not significantly different from each other. F-test ($F = .127$, $df = 3$, $p = .944$) indicated that there was no statistically significant difference within the four demographic groups.

As to the assessment for learning dimension, the mean scores of the four groups were ($M = 3.56$, $SD = .768$), ($M = 3.31$, $SD = .649$), ($M = 3.56$, $SD = .909$), and ($M = 3.66$, $SD = .849$) for groups with ≤ 5 , 6-10, 11-20, and > 20 years, respectively. Comparing the results within the four groups, F-test ($F = .521$, $df = 3$, $p = .670$) indicated that the groups were not significantly different. Furthermore, the mean scores of the four teaching experience groups to assessment as learning dimension were ($M = 3.93$, $SD = .546$), ($M = 3.77$, $SD = .818$), ($M = 3.52$, $SD = .9017$), and ($M = 3.78$, $SD = .663$) for groups with ≤ 5 , 6-10, 11-20, and > 20 years, respectively. Similar to the other two dimensions, F-test ($F = .833$, $df = 3$, $p = .481$) showed that the four teaching experience groups were not significantly different in using assessment as learning purpose.

Regarding assessment method, the mean scores to the alternative assessment method dimension were ($M = 2.94$, $SD = .763$) for instructors with ≤ 5 years, ($M = 2.88$, $SD = .659$) for instructors with 6-10 years, ($M = 3.11$, $SD = .620$) for instructors with 11-20 years, and ($M = 3.15$, $SD = .614$) for instructors with > 20 years. When the results of the analysis of the responses were compared, it was observed that the groups were not significantly different from each other. As shown in Table 22, F-test ($F = .518$, $df = 3$, $p = .671$) showed that there was no statistically significant difference within the four demographic groups.

In terms of traditional assessment method, the mean scores of the four teaching experience groups were ($M = 4.58$, $SD = .368$), ($M = 4.06$, $SD = .959$), ($M = 4.63$, $SD = .474$) and ($M = 3.83$, $SD = .878$) for instructors with ≤ 5 , 6-10, 11-20, and > 20 years of teaching, respectively. Comparing the results within the four groups, it was observed that the groups had different levels of practices to traditional assessment methods, proved to be statistically significant. F-test ($F = 4.777$, $df = 3$, $p = .005$) showed that there was statistically significant difference within the groups in using traditional assessment methods. Hence, post-hoc was conducted to locate where exactly the difference/s lie/s as shown in table below.

Table 23. Tukey HSD Post-hoc Multiple Comparisons of Group Means

Groups	Mean difference	Standard error	p-value
Between ≤ 5 and 6-10 years	.51693	.26620	.222
Between ≤ 5 and 11-20 years	-.05387	.24607	.996
Between ≤ 5 and >20 years	.74242	.30058	.075
Between 6-10 & 11-20 years	-.57081*	.21299	.045
Between 6-10 and >20 years	.22549	.27416	.844
Between 11-20 & >20 years	.79630*	.25466	.014

*The mean difference is significant at the 0.05 level.

As a post-hoc analysis of the F-test result conducted using Tukey HSD showed, the mean scores of instructors in 11-20 years of teaching experience groups were significantly higher than the mean scores of instructors in both 6-10 and >20 years of teaching experience groups. The results indicated that instructors in 11-20 years of teaching experience significantly used traditional assessment methods than both instructor groups having 6-10 and >20 years of teaching experience. However, significant differences were not observed between the ≤ 5 & 6-10, ≤ 5 & 11-20, ≤ 5 & >20 , and 6-10 & >20 teaching experience groups.

Table 22 further illustrated differences in assessment feedback according to the four teaching experience groups. As to the verification feedback, the mean scores to the groups were ($M = 3.98$, $SD = .374$), ($M = 4.12$, $SD = .347$), ($M = 3.73$, $SD = .940$), and ($M = 3.90$, $SD = .271$) for instructors with ≤ 5 , 6-10, 11-20, and >20 years, respectively. In comparing the results within the four groups, it was observed that the groups were not significantly different from each other. F-test ($F = 2.147$, $df = 3$, $p = .104$) indicated that there was no statistically significant difference within the four demographic groups in using verification feedback type.

As to the directive feedback, the mean scores of instructors with ≤ 5 , 6-10, 11-20, and >20 years of teaching experience groups were ($M = 3.44$, $SD = .528$), ($M = 3.69$, $SD = .609$), ($M = 3.20$, $SD = .813$), and ($M = 3.52$, $SD = .483$), respectively. Comparing the results within the four groups, F-test ($F = 1.951$, $df = 3$, $p = .131$) indicated that the groups were not significantly different in using directive feedback. For the final scaffolding feedback dimension, the mean scores were ($M = 2.96$, $SD = .430$), ($M = 2.75$, $SD = .424$), ($M = 2.94$, $SD = .459$), and ($M = 3.05$, $SD = .453$) for groups with ≤ 5 , 6-10, 11-20, and >20 years of teaching, respectively. Similar to the other two

dimensions in this category, the groups had no significant difference to scaffolding feedback practices. F-test ($F = 1.132$, $df = 3$, $p = .343$) showed that the four teaching experience groups were not significantly different in using scaffolding feedback.

Table 24. ANOVA results of instructors' assumptions comparison based on teaching experience

	Dimensions									
	Traditional assumptions					Alternative assumptions				
	M	SD	F	Df	P	M	SD	F	Df	P
≤ 5	3.23	.576	1.583	3	.203	3.99	.359	15.835	3	.000
6-10	3.47	.544				3.97	.382			
11-20	3.12	.470				3.41	.495			
>20	3.26	.510				3.12	.394			

As shown in Table 24, the mean scores to compare the responses of the four teaching experience groups to the traditional assumptions dimension were ($M = 3.23$, $SD = .576$) for instructors with ≤ 5 years, ($M = 3.47$, $SD = .544$) for instructors with 6-10 years, ($M = 3.12$, $SD = .470$) for instructors with 11-20 years, and ($M = 3.26$, $SD = .510$) for instructors with >20 years. When the results were compared within the four groups, it was observed that the groups were not significantly different from each other. As shown in the table, F-test ($F = 1.583$, $df = 3$, $p = .203$) revealed that there was no statistically significant difference within the four demographic groups. This indicated that the four teaching experience groups had similar levels of assumptions underlying traditional assessment approach.

When it comes to the second dimension, the mean scores of the four groups to the alternative dimension were ($M = 3.99$, $SD = .359$), ($M = 3.97$, $SD = .382$), ($M = 3.41$, $SD = .495$) and ($M = 3.12$, $SD = .394$) for instructors with ≤ 5 , 6-10, 11-20, and >20 years, respectively. When the results of the analysis of the responses were compared within the four groups, it was found that the groups had different levels of assumptions, proved to be statistically significant through a one-way ANOVA. As can be seen in Table 24 above, F-test ($F = 15.835$, $df = 3$, $p = .000$) proved that there was statistically significant difference between the groups mean scores.

However, for the test does not show where exactly the difference/s lie/s, it was necessary to carry out a post-hoc analysis of the F-test result. Accordingly, the Tukey HSD post-hoc was used to perform multiple comparisons to locate the differences as seen in the following table.

Table 25. Tukey HSD Post-hoc Multiple Comparisons of Group Means

Groups	Mean difference	Standard error	p-value
Between ≤ 5 and 6-10 years	.02032	.16251	.999
Between ≤ 5 and 11-20 years	.58350*	.15022	.001
Between ≤ 5 and >20 years	.97091*	.18350	.000
Between 6-10 & 11-20 years	.56318*	.13003	.000
Between 6-10 and >20 years	.95059*	.16737	.000
Between 11-20 & >20 years	.38741	.15547	.071

The Tukey post-hoc test revealed that the group mean score of instructors in ≤ 5 years of teaching experience was significantly higher than the group mean scores of instructors in 11-20 and instructors with >20 years of teaching experiences. In addition, the mean score of instructors in 6-10 years of teaching experience group was significantly higher than the mean scores of instructors in 11-20 and with >20 years of teaching experience groups. Instructors in both 11-20 and >20 years of teaching experience groups had lower levels of agreement than the instructors in ≤ 5 and 6-10 years of teaching experience groups to the dimension.

The results indicated that less experienced instructors scored significantly higher than did the more experienced instructors to statements reflecting assumptions underlying alternative assessment approach. Further analysis by merging the first two (≤ 5 & 6-10; N= 28; M= 3.98) and the last two (11-20 & >20; N= 37; M= 3.30) experienced groups also proved statistically significant difference between the less experienced and more experienced instructor groups ($t= 6.22$, $df= 63$, $p= .000$). However, statistically significant differences were not observed between the ≤ 5 and 6-10 groups and 11-20 and >20 years of teaching experience groups.

4.3. Analysis and Findings of the Qualitative Data

This part presents findings from the qualitative aspect of the study which was categorized into two basic sections. First, the analysis and findings of assessment practice with its various themes

and sub-themes is presented, and this is followed by the analysis and findings of the participants' assessment assumptions.

4.3.1. Assessment practices

4.3.1.1. Purposes of assessment

Under this major theme, the study generally reflects on the first research question, the purposes for which EFL instructors conducted assessments in communicative English language skills course. The analysis of the instructors' interview responses were presented under three sub-themes: assessment for learning, assessment of learning and assessment as learning purposes.

Assessment for learning

Assessment for learning is understood as assessment that is used for the purpose of enhancing instruction. Such assessment is mainly used to inform instruction by identifying instructors' areas of strengths and weaknesses upon which they take the necessary measures. One of the recurring types used under this genre that the study instructors commonly reported was conducting assessment for understanding purpose. In all nine instructors' responses, there were some kinds of mentions that reflected that assessment is conducted primarily to know how much students understood what they are taught. For example, two instructors maintained that they usually conducted assessments "to get students' understanding" (Instructor 3) and "to check up my students' understanding" (Instructor 7).

However, in such brief answers, the practice mentioned can be identified as assessment of or for learning purposes. Thus, further question, that sought the instructors' subsequent actions once they knew their students' understanding level through assessment, was forwarded to help categorize the codes under either of the sub-themes. Accordingly, it was only in some of the instructors' discussions that understanding purpose of assessment reflected the assessment for learning idea. In this regard, the instructors used assessments mainly to assess and monitor teaching effectiveness and to provide feedback to students.

As to assessing and monitoring teaching effectiveness, instructors described that they carried out assessments to understand whether students have developed English language skills taught, through which they guided their instruction based on the information obtained. Along with

determining how well the students were able to apply knowledge of a given topic, the instructors also conducted assessment to evaluate the effectiveness of their teaching. This suggests that instructors used assessment as a tool to inform their instruction. In relation to this, Instructor 4 stated:

For me, assessment is more of a checkpoint to make sure that everybody has acquired the skills taught. If students do well on the lesson, we push through. But if not, it informs me to go back and review a little bit, or to do something to get those skills up.

Another interviewee also explained that he had an experience of re-explaining the topic on which the students did not do well and announcing a make-up test when his students showed poor performance on a quiz. “There was some problem; the students did not perform well on the test; so, they needed more explanation.” (Instructor 1) That is, results from the test informed his teaching strategies. Occasionally, assessment was also deliberately conducted by instructors to check teaching effectiveness. A series of post-lesson oral questions were raised purposely for the sake of guiding instructions.

In the context of grammar, one instructor recounted that she usually posed short answer questions, and if the students were up to the point, she built upon it and escaped. However, when not, she was required “to revise my teaching method” (Instructor 5). Adding, “when most of the students failed to respond, I tried to decrease the pace of my explanation, repeat important points, give different examples ...” (Instructor 5). She used this strategy to recognize where and why her students failed to understand the lesson better and modify her teaching methods. As to the instructors, even when their assessment evidence shows that students are doing well, they preferred expanding on the topic than just moving to another point. In such circumstance, assessment was also used to elicit more information and move the lesson forward.

The other assessment for learning purpose type used in communicative English language skills classrooms relates to providing ongoing feedback. That is, the instructors conducted assessment to provide feedback for students. In this regard, Instructor 3 talked about the written feedback purpose behind writing skill homework, saying, “when I give them homework, or assignments, that is to offer detailed written feedback on their learning progresses and weaknesses”. Another instructor also used oral questioning just to provide oral or rarely written feedback to students.

The purpose of such question is to assess them, but not directly. I mean they have to do it, and then I can see the answers based on my own criteria; I mean, I would provide them with feedback on their weakest points. ... So, my purpose of assessment is to offer feedback to students. (Instructor 7)

Generally, the practice of assessment for learning purpose was indicated in five of the interviewed instructors' discussions. Some instructors used assessment to understand the students' level of understanding upon which they determine whether to revise or not their instruction. Others purposely conducted assessment to assess if their teaching was clear as a means to diagnose their strengths and weaknesses in the course delivery. Further, assessment was also conducted by the instructors to provide feedback to students. The findings indicated that results from assessments informed the instructors about effectiveness of their teaching and their students' actual learning status. As such, assessment of student learning gave them useful information that helped them continue to improve in their efforts and students' achievement. So, the instructors used assessment as a tool to facilitate instruction and learning.

Assessment of learning

Assessment of learning is understood in this study as assessment used to summing up learning achievements of students. The purposes of assessment of learning are basically to collect and interpret learning evidences so as to grade and report learning to the stakeholders. Based on the instructors' interview discussions, measuring achievement, reporting, and ranking and grading types of assessment purposes could be identified. Using assessment as a safe-guard was also emerged from the interview data.

Assessing to measure students' achievement is the most common and predominant type of purpose of assessment emerged under assessment of learning theme. Several instructors declared that their main purpose of conducting assessments in English language classrooms was to evaluate students' achievement on specific paper-based tests and examinations. Instructor 9, for instance, illustrated that "I assign written assignment for students to measure their proficiency in writing".

In addition, evaluating students' learning and comparing it vis-à-vis the course goals was commonly reported practices and what the instructors sought to achieve when assessing students.

The instructors maintained that there are set learning objectives that they demanded their students achieve, and they usually assessed the students based on these objectives in order to check whether that happened or not. For example, Instructor 2 stated:

I assessed students to know whether or not they achieved the goal set. ... we want our students, for example, to be able to speak frequently in English; so, when we have assessments, our goal is to evaluate whether or not students can do.

For these instructors, the main objective of assessment was to check how pupils used what they are taught. This is also claimed by another interviewee, who noticed that assessment helped him see just how much the students “integrated” what he “transmitted” to them (Instructor 6).

The reporting purpose of assessment was also the other most recurring type. Seven of the nine instructors indicated that they assess their students because it was required by the university. The instructors noticed that they often conducted summative assessments because they had to submit the results to the department and others in response to the university policy. For example, besides the major purpose of measuring students’ achievement, Instructor 9 assessed students because he “need scores which are required by the university in order to determine the students’ success”. A comment from another instructor further indicated that he usually conducted assessment for reporting purpose through which quality of the program is controlled.

... we use it to provide evidence that students are gaining value from the experiences we provide; you know, it is through assessment that the quality of the program is evaluated, and mostly based on what we report that the departments and others control and judge the quality of this course [communicative English language skills]. (Instructor 3)

Related to the reporting purpose of assessment is the use of assessment evidences for determining the students’ final grades. Several instructors maintained that they needed marks in order to determine whether and how many of the students failed or passed the course, which is also demanded by their universities. One instructor, asserting that he conducted assessment essentially for the purposes of assigning grades, described:

Without assessment, you cannot evaluate and you don’t have the scores for the students because at the end you must produce students’ communicative English grades by adding

all the assessment you gave to them along with the final exam. This is a must. (Instructor 2)

The instructors tended to indicate that there is no opportunity to monitor students' learning progress on a continuous basis through assessments, blaming the established system that emphasized highly on documentation. Instructor 4, for instance, sarcastically reflected the situation saying, "more than we use it for instructional activities, our concern is on what [grade] we report to ... direction of the university". This suggests that the instructors critically needed students' assessment results in order to determine their final grades for the course, which they are required to submit at the end.

Finally, conducting assessments to defend themselves was also emerged in two of the instructors' interview data. The instructors argued that they collected students' assessment results in order to shield themselves against any accusations. For example, for Instructor 1 assessment serves not only to evaluate students' achievement but also to record their learning gathered through formal paper-based assessments, which could allow him provide as evidence whenever required, as "a safe-guard" in his own terms. Instructor 7 also shared his experience as follows:

... I was accused of making fault for the reason around that 40% of the students in a class failed.... No any information regarding how many percent should pass or fail; not on legislation. ... I was asked to show them all the assessment papers; seven assessments were requirement. Fortunately, I able to show them all and defended myself. Imagine what happened if I couldn't provide all the evidences.... After that, my primary concern is to protect myself anytime I assess learners.

Such practices may push instructors to emphasize on the summative paper and pencil tests to measure students' achievements. This is because to defend him/herself, the instructor should able to provide tangible and concrete evidences of students' assessment results if complaints either from the students or from the university management appear. As a result, the use of assessment to facilitate learning can be undermined and forgotten, and summative assessment that is basically to evaluate students' learning may become dominant institutional culture.

To sum up, to measure students' EL skills achievement based on end of term, class, or lesson tests and examinations was the most common purpose of assessment for the instructors. For

better or worse, the instructors are always required to provide evidence that their students are gaining value from the experiences. So, conducting assessments to come up with the final grade, which is seen as a prime indicator of students' success or failure to the course, they are required to communicate to the relevant principals became the most common practices. This suggests that instructors are expected to emphasize on accountability purpose of assessment.

Assessment as learning

The final major theme that emerged from the instructors' interview data regarding the purpose of assessment in communicative English skills course was in relation to the potential benefit involving students in the assessment bears- technically termed 'assessment as learning'. In this study, it is understood as assessment basically conducted by emphasizing on the role of students in order to help them self-monitor and self-correct their learning. The main objective of such assessments is empowering students so that they can make decisions and solve problems independently being critical assessors of them-selves.

Reporting that effective assessment practices are those that empower and aid in the development of student autonomy, two instructors indicated that they also conducted assessments to help learners recognize what aspects of their own work need to improve. One of the instructors recounted that rather than correcting students' papers for himself, he occasionally allowed students to edit each other's work by choosing their peer editors for the writing task they are assigned. Such practice was to help the learners "develop their English skills by relying on themselves by engaging in the independent editing. ... due to this, marks are not given" (Instructor 4). Thus, the purpose of his assessment was just to empower learners by actively engaging them as assessors. Another instructor also conducted assessment to facilitate self-learning of students. She explained what goal she wanted to achieve by allowing students to comment on each other's presentations as follows:

First of all, my interest is to help them that they understand everything by themselves. The other thing is to promote learner's autonomy; this is when instructors encourage students to depend on themselves in learning and answering questions. (Instructor 5)

The interview data regarding the instructors' purpose to conduct assessments in the target course indicated that it is only among few instructors that assessment is used to support students'

independent learning in communicative English language skills course. This suggests that the use of assessments to encourage learners to take ownership of their learning in the target course was restricted.

4.3.1.2. Assessment methods

This section generally reflects on the second research question that sought to understand the assessment methods employed while assessing students' communicative English language skills. Analysis and findings from instructors' interview responses are presented in order of their frequency of employment, from the most to the least frequently used assessment methods.

The conventional assessment methods: written test, mid and final exams

The interview findings indicated that, to a great extent, traditional form of assessment dominated the instructors' assessment practices, given various reasons. To be specific, written tests, mid and final exams were used by one hundred percent of the instructors. All of them even admitted that these are the main assessment methods used in their communicative English skills classrooms. For instance, instructor 2 asserted, "we give students tests including midterm and final exam". He also tended to describe that these are exclusively used in his classrooms. In other words, these methods are applied to assess all the language skills.

Many instructors also confessed as these methods are mainly used to assess the students' grammar, vocabulary, writing, speaking or reading proficiencies. Some instructors maintained positive perspectives towards the use of the conventional methods. They rely on the methods since they benefit weaker students, they are the easiest and fastest methods, and they are less demanding than the other assessment methods. Their efficiency in measuring student achievement and capability in preventing plagiarism were also reasons for their reliance on tests and exams. So, these methods were what worked for the instructors.

However, although commonly practiced in the universities, some instructors doubted the effectiveness of these methods of assessment, illustrating that they hardly accounted for a holistic observation of students' progress. One senior instructor commented she "observe an issue from one angle, assessing students' achievement of the texts that are taught through midterm exam and final exam and ... we do not assess the performance of students" (Instructor 3). Similarly,

stressing on the importance of utilizing differentiated assessment methods with respect to diversity of the students, Instructor 7 insisted that this is not applied in his university. He asserted that he relied highly on conventional assessment methods “which are not appropriate even for schools, but it is what was accustomed in this university”.

Despite their interest and philosophy, many instructors seemed keeping the utilization of formal tests, mid and final exams to measure students’ understanding and to judge their learning through grades. Their persistence likely stemmed from the top-down nature of control in the structure of higher education system. Apart from implementing what came to them from MoE, the instructors may not have a control over those practices. In line with this instructors’ autonomy in terms of students’ scores and assessment methods, Instructor 6 illustrated:

We are implementers; distribution of scores depends on the rules and regulations of the MoE. ... students had one midterm and one final; there were also who test students based on a final exam. But in the past few years, students obtain 60% to 70% of their scores through continuous assessments and the rest for final exam. But in the new system, we are required to give one or two tests, a mid and final exams.

It can be understood from the above excerpt that whether single or multiple assessments are adopted to assess students, it is apparent that the traditional methods are commonly employed.

The analyzed documents similarly depicted that the conventional assessment methods are the predominantly used method in the course. According to the information gathered from the communicative English language skills course outlines, tests/quizzes, mid-term and final examinations are found to be the most recurring methods. All the instructors indicated through their course outlines that they employed tests and final exams to assess students’ communicative English skills. Except one, all of them also mentioned midterm exams in their course outlines. Compared to the other methods encompassed in the course outlines, more than half of the assessment methods planned (54.7%) is based on written tests or quizzes, mid-term and final examinations.

Informal oral questioning

The other regularly implemented assessment method in communicative English language skills classrooms was asking informal oral questions. This was found to be frequently recurring method in all the nine interviewed instructors' discussions. Such questioning, which were targeted at individual students or the whole class, were used at the beginning, during or at the end of instructions. Thus, oral questions tended to serve multiple purposes including, to elicit information, monitor student understanding, evaluate instruction, and, for few instructors, to create opportunity for the language use and manage behavior.

Instructors regularly used questioning method in the course in order to elicit students' understanding about the topic. In the context of reading skill activity, Instructor 3, for instance, reported that "prior to reading activity, I ask series of questions to let students elicit information they have about the topic which informs me their understanding level". Instructors also used this formative questioning for making instructional decisions. They often used it as a powerful tool to recognize students' learning difficulties apart from assessing retention of the day's lesson. "Oral questioning is the fastest method to get learners' feedback on my teaching" (Instructor 5).

Instructors further conducted oral questions to widen students' opportunity of exercising their English language communicative skills in the classrooms. They maintained that their use of oral questions gave their students the chance to use the language in the classrooms. In this regard, Instructor 1 maintained:

... as the course emphasizes on developing learners' English skills, [oral] questioning for us is more than understanding what students learned or not and ..., it is used to create opportunity for students to use and practice the language in the classroom.

According to the instructors, although their use of oral questions was to enhance students' oral skills and to help them overcome fear of using the language, many students did not take advantage of this because of fear of using the language in front of their classmates. This suggests how external contextual factors inhibit EFL instructors from implementing their beliefs about assessment.

Observation

Observation was also the other assessment methods instructors commonly used in the target course. Many of the interviewed instructors indicated that they frequently used classroom observations as a method of assessing students' learning mainly to gather feedback on their teaching strategies and on what was happening in the classrooms including students' participation. According to Instructor 5, such feedback from observation was vital to "adjust" her teaching or "to provide timely interventions" to support her students' learning. Similarly, Instructor 7 usually used observation since it was "the easiest way of obtaining feedback on my teaching". He recounted, "... for example, while I was teaching, their [students'] body language tells me that they are tired, so I was required to adjust my teaching strategies to engage students in learning". That is, using observation, the instructor could understand what students transferred non-verbally that helped him take the necessary measure.

Further, observation was also critically used by instructors to monitor learning particularly during students' group discussions. Instructors always visited each group to observe and listen to their discussions through which they detect students who were not actively involved in order to engage them back. Observation was even used to record students' class participation, which is finally allotted scores and included in grading their achievements (Instructor 3). He explained that this was done to encourage his students to commit to study during the whole semester. Generally, classroom observation was often used by instructors so as to intervene the instruction as needed and keep the students in the learning track.

Authentic and performance-based assessment methods

In the instructors' interviews, performance-based assessment methods that require students to produce language through writing, speaking, or doing something else and prepare them for practical situations were indicated as utilized. Among the few instructors that used such methods, Instructor 7, for instance, explained:

The good thing is promoting long term learning. You see, our learners are future managers, doctors, engineers, or ... and will be faced with English skills related questions on a daily basis, and if we use assessment methodology that only demand to memorize, they will not be prepared for the situation. So, sometimes I use methods that is based on practical things and prepare them for the real-life context.

In the same vein, others used some methods as it is the case of practical work in groups or individually that encourage students to develop and demonstrate the language skills. It was indicated that written assignments and oral presentations were rarely used to assess students. Instructor 2 creates small groups and gives “one assignment in groups on the grammar points of the course ...; they present at the end of lesson”. Correspondingly, Instructor 5 asserted, “I assign students in an assignment that is presented and submitted because a lesson that I taught may need some revisions from the students’ part”.

The instructors’ interview data also showed tendency of using authentic assessment methods. The use of role plays was raised although only two of the nine interviewed instructors declared that they employed the method to assess students’ oral English language skills. According to the instructors, at least one role play is used in a class to assess students’ English language performance. Instructor 4 pointed out that he organized the class into small groups and required each of them to prepare a case on functional languages such as greetings, introducing, advising, and perform in the classrooms. On the other hand, Instructor 7’s use of role play seemed to align with his teaching philosophy for he also valued “authentic learning” and learning by doing. His reasons for using role play were to help students practice communicating using the target language and become immersed in real-life situations. “Using role play activities ... also allowed me realize the skill gap and adjust them if necessary” (Instructor 7). Besides for evaluation purpose, this instructor employed the method to evoke students’ existing performance and to adjust his teaching accordingly.

As data from the documents analyzed revealed, there were instances of performance-based assessment methods that require students to produce language. In this regard, the course outlines collected included individual or group-based written assignment, oral presentation, and project work that prepare learners for practical situations. Other authentic or performance-based assessment method types such as interview, role play, and debate were not listed in any of the outlines. Individual or group-based written assignment was found to be the most frequently recurring method, which was mentioned in all the eight documents. On the other hand, oral presentation, project work, and portfolio were respectively mentioned in three, two and one of the course outlines.

Overall, it can be understood from the interviews and documents analyzed that authentic and performance-based assessment methods were rarely used. This suggests that it is hardly possible to verify that students can apply what they have learned to the real-world setting through existing assessments.

More student-led assessment methods used: peer and self-assessments

Peer and self-assessments were not common practices amongst most interviewed instructors. Assertion by one instructor describing the use of self and peer-assessment as “exceptional practices rather than accustomed” (Instructor 3) in his communicative English language skills course can be extended to most of the instructors interviewed. Some instructors did not employ peer and self-assessments, as they had no trust on such methods since students do not fairly assess themselves and each other. Instructor 1 and Instructor 3, in this regard, reasoned out that “everybody merely assigns high marks for each other when allowed to do” and “because they are classmates and they all wanted to gain high marks, they do not evaluate each other appropriately”.

However, some of the instructors had positive opinion about peer and self-assessments. They consistently argued that such practices in assessment motivates students, makes them feel part of learning and promotes their autonomy. Nevertheless, they raised constraints in relation to time, work load, student resistance, and the equal status between peers prevented them from conducting peer and self-assessment in communicative English language skills classrooms.

As to the peer-assessment, activities some instructors refer to was more of peer-feedback that students provided while presentations or performances, and many admitted that they did not allow students to mark or grade each other’s work, for example. In this regard, Instructor 6 responded, “I provide them opportunities that they can make comments on their classmate’s presentations”. The peer-assessment instructors described to have used also seemed to be mixed with the idea of group discussions. Nevertheless, in few of the instructors’ interviews, actual practice of peer-assessment could be realized, which they indicated that they rarely used in their communicative English skills classrooms. For example, Instructor 5 involved students in “assessment marking” by allowing them to “evaluate each other’s work”. In another interview, it

was similarly indicated that students are asked to exchange their papers and cross-check answers for the written questions instructor provided:

“I give them a set of grammar questions on the board; ... then, I order them interchange their answers. Finally, the students correct each other’s answers and assign total mark based on the class-wide discussions we make on the questions”. (Instructor 9)

Comparable to peer-assessment, current use of self-assessment often remains a relatively small part of the assessment practice in the universities. It was only among two instructors that self-assessment was reported to be incorporated into instructors’ assessment practices in communicative English language skills. Instructor 8 utilized self-assessment because it transmitted message to her students that she trusted them and “to encourage students to depend on themselves in learning and answering questions”. Instructor 4 also allowed students to mark their own group’s performance. Regardless of students’ inaccurate self-assessment results, the instructor continued to use the strategy in the classrooms along with teaching his students how to self-assess. Although self-assessment was incorporated in the form of self-evaluation activities both at individual level and within students’ groups, he insisted that he does not give any marks for such activities.

In similar manner, peer-assessment and self-assessment are rarely mentioned in the instructors’ course outlines. Regarding peer-assessment method, it was mentioned only in two of the eight course outlines collected. In other words, most of the instructors did not include peer-assessment in the assessment methods planned in their course outlines. When it comes to the self-assessment method, it is non-existent. None of the eight course outlines collected mentioned self-assessment as a method of assessing students’ English language in communicative English skills course.

To sum up, the qualitative data collected through instructors’ interviews and document analysis revealed that the instructors’ practices of assessment were dominated by the traditional forms with more reliance on the paper-based tests, mid-term and final exams. In such situation, assessment is hardly aligned with instruction and students will have little opportunity to holistically develop the language, as such methods focus on only particular skills. It could also be realized that authentic and performance assessment methods were used only rarely although many instructors often acknowledged importance of such assessments. That is, methods that

require students to produce language through writing, speaking, or doing something else were not emphasized in assessing students' communicative English skills.

4.3.1.3. Provisions of assessment feedback

This section reflects on the research question that sought to understand how feedback is practiced in students' communicative English language skills assessments. Accordingly, analysis and findings from instructors' interview responses and document analysis that emphasized on mode, source and focus of assessment feedback will be presented.

In their interviews, all the participants indicated that assessment feedback is carried out in communicative English skills classrooms. While some instructors responded to have used both oral and written forms and instructor and student-based feedback types, others' responses seemed to indicate that they did not use all these recommended forms and types proportionally. Instructor-led assessment feedback provision was found to be the most recurring type in the interviews. The students relied heavily on feedback from instructors, and this put the instructors under pressure to provide feedback by them-selves than allowing students (e.g., Instructor 2, Instructor 5).

Particularly, instructor written feedback appeared to be more preferable, and the interview discussions illustrate the dominance of this form of feedback. All the instructors reported that they used written comments as feedback for every assessment activity of students. Although some instructors believed that it was good to frequently practice feedback of any kind to enhance students' learning, their feedback "emphasized more on written comments", reasoning out "students' linguistic difficulty in comprehending oral feedback" (Instructor 8).

As in the case of selected response assessments, while just correct and wrong signs are merely used on students' assessment papers for most instructors, few of them reported to provide more written comments in the margin and as endnotes. However, irrespective of these instructors' differences in the magnitude of details of feedback they provided, they all described to have detailed comments for their students as in the case of written performance tasks. For example, Instructor 7 asserted, "I indicate what they missed and how could be corrected in detail when I correct students' tasks about sentence, paragraph or essay writing." Two other instructors also tended to show that they spent a lot of effort in providing written feedback for students.

Oral feedback was also the most recurring type raised to be used by the instructors. They all claimed to use oral feedback that targets whole class or individual students in the classrooms. In both situations, the instructors maintained the effectiveness of the approach. For example, Instructor 9, who emphasized on the use of oral feedback, commented: “I get very hard to communicate in written form what I get to do it orally”. Others, however, cautioned its large-scale use on the basis that “students can forget many things” (Instructor 5). This whole class based oral feedback was often used by instructors when they returned the commented and marked students’ papers and following students’ presentations and group activities. The instructors also used this type of feedback when they felt students needed more elaborations on certain areas and upon the students’ request.

Individual based oral feedback was also reported by two instructors to be less frequently employed. Instructor 5 pointed out that she used such approach as a “dialogue with individual students in classrooms and outside”. Both instructors indicated that individual-based oral feedback was mostly conducted based on students’ request. Instructors who did not use individual oral feedback or did so infrequently, but still believed in its effectiveness, blamed this on workload, class size and time constraints.

Despite the dominance of instructor-based feedback practices, there were also instances in which students were provided opportunities to practice self and peer-feedback. Concerning peer-feedback, for instance, an instructor stated, “I also sometimes give them the chance to comment on their peer’s papers when I give them back their exam papers” (Instructor 9). It was also clarified that peer-feedback is utilized during presentations, written assignments and class activities. Along with this, Instructor 4 commented, “yes, students took part in feedback provisions; they have to do this before my feedback. Mine comes usually at end whenever they have class activities.” In addition to inviting students to comment on each other’s works, Instructor 4, who is enthusiastic about peer-feedback, also used different approach. He clarified that he sometimes provided anonymous texts for students upon which they are required to give their comments, although this was limited to written texts including sentence and paragraph correction, focusing on structure and grammar.

On the other hand, Instructor 6 maintained as he did not engage students in feedback provisions, reasoning: “it was hard for poor students to give feedback for good students”. Other interviewee,

however, used aids such as assessment form having checklist for guidance to overcome a barrier mentioned by Instructor 6. The assessment form, as he claimed, was set up based on criteria the instructor himself specified, and “peers write their comments in a form provided” (Instructor 7).

Regarding self-feedback, however, it appeared that most interviewed instructors did not incorporate in their instructions. It was apparent in Instructor 4 and Instructor 8 discussions, who reported that they seldom used the technique. Instructor 4 employed it in writing activities “to help learners identify their own weaknesses and strengths by themselves”. Both instructors reported that they provide assessment criteria to students based on which they self-evaluate their work prior to submitting it.

The interview data, furthermore, showed that EFL instructors simultaneously employed verification, directive and scaffolding feedback types although frequency of use of each type varies. The interview data indicated that verification feedback was the predominant feedback type employed among the instructors. The most recurred types of verification feedback raised by the instructors were the use of correct and wrong signs and assigning scores on students’ assessment papers as judgment of their academic performance. For instance, Instructor 9’s feedback was mainly to indicate learners’ “mistakes ... I mean, by crossing where they failed to appropriately respond ... put total marks on their tests”.

The purpose of such feedback is pointing out where students are right or wrong. In this regard, Instructor 2 elaborated that “using some marks I must indicate which questions are correctly answered and which are not whenever I check their papers”. For some, the feedbacks focus more on pointing out students’ weak points, which was particularly employed during oral presentations. Instructor 5 provided oral feedbacks at the end of students’ presentation, “like on their grammar usage, pronunciation, and other weak points”. The reason for feedback to focus on students’ weaknesses was because “students will try to improve their learning when they know their weaknesses”. (Instructor 3)

The target of such instructors’ assessment feedbacks was on providing information that describes whether or not answers are correct by telling or marking wrong on the locations of the errors and telling or marking right on those properly answered. Analysis from the students’ assessment papers similarly showed instructors’ widespread use of such assessment feedback types. This

assessment feedback type covered more than half of the feedbacks included in the assessment documents analyzed, that comprised of tests, major exams and assignments collected from the instructors. Out of the 681 feedbacks provided, 479 or 70.3% of them were verification feedback. The documents revealed classroom-based assessment feedback culture that revolves around verifying the right and wrong responses and justifying marks.

Directive feedback, which refers to instructor-led direct written or oral response provisions, was also the other dominantly employed feedback type, as it was found to be the most recurring in the instructors' interviews. That is, the instructors often focused on directly providing and supplementing the correct answers for the students' incorrect and incomplete answers. The instructors conducted such direct supply of the correct answers to a given question either through telling or writing them when they turn students' papers back. This was a habitual practice for Instructor 9, whose feedbacks focus on providing answers directly to students "by jotting down the correct answers on the board ... to allow crosschecking". Such directive feedback was also utilized while correcting students' papers and during presentations. "When I correct papers, usually what I do is indicating ... and also I have to put the correct responses." (Instructor 6) In writing tasks, Instructor 8 also recounted that she immediately corrects it when students provide "partial information" to their writing and "give them written feedback straight away".

The analyzed assessment documents also revealed that directive or corrective written feedback was the other feedback type used by communicative English skills instructors. Although it was not as frequent as the feedback that focuses on verification, instructors provided directive written feedback on students' assessment papers. The data from the documents reviewed indicated that direct correction of students' errors and supply of correct answers to students' incorrect or no responses were the second largest feedback type employed. This is represented by 137 or 20.1% out of the total 681 feedbacks provided.

Finally, as for scaffolding feedback, the prevailing situation in this study is that it was appeared not to be adequately utilized by the instructors, on the basis that students are not ready to be challenged and claiming their reluctance. In the interview responses of few instructors, it was indicated that they sometimes focus on directing and guiding students to independently generate answers to questions. The instructors conducted such feedback under two different conditions:

one is when assigning tasks to students and the other is when they are working with the activities or assignments.

For instance, while assigning writing activity to students, Instructor 7 provides them some “basic terms” for that particular topic, which helped him to “motivate learners and easily get them into producing the texts”. He further stated that he particularly utilized such feedback with students with lower proficiency in English, as “they need extra support” (Instructor 7). The scaffolding feedback was also utilized while students are working with their tasks or assignments. Instructor 4 asserted that “learners have opportunity to consult with me and I give them hints; tell some techniques that may guide how they need to pursue their work”. This means that the instructor allowed students to interact with him and support them as they prepare their assignments.

Correspondingly, the analyzed documents indicated that EFL instructors rarely provide scaffolding or facilitating feedback, which focuses on supporting learners to complete or solve a given task independently, on the students’ assessment papers. This feedback type had the least coverage among the feedback types provided in the documents. Out of the 681 feedbacks observed in students’ tests, exams and assignments, only 65 or 9.6% of them were scaffolding or facilitating feedback type (details are found in Appendix L).

Such scaffolding feedback types provided in the documents were presented usually in the form of signs and at the level of words and short phrases giving clues or hints to students to revise and retry the questions they either incorrectly or partially answered. For instance, brief comments such as “what?” and “second major point?” or just question marks (“?”) were provided. Such feedbacks, though implicitly, were to drive learners toward re-carrying out the tasks and further exploration. In other words, the instructors hardly used descriptive elaborated feedbacks that more encourage further learning as such feedbacks were unevenly utilized in the assessment documents analyzed. Moreover, it could be understood from the documents that assessment feedback provisions were taking place mostly from closed questions, and feedbacks to open-ended questions seem to be given little attention by instructors.

4.3.2. Assessment assumptions

This part of the study deals with the second basic research question that emphasized on exploring assumptions underlying assessment practices in communicative English language skills course.

The interview responses from instructors and students are organized according to knowledge/content assessed, value and judgment, and power and control themes. Most of the assumptions under each theme are dichotomously paired with each other, some being truly bipolar and others not. Hence, although the assumptions in each pair are presented separately, some of them are not mutually exclusive and can be held simultaneously by the same participant and acted upon in response to a specific context and situation.

4.3.2.1. Assumptions about content assessed/knowledge

Focus of assessment on knowledge construction process

The view that assessment should engage students in the knowledge construction process rather than in repeating the existing one was widely reflected in the participants' reasoning although most were from instructor participants. This was described in relation to assessment purpose, method selection, task assignment and view of good/bad work. For example, the instructors were critical of using selective item exams since "they foster memorization; just recall of facts" (Instructor 4) and preferred performance-based assessments as they "lead to greater learning experience" and provide opportunity to see students' "knowledge creation ability" (Instructor 2). This view was also reflected in their perception that a good work lies in the students' ability to reveal the unseen perspectives by justifiably presenting them rather than in their ability to duplicate the established knowledge. Instructor 8 expected students to take a critical approach in writing assignment because students should "not be allowed to copy and simply follow the way presented somewhere else; they have to produce unique piece of work. ... You have to force them to show something genuine, different views".

The participants further believed that assessment should monitor students' ability to develop their own understanding by transforming and reorganizing existing knowledge. Describing that the main goal of the target course is for the learners' academic success, Instructor 5 maintained, "what we do through assessment is ensuring that our learners can take over what they have learned to ... based on understandings they developed". Another instructor recounted that he engaged students in writing tasks describing a particular building in their campus after exposing them to a lesson because "they must be asked to shift their understanding to a new context. This requires them creativity" (Instructor 7). These instructors recognized that learning is ensured

when students ably develop their own understanding based on the established knowledge. Thus, what they assessed focused basically on knowledge construction processes, which signify their expectation for students to be creators of knowledge.

However, such views are hardly reflected in students' interview data. Only two of the student participants emphasized in their discussions of good/bad work or answer that students have to produce a unique piece of work rather than repeating what is already established. Their reflections suggested that they are expected to construct knowledge that should be accounted for in the assessment process. For instance, Student 11 felt that good answer is that:

gives a detailed explanation and has my own examples ... I don't think that a good answer when we come with assignment with a list we learnt in class and something provided in the handouts. ... A good answer should be that tells the instructor that I really had grasped the grammar points.

Assessment focuses on students' ability to recall the contents taught- Knowing is reproducing

The participants' rationales behind their assessment decisions and practices also illustrated the assumptions that there are essential and fixed contents that students are expected to accumulate, and that assessment measures students correct reproduction and straightforward use of these contents. A response from Instructor 1, "I wanted them [students] to bring the essential points that I taught them during exams ..." best illustrates the view. The notion was also reflected by Instructor 9, whose oral skills assessment usually emphasized functional languages and tasks covered in the module or practiced in the classroom. He stated, "Students have to follow those best expression and procedure examples if they are wanted to be speakers of the language like natives" in justifying the practice.

Such anticipation for EL learners to accumulate rather than to produce knowledge was further signified in relation to answers expected from students. Reasoning of two instructors suggest that there are verified ready-made answers that students should adhere to while answering the given questions. For example, Instructor 6's rationale for correcting students' answers to class activities was to ensure that they have "the right answers that they need to use in the future perhaps for the mid or final exam". That is, each question of the instructor has one acknowledged

correct answer that he equips students with, allowing no room for the emergence of varied perspectives.

The two instructors' view of uncontested nature of answer for what is assessed implicitly indicates the need for assessment to focus on measuring student's ability to correctly reproduce established knowledge. When students are taught in a setting of certainty, where they need to follow set generalized rules, problems are thought to have homogeneous solutions. Hence, instructors equip students with those solutions to build correct knowledge which is universal throughout contexts. What is assessed, thus, expected to have one verifiable correct answer. In such environment, instructors assess to ensure that what had been taught was accurately retained, and students are required to recall and demonstrate the pre-defined universal knowledge.

When it comes to students, the assumption that assessment should measure students' correct reproduction and straightforward use of the essential and fixed contents that they accumulated was remained dominant throughout their discussions. Several students tended to describe that assessments that require students to correctly reproduce information presented in lectures and modules are considered good, and they believed that able to do so is seen as a sign of learning. In this regard, Student 2 responded that he considers it good when assessment "requires us to apply what we have learnt from teacher exactly the way we have been taught".

The students further expressed the idea that any question has one acknowledged correct answer that is already known and predefined. Student 8 commented, "... teachers should prepare exam questions that enquire students to answer regardless of differences." In the same vein, while answering to the question the researcher posed on whether there could be two or more differing but valid explanations for the same question, few students responded that this should not exist. For example, Student 4's rationale was that "one question must have one right answer." Based on such taken for granted beliefs, they insisted that question with verified one clear answer needs to be designed. It was also remarked that there are standardized approaches that students should adopt for a given question that enable them to arrive at a predefined clear answer. "We have examples in our handouts that we need to follow during speaking tests; so, we cannot argue saying this is the correct and that is not." (Student 12) That is, assessment tasks that demanded students one correct way of answering are more valued.

Generally, these instructors and students equated learning with correct repetition of established knowledge. They also showed that the soundness of students' answers is to the extent that they align with the contents (factual and procedural knowledge) as presented in module, textbook, or lecture: apple to authority. Such practices of assessment explicitly define the existence of universal knowledge and the specified form in which it is demonstrated.

Teaching and assessing just one skill at a time

The reasoning of some participants of this study also tended to reflect the assumption that language is a collection of component skills that are taught and assessed primarily independently. These participants were also certain that students' communicative competence can be captured in additive assessments of writing, grammar, reading, vocabulary, and other discrete points of the language, and the development of each skill needs to be ensured before moving to the next.

For instance, independently assigning values and separately assessing each language skill and sub-skill were common practices among instructors. Although some admitted that the practice hardly accounted for a holistic observation of students' use of the language and held the university assessment system and students' ability responsible, others maintained positive perspectives. Instructor 1's rationale was that "the skills are taught separately. ... Similarly, that [testing the skills separately] should follow; allow me to accurately measure their capability of using the language." This led the instructor to assess students' English communication skills by focusing on discrete skill.

In addition, three other instructors thought that each skill needs to be assessed discretely since this allows them to look for efficient transfer of the skills taught. In this regard, Instructor 6 contends that "by mixing these skills in testing, you may lack focus; your understanding of students' extent of achievement comes shallow; you can't grasp deep knowledge of their level of learning on each skill". Instructor 3 also held the idea that language components should be treated separately.

My assessment usually focuses on all skills. I usually allot ten points for each of the skills. That is how I design ... for example if I assess reading by the first test, my next test may be on writing skill, then grammar and it continues. I have to do this way to ensure students understanding of one skill before jumping to the other.

Consistently, the students also indicated that the language skills need to be separately assessed in order to effectively measure their English language ability. Resonating with this idea, Student 11 strongly believed that “separate skills assessment is a must to well understand how far we retained what we had learnt and allow us to easily learn the language step-by-step”. In addition, based on the thinking that assessing the skills holistically is unrealistic, two students also demanded to be assessed on each skill. For example, Student 8 reasoned out that since “assessing all the skills in one spot is not suitable, I think ... fair to focus on each skill one after the other”. Overall, such participants’ reflections further suggest the implicit assumption that complex knowledge should be broken into its components to effectively transmit and assess its efficient transfer.

Integrated skills assessment

For some participants, assessment of English language that integrates different skills together was reasonable practice on the basis that such assessment takes students opportunity of using the language into account. However, this view was held almost exclusively by instructors, as it was only evidenced in one data of the interviewed students. The instructors placed greater emphasis on a more holistic approach to language acquisition and assessment, in light of communication function of language. They pointed out that the current communicative approach to language teaching necessitates assessment to focus on students’ ability to use language skills comprehensively. For example, Instructor 7 stated, “... communicative competence requires such integration; ... it can’t be captured in separated tests of the skills...”. This assertion denotes the instructor’s reliance on integrated skills assessments in order to effectively capture students’ communicative competence.

Furthermore, the instructors realized that effective assessment should be mock of real-life situations, so it requires students to apply several abilities. Instructor 5’s assessment required students to deal with complex tasks that demanded listening to news, writing a summary, and presenting it in the class. Her rationale was that “the task encompassed various skills in one shot because natural setting imposes bringing different skills and abilities together to use the language. That’s the way I feel right.” In his recount of ineffective assessment practices, Instructor 4 similarly illustrated his view as follows:

commonly we often teach and evaluate each skills distinctly. I don't know where the practice emerged, but I see it malpractice that goes against the nature of the language itself that naturally works as a complete system. ... must be in ways language is used in real life. It needs to work on it, to disconnect from practices that treat skills independently.

However, only one student participant shared the idea of integrated skills assessment. The student raised the view in relation to how the language works in real-life situation, saying “we have to be assessed in a way we use the language outside classroom.” (Student 1) According to her, assessment should integrate language skills rather than focusing on isolated skills.

In summary, these participants shared the idea that it is better for assessments to become more integrated and stop being so fragmented into skills. Their reasoning behind their assessment decisions and practices appear to reflect the assumption that integrated skills assessments that account for a holistic observation of students' use of the language are effective, suggesting their holistic view of knowledge. However, it should be noted that the assumption was rarely evidenced in students' accounts.

4.3.2.2. Assumptions about values and judgments in assessment

Objectivity of assessment

The assumption that students' learning can and should be measured objectively is the most widely reflected among the study participants. The participants tended to reflect that the assessment process and the final grade are carried out in a value-free approach and information from students learning needs to be scientifically objective. According to the instructors, this is thought to be maintained through communal practices, using structured tools, quantitative methods and by restricting students' involvement. Such approaches allowed them to overcome possible sociocultural influences on their assessment practices and to objectively measure what a student has learned. For instance, in response to why exams are so a collective action, Instructor 6, department head, responded: “university necessitates objectively measuring learning ... crucial for large student population”. Correspondingly, Instructor 9 described:

In this course we need to ensure that we are able to objectively measure what a student has learned. That is where we reach consensus with students and others. The marks should be free from any influence. ... we set some 15 objective questions in mid exam, each having two marks. ... leave no room for our subjective judgments.

The intention to objectively measure learning was also evidenced in the instructors' justifications to rely on paper-based examinations and to distrust open-ended items, that is, trust in highly structured instruments. Their reliance on examination relates to the rationality of its effectiveness; it permits "accurate evaluation" and allows students to gain clearer information on their learning than other methods (Instructor 3). On the other hand, open-ended questions were seen as unfair because they are not appropriate for "objective scoring ... and bear biases" (Instructor 2).

Further, the instructors seemed to place greater faith in quantifiable data and numeric descriptions of students' learning in order to avoid subjective bearings and provide reliable information about grades produced. They held the belief that assessment needs to result in numeric data in order to objectively and effectively measure students' communicative English skills learning. A comment from Instructor 1, "I have to evaluate them by grades ... objective expression of how well they are doing. So, numeric value is a must. ... otherwise, you can't produce sound grades" illustrates the view well. His rationale was to produce unbiased grades that could accurately express students' English language abilities.

The final group of constructs that reflects the assumption of objectivity relates to involving students. For example, the instructors were unwilling to involve students in the assessment processes, as they are concerned about students' ability to accurately carry out assessment issues. They justified their concerns by drawing on the view that students' engagement in assessment precludes the possibility of producing valid learning results. This also presupposes the assumption that involving students in the assessment process renders the results invalid. "You may not involve them because you don't need to influence the results." (Instructor 3) "Students are inaccurate while evaluating themselves. ... It makes impossible to get accurate information about the marks I need at the end." (Instructor 8) Such assertions specify that students are not made to have a significant role in assessment because of the instructors' motive to have value-

free assessments and lack of that from the students' part, suggesting the instructors' assumption that they are able to objectively measure what a student has learned.

The student participants have also widely endorsed and shared the notion of objectively measuring learning. For instance, regarding the types of questions to be embraced in their assessments, many of them indicated as they prefer if open-ended questions are not included. They justified that such questions are not appropriate "to accurately measure our skills" (Student 7) and "because lead the instructor himself to biases" (Student 2). These students indicated their expectations for assessments to measure their learning objectively.

This view of value-free assessment was also raised in relation to bad assessment they had experienced. In this regard, two students showed their objections to the practices in which students' efforts and progresses in learning are included while making grades. According to them, grades that constitute only formal assessment results are fair, and inclusion of efforts and progresses they made in learning in grade calculation results in inaccurate outcomes. "You should not be evaluated because of something good you have done during the term or your efforts; that is not fair." (Student 2) Student 4 also stated, "Unless my results from assessments alone are summated, the grade cannot truthfully represent my performance." The assumption that assessment objectively measures students' learning seems to shape these students reasoning behind grading practices.

This assumption was also evidenced in two students' discussion of their hesitance to take part in the assessment processes, justifying that their involvement may result in inaccurate and invalid decisions. This was evidenced in their complaints with the subjectivity with which their instructor corrected assessments. They demanded that instructors need to be objective by not involving students in the assessment process. In this regard, Student 10 required that her work is "corrected by instructor alone", following her desire to be assessed without bias. This was reinforced by Student 6 in his explanation of how to produce more reliable final grades by involving no one.

I think, their [instructors'] major goal is making quality grades. But if students are allowed to score and give marks, the grades they make cannot be free from error. ... That is why I prefer if we are not [engaged in assessment].

Qualitative descriptions and subjective interpretations of students' learning progress

Regardless of its representation by few comments, exclusively by instructors, it was also reflected that qualitative description and subjective interpretations of students' learning progress and performance is illuminating than numeric descriptions. The instructors felt that, verbal data-based assessment practices offer deep insights for instructors and students regarding students' learning progress. They also distrusted assessments that adopted objective methods and inquiries on the basis that they do not provide a trustworthy interpretation of students' learning. In this regard, Instructor 7 stated,

It is more important to see what students do through like portfolios which carry thorough description. ... if possible, I don't give grades. That's senseless. You may face huge challenge if you use your own judgmental approach to determine their learning but more justifiable. ... If learning is end goal, we must go beyond counting numbers and quit judging learning by grades.

Similarly, questioning the existing system, another instructor suggested that he did not accept the notion of objectivity. "... you do since you have to. The relation between grade a student holds and his actual English ability tells you much how ineffective procedure we are made to follow ... and why I don't trust number-based judgments" (Instructor 4). Being suspicious about the soundness of number and grade-based assessment practices in representing students' performance, what worked for the instructors was judgmental approaches, as these enabled them to more credibly represent students' performance. Even when grades are mandatory, it was remarked that students' efforts and progress during their learning should be considered. This was reflected by Instructor 8, who went beyond ordinary practices to grade students' performance, justifying the following:

I don't rely on those results because an hour exam can't be good indicator of learners' performance of the course. You may call it personal judgment or that I am biasing results, but I consider not just results but the efforts students make ... their progress during learning while giving grades. That I believe more realistically represent their English performance.

These instructors rejected the objectivity notion, since assessments based on such inquiries are less likely to yield deeper information about students' learning. Their motivation to gain deeper understanding of students led them to rely on subjective interpretations of students' learning progress and performance irrespective of the influence of their values, which may also suggest the assumption that inquiry in and interpretation of assessment results are subjective, but more important.

Consistency of assessment

Participants reasoning behind their assessment practices and decisions also revealed existence of high demand for assessments to be uniform. Several participants of this study felt that all students taking communicative English skills course should sit for and pass the same assessment in terms of content, method, scoring criteria and timing. They could avoid favoritism and ensure equality when students are uniformly assessed, which remained central in the assessment practices of communicative English skills course. The participants realized that this is the essential feature of assessment practices in the course that enables and leads to gaining more reliable data. Underlying such practices and awareness is the assumption that assessment measures should be consistent across student populations and contexts in order to gain the most objective data.

Such underlying beliefs about assessment were widely endorsed by many instructors. The instructors felt that any activity in assessment should yield consistency. Specifically, with regard to the question if instructors may allow differential assessments for different students and why or why not, many of them responded 'no' with different justifications. For example, Instructor 3 explained his justification as follows:

... because, on same course I think assessment has to be administered uniformly; test items, tasks, types, time of the tests has to be the same for them to secure reliable results. Could it be reliable if, for example, some students take classroom tests and, simultaneously, others take home take assignments? ... It also needs to make learners not feel distinguished.

In pursuit of impartiality of assessment two instructors also claimed that they do not assess their students differently. "Because you have to be impartial; reasonable if they all be measured on the

same task” (Instructor 8) and “there must be similarity to avoid privileging some students over the other” (Instructor 9). Further, being detached from the mass and assessing one’s own students alone was a concern for an instructor because such practice is “counter to the regulation. ... All conditions must be equal for every student for fair assessment.” (Instructor 6)

The assumption that the course assessment should be consistent across students and settings was also endorsed by student participants. For instance, a student did not want his peers’ engagement in evaluating his work since the practice prohibits consistency. He pointed out, “I want our tasks are uniformly marked and judged by our instructor.” (Student 10) In addition, Student 7 does not feel right unless “we all taking this course are assessed on the same questions and in a similar manner”. The idea of differentiated assessments was also rejected on the basis of maintaining impartiality. This was raised by one student in his complaint regarding administration of alternative tests for different sections. He claimed that “... the tests must be similar for everyone to avoid favoritism” (Student 2) in response to bad assessment he felt experienced in the target course. Generally, central to the participants’ assessment practices and views was ensuring consistency, which they knew was an important aspect of language assessment. Thus, assessment practices that took the ‘one-size-fits-all’ approach into account were what worked for them.

4.3.2.3. Assumptions about power and control

Instructors power dominance

The participants commonly shared the basic assumption that much of the power in assessment needs to remain with the course instructor. This was evidenced in their comments, which showed the course instructor’s sole authority and role as taken for granted in assessments. For example, Instructor 6 stated, “... just teachers know about assessment”. What was observed in instructors’ responses, in which their experiences appeared to impress on them that assessment is instructor’s responsibility, also supported the view. In this regard, Instructor 1 reasoned out, “Because from my experience teachers assess students. That’s what I know. ... You may involve students, but this doesn’t mean that they are in charge of it”. Such authority instructors bestowed over assessment was further captured in their claim that of all educational processes, assessment should be under instructor control. “It is assessment, not another thing, and I would like to say it

is my duty because after all, I am responsible to ... not the students. I can't see their place here.” (Instructor 3)

Even hinting at policy documents, the instructors showed their expectations for the course instructor to take control of assessment. “... those written in legal documents are for instructor, makes you accountable not students ... I mean, whatever ... I have legal reason and responsibility to control it” (Instructor 9). This indicates instructors' interest in sustaining the existing institutionally armored hegemony that privileges them to possess more power in assessment. Through the assertion that “... because there must be some kind of distance between instructor and student in assessment”, Instructor 8 also explicitly indicated her demand to control and have authoritative figure in assessment.

This instructors' privilege over students in assessment was backed up by more than half of the student participants. It was evident in their interviews that they were also happy for instructors to take on the dominant power in assessment. For example, Student 9 confidently responded that assessment is “instructors' work; we used to be assessed by instructors”. This indicates that the student, from her experience, knew that assessment is just the duty of instructors, and this led her to accept the powerful authority of instructors in assessment. In line with this, Student 6 also commented that “I think providing feedback is part of assessment so the teacher has to give it, not us.”

Moreover, students' involvement in different assessment activities was unacceptable practice for two respondents since students lack the knowledge. For instance, Student 5 justified his argument saying, “in this subject I don't think students can assess themselves well enough. It is what our instructor should do because he knows more than anyone else because he is the course instructor.” This shows that instructors are regarded responsible for providing assessment feedback since they are more knowledgeable than students.

Generally, this study participants' reasoning behind assessment practices and views show their demand for communicative English skills course instructors to control power in assessment. Besides, instructors' dominant power and exertion of it against students were also implicitly conveyed through their over utilization of non-performance-based assessment methods that potentially restricted students from expressing their views.

Sharing power

Although the assumption that reflects instructors' dominant power position in assessment was prevailing, few of the interview participants were not in a position to accept. Relating with multidimensionality nature of learning, these participants described that they consider students as part of learning, who able to make effective instructional decisions that can enhance their learning. In particular, instructor respondents even condemned the hegemony instructors traditionally have in assessment and said that they tried to promote democratic assessments in the classrooms by actively engaging students. The assumption that instructors should share power with students seems to underlie such assessment practices and views of the participants, which was evidenced among seven of the twenty one interviewees.

This assumption was reflected in instructors' assessment practices in which students were allowed to comment on each other's presentations, group discussions, and anonymously written paragraphs. In doing so, instructors recognized that students are not passive recipients but are capable of generating and communicating comments to their peers. Instructor 7's rationale was that "their [students'] role is immense. They have a lot to share to each other and must be given that chance". Instructor 4, who employs peer and self-assessments in his classroom along with empowering students by teaching them on how to conduct, similarly justified that "learners are capable to learn by themselves if they are engaged".

This view was further supplemented by another instructor while describing his relationship with students. Reminding students' current power position in education, Instructor 2 commented that "we have smooth relationship; I do not need to dominate them. ... I think the view that since instructors are more knowledgeable, they should control assessment is incorrect. Learners should be part of decisions."

Resonating with this view of empowering students in assessment, two students, who were aware of their essential roles in assessment, shared that learners should be part of decision-making processes. They claimed provision of opportunities for students to voice their opinions rather than waiting for only the instructor comments. Student 10 stated, "It is good if both instructor and students are involved in assessment collaboratively" in his view of good assessment. Student 7 also commented that "sometimes, as we are also doing it sometimes, we must be allowed for

example, to comment on one another's work ...” These students' comments, from the researcher's point of view, relate to claim of reasonable power for students. Generally, despite its representation by few comments, these participants placed greater faith in joint assessment processing, indicating their expectations for shared power relationship between instructor and students in assessment.

4.3.3. Factors inhibiting instructors from interpreting their assessment assumptions into practice

The instructors' interview data signified discrepancies between their assumption in favor of alternative assessment approach and its implementation. Even though the instructors' interviews showed their possession of few assumptions underlying current assessment approaches, they were unable to interpret this into practices and to employ the innovative assessments. That is, some of the assumptions that have roots in the alternative assessment culture that the instructors expressed were left unimplemented. These instructors attributed the reasons for the discrepancies observed between their assumptions and practices to different institutional and contextual factors.

Institutional factors

According to the participants, one potential reason that contributed to these findings is the established assessment system. In this regard, some instructors indicated in their interviews that they had scarce opportunity to monitor students' learning progress on a continuous basis through assessments, blaming the established system that emphasized highly on documentation. For instance, Instructor 4, who faces challenge in translating his assumption, described the situation saying, “more than we use it for instructional activities, our concern is on what [grade] we report to ... direction of the university”. This shows that instructors basically conduct assessments to determine students' final grades for the course irrespective of their assumptions.

This view was also reflected in relation to the way the instructors decide on the assessment methods they use. In line with this, the participants tended to indicate that their university assessment culture led them to reserve themselves from implementing alternative assessment methods in their classrooms. For example, although Instructor 7 had no doubt on the importance of using alternative assessment methods, he relied highly on the conventional methods reasoning

this is “what was accustomed in this university”. That is, he kept the utilization of paper and pencil tests, mid and final exams to measure and judge students’ learning through grades regardless of his interest and belief. It is important to note that these instructors must abide by the assessment plans of their universities.

Their persistence also likely stemmed from the top-down nature of control in the structure of higher education system. Apart from implementing what came to them from MoE, the instructors may not have a control over those practices. In line with this instructors’ autonomy in terms of students’ scores and assessment methods to calculate grade, one instructor illustrated the following. “We are implementers; distribution of scores depends on the rules and regulations of the MoE. ... In the new system, we are required to give one or two tests carrying 20%, a mid 30% and final exam 50%.” (Instructor 6)

As instructors of this study also pointed out, what is viewed as typical in their context influenced their assessment practices. In this regard, although assigning values for and assessing each skill discretely were common practices among the instructors, some admitted that such practices hardly accounted for a holistic observation of students’ use of the language. Instructor 4, for instance, illustrated his view as follows in his recount of ineffective assessment practices:

Commonly we often teach and evaluate each skill distinctly. I don’t know where the practice emerged, but I see it malpractice that goes against the nature of the language itself that naturally works as a complete system. ... must be in ways language is used in real life. It needs to work on to disconnect from practices that treat skills independently.

The instructors could not practice assessment deviating from what is learned, as this could bear criticism and accusation from the management and students. This indicates how their assessment practices are influenced by external factors rather than their personal assessment theories. A point from Instructor 8 indicates how she fears differing from the existing practice:

If you use other methods [performance-based assessments], you are subjected to criticism from students and others. They complain as if you are impartial, inaccurate and subjective in judging their grades. So, you must adhere to the trend.

Two instructors also tended to specify that such institutional challenges forced them to employ assessments that adopt objective methods and inquiries irrespective of their thoughts. Instructor 7 stated,

If possible, I don't give grades. That's senseless. You may face huge challenge if you use your own judgmental approach to determine their learning, but more justifiable.... we must go beyond counting numbers and quit judging learning by grades.

Similarly, questioning the existing system, another instructor suggested that he did not accept the notion of objectivity. "... you do since you have to. The relation between grade a student holds and his actual English ability tells you much how ineffective procedure we are made to follow ... and why I don't trust number-based judgments" (Instructor 4).

Contextual factors

The interviewed instructors also raised some contextual factors that inhibited them from interpreting their assessment assumptions into practice. For example, they raised some student-related factors as reasons mediating between their assessment assumptions and practices. The participants reported that lack of language proficiency, assessment competence, and willingness from students' part influenced their assessment related decisions and practices.

The low language level of students, which many instructors were complaining about during the interviews, was a plausible reason for the instructors to practice assessment irrespective of what they think and judge to be worthwhile. Instructor 2 described that although he is interested in using oral questioning in the classroom thinking that the technique enhances students' oral language skills, he refrained since many students do not respond "because of fear of using the language in front of their classmates". This language problem of students was also a factor in using feedback of different modes. Two instructors believed that it was good to frequently practice feedback of any kind to enhance students' learning; however, their feedback emphasized more on written comments reasoning out "students' linguistic difficulty in comprehending oral feedback" (Instructor 8).

Students' overall competence was also found to be a barrier that influenced the instructors' practical choice of assessment. The participants reported that their decision to use any classroom

assessment was influenced by students' level. For instance, Instructor 5 explained why some of the assessment methods were not appropriate and not used in her classrooms:

I don't ask the students to comment on each other's papers; it would be a waste of time ...they are not going to do it. In the same manner, you can't use self-assessments when you mostly have weak students in the class...

Others also attributed the reason why they did not engage students, particularly in feedback provisions, to students' lack of competence. Instructor 2 reasoned out his refrain from giving students opportunity to assess their peers and to overlook the use of assessments that emphasize on students' role saying, "It is hard for these poor students to give feedback." The students' heavy reliance on feedback from instructors seemed to put the instructors under pressure to provide feedback by themselves than allowing students. Unwillingness from students also forced them to be adhered to the conventional methods. "They refuse when you bring them activities which can help them develop their communicative skills; they are not volunteer." (Instructor 4)

The participants, further, identified large class size and workload as factors inhibiting them from translating their assumptions into practice. For example, instructors who did not use individual oral feedback or did so infrequently, but still believed in its effectiveness, blamed this on large class size and workload. For example, an instructor clarified that he would have provided more individual-based oral feedback if such constraints were not inhibiting him. In this regard, Instructor 7 explained:

Well, I am positive if my students actively take part in the assessment issues. Student active participation in his learning has ground in the current theories.... But an important question here is if our institutions are really ready for this type of practice. Just to see a couple of examples, can the large number of students in class- the common challenge in the universities, and the load allow me to do?

Shortly, instructors of this study faced institutional and contextual related challenges although they acknowledge, sometimes through their assumptions, the importance of utilizing the innovative assessments. As data from the interviews indicate, instructors have limited autonomy and may not have control over assessment practices rather than adhering to the existing system. Assessment policies and the institutionalized organizational culture of assessment (what is

viewed as typical in their context) were inhibiting factors influencing EFL instructors from interpreting alternative assumptions they hold about assessment into classroom practices. In addition, students' language proficiency and overall competence, large class size, and workload were raised as challenges to implement their underlying assessment assumptions.

4.4. Discussions of the findings

This section of the study presents and reflects on findings of the study that sought to investigate assessment assumptions, practices and challenges in the enactment of communicative English language skills course. Accordingly, first, discussions of the findings in relation to assessment practices will be presented, and this is followed by discussions on the assessment assumptions.

4.4.1. Assessment practices

4.4.1.1. Purposes of assessment

The study revealed that assessments are employed for different purposes in communicative English language skills course. EFL instructors simultaneously employed assessment of learning, for learning, and as learning purposes in the classrooms although the extent each purpose practiced was different. The qualitative findings showed instructors' dominant use of assessment of learning followed by assessment for learning and the least used assessment as learning although the quantitative results indicated their frequent use of the three purposes. Yet, instructors' practices of assessment for the first two purposes were not prioritized from the students' perspective, substantiating the findings from instructors' interview data. Shortly, the study indicated a lack of balance in using assessment results to improve instruction and learning on one side and to measure students' learning achievement on the other side.

Measuring students' learning to make judgments about their achievement and to confirm whether or not they have met curriculum outcomes is a must in academic study regardless of assessment format or approach. However, highly emphasizing on this purpose clearly runs counter to recent developments in the assessment culture supported by experts and the new national curriculum (Black & Wiliam, 2009; Hill & McNamara, 2012; Fulcher, 2014; MoSHE, 2019). Focusing only on this purpose could be viewed as questionable in the classroom context, as it suggests that an authoritative nature of interaction and limited self-learning opportunity for students existed in the EFL classrooms, overall, minimizing students' role in the co-construction of knowledge.

Particularly, viewing the finding from the angle of the nature of the course the current study focused on makes the results disappointing. Compared to other content courses, the role of students magnifies in communicative English skills course, as much is expected from them than from instructors throughout the course delivery. However, assessment of learning, which aligns with summative assessment, emphasizes on accountability purposes primarily based on end of lesson, chapter or term exams. This means that students have limited opportunity to engage in different language activities, which are, however, considered critical in developing their EL skills. Module for communicative English language skills was designed to develop students' EL by engaging them in different learning activities (MoSHE, 2019b). Hence, overemphasizing the practice of assessment of learning in the target course would mean countering the goal of the course.

These findings are largely consistent with the available research findings. Previous studies reported that university EFL instructors used assessments for making judgments about learners' achievement more than any other purposes (Saefurrohman & Balinas, 2016; Kitaw, 2017; Dashti, 2019). Yet, findings of this study contradict with results from Shohamy et al. (2008) and Cheng et al. (2004) studies, where assessments are predominantly used by university instructors for instructional and student-centered purposes rather than for accountability. This difference in findings can be caused by the country context where the studies are conducted. Compared to Ethiopia, ESL/EFL instructors in USA and Canada, Hong Kong and Beijing contexts may autonomously practice and decide on assessments. As also presented in the above section, in Ethiopia, instructors have limited autonomy and may not have control over assessment practices rather than adhering to the established system that emphasizes more on documentation purpose of assessment. Contextual constraints in terms of resources and workload, that significantly determine instructors' practice of assessment for intended purpose, are also more of a concern in Ethiopia than in those countries.

As viewed persistently in the literature, central in classroom assessment is improving instruction and learning (Black & Wiliam, 1998; Alderson, 2005; Al-Bahlani & Ecke, 2023). However, aligning with Sara (2019) and Zelalem & Emily (2018), result of this study suggested EFL instructors' constrained approach towards using assessment to drive instructional decisions in communicative English skills classrooms. For example, Zelalem and Emily (2018) reported that

university EFL instructors did not fully utilize assessment for learning purpose in writing skills. Limited emphasis on the assessment for learning implies instructors' and students' limited use of assessments as investigative tool to find out and fill gaps in learning, a process that tremendously enhances students' English language learning.

The 21st century learning skills such as critical thinking, problem solving, reflective learning and creativity, which education policy of several countries including Ethiopia stress to foster, are effectively developed through the use of assessment as learning purpose (Earl, 2003; Stiggins & Chappuis, 2004; Price, et al., 2014). Unfortunately, this study found that assessment as learning was the least prioritized purpose of assessment in communicative English skills course. This suggests limited use of assessments to help students independently develop their English skills being critical assessors of themselves, or generally to take ownership of their learning by promoting their metacognition. For instance, students' creativity in learning is more promoted if they are provided the opportunity for self-criticism and self-evaluation than by others. So, the existing assessment practices put the strive to prepare English language learners for lifelong learning skills under question, as this is best facilitated when assessment as learning is primary and assessment for evaluation is of secondary importance in the classrooms (Earl, 2003).

Generally, EFL instructors' such polarized use of assessments for accountability purpose while limiting the use of assessment to support students' learning also contradicted what is stated in the recent policy documents. They stipulate that assessment should be carried out in the classroom for a variety of purposes and suggest balanced practices of diagnostic, formative and summative assessments both to support students' learning and to measure achievement (MoSHE, 2019; EETA, 2023). Since each of them have distinct functions to serve, employing any purpose of assessment and their hand-in-hand practices are essential.

4.4.1.2. Assessment Methods

Regarding the methods of assessment used to assess EL learners in Ethiopian universities, two major forms of assessments were witnessed including some level of alternative assessment and primarily traditional assessment. The study indicated EFL instructors' greater emphasis on particular assessment methods. Analysis of both the qualitative and quantitative data revealed that paper and pencil-based tests, mid, and final exams were the most common and

predominantly used methods within communicative English skills classrooms. Written assignments, observation, and informal oral questions were also moderately used. Even, there were hardly any mentions by the instructors during interviews in terms of using portfolio and debate as assessing methods. However, the quantitative results from both instructors and students showed that both were rarely utilized similar to roleplay, interview, self and peer-assessments.

This suggests the still overwhelming dominance of traditional assessment methods in current educational practices in Ethiopian universities. However, the recent policy documents underline that a variety of assessment methods should be used so that assessment is well aligned with instruction and optimum condition for learning is provided (MoE, 2018; MoSHE, 2019). According to Nassab (2015), alternative assessment makes a link between assessment and instruction by taking account of a feedback loop which allows instructors to monitor and modify instruction continuously based on what is already assessed. So, using alternative assessment methods that emphasize on written, practical, and oral approaches is highly recommended in the current classrooms to assess students' EL communication skills. Based on the findings, this study suggests that EFL instructors' methods of assessment were incongruent with the currently recommended practices.

Weighty emphasis on particular modes of assessment also suggests that a limited range of learning was actually being assessed. Hence, the instructors' strong emphasis on paper-based assessments may indicate that assessment of students' communicative English skills has enormously focused on assessing factual and procedural knowledge at the expense of the other domains. In this regard, the assessment practices are encouraging surface than deep approach to learning (Brown, 2009; Sheehan & Munro, 2017). By highly emphasizing on the paper and pen assessment methods, instructors are also less likely preparing EL learners for real-life situations since such methods are unsuitable to evaluate some of the skills such as listening and speaking accurately. This is also since preparing EL learners for real-life situations is best achieved through the use of authentic assessment methods (Struyven et al., 2006; Katz & Gottlieb, 2013). In this regard, the new Ethiopian education road-map, commenting to detach from paper-based-only assessments, also suggests the use of authentic assessment methods on the basis that the assessment of the learners' performances must provide an opportunity to verify that they can apply it to the real-world setting (MoE, 2018).

Although not with the policy and curriculum intentions, the findings are not surprising and are consistent with previous studies conducted in Ethiopia tertiary EFL context. For instance, Dagne (2009), Nibret (2013) and Kitaw (2017) reported dominant use of the traditional methods in EFL classrooms. Similarly, the findings are in line with studies conducted in different parts of the world. Sara's (2019) study indicated how deeply embedded traditional assessment methods are in EFL classrooms of six colleges in Oman. Along with this, Elshawwa, et al. (2017) study, which focused on separately investigating each language skills, found that majority of the Malaysian university English language instructors assessed students' language ability by relying on paper and pencil assessments. Moreover, a study in Canada and China universities reported that more than half of instructors used traditional methods in assessing ESL students' reading skills (Cheng et al., 2008).

Heavy emphasis on traditional assessment methods could be due to their easiness in execution and marking, i.e., less demanding than alternative assessment methods (Elshawwa, et al., 2017; Sara, 2019). Their efficiency in measuring student achievement and in preventing plagiarism was also reason of the instructors. Thus, the instructors preferred to use the methods in assessing their students' communicative English skills, as these were what worked for them. This suggests that the essential role alternative assessment methods play in promoting students learning is being compromised and undermined while looking for easy way of assessing and technical issues, perhaps, which relates the findings to the gap in instructors' assessment knowledge.

4.4.1.3. Provisions of assessment feedback

Regarding the sources of feedback, the findings revealed that both EFL instructors and students provided assessment feedbacks although the provisions were dominated by instructors. Even though the instructors indicated through the questionnaire that both groups almost equally took part in assessment feedback provisions, data from other sources and methods did not proof. For example, instructor-led feedback was found to be the most recurring in the interview data of the instructors, where limited use of students as a source of feedback was also depicted. Corresponding to the instructors' qualitative findings, data from students also indicated over-dominance of instructor-led feedback.

This suggests that EFL instructors prioritized provisions rather than engaging students in assessment feedback. EFL instructors' restricted use of students as a source of feedback also suggests that existing feedback practices promote knowledge transmission than creation, students being passive consumers of feedback. That is, the instructors' feedback strategies do not apprentice students into becoming active participants in their learning. Similar findings were also reported by several previous studies (e.g., Tasdemir & Yalcin, 2018; Brooks et al., 2019). In spite of the fact, scholars caution the danger of dominantly relying on instructors as a source feedback and underscore that feedback is only helpful and support students' learning when students attend to and act upon it (Black & Wiliam, 1998; Hattie & Timperley, 2007; Fulcher, 2010).

It was also indicated that both written and oral assessment feedback modes were used by instructors. This finding is confirmatory of other studies on feedback (Cheng et al., 2004; Cheng & Wang, 2007; Saefurrohman & Balinas, 2016). Although both modes were frequently used, data from instructors' interview further put insights into how situations can influence them to dominantly use one type over the other. That is, the dominant use of either written or oral feedback was dependable on particular situations and varies across individual instructors. For example, for some instructors, justifying students' linguistic difficulty in comprehending oral feedback, their feedback emphasized more on written mode. For others, oral feedback was more emphasized following students' request after just returning the assessment papers or when the instructors felt that students needed more elaborations. In addition, the findings indicated that the dominance of any type was dependent on the type of assessment method used. While written feedback was the overriding mode in paper-based assessments, oral feedback was primarily used mode in non-paper methods such as presentations and group activities. This suggests that assessment feedback practices in communicative English skills course is not as such straight forward, suggesting context dependability and complexity of instructors' mode of provisions.

Regarding the focus of feedback, the findings indicated the overwhelming dominance of verification feedback that focuses on providing binary information describing whether or not answers are correct. This was followed by the provision of directive feedback, which refers to direct written or oral provisions of correct answers to students. Restricted use of scaffolding feedback was also indicated. That is, feedback that focused on providing supports in the form of

hints, clues, questions, or reminders to students to help them complete or solve a given task independently and that drive them toward self-directed learning has received little attention from EFL instructors. This suggest that the instructors feedback directed more on telling and providing the right and wrong knowledge rather than guiding students to express their own understanding and encouraging them to come up with new insights.

These findings seem to be consistent with previous studies (Shute, 2008; Brooks et al., 2019; Dashti, 2019; Guo, 2020). While Shute (2008) and Dashti (2019) findings revealed more focus of instructors' feedback on justifying marks and grades and their limited practices of self-regulatory level feedback to build deeper and conceptual understanding, Guo (2020), from students' perspectives, reported that instructors frequently practiced directive feedback. The findings of this study also contradict with Gan et al. (2021) study that documented frequent use of scaffolding feedback among instructors of two universities. The fact that most instructors' feedback focused on describing whether or not answers are correct and explicitly correcting the students' errors than on scaffolding suggests that the instructors were mainly concerned with students' mastery of the task and justifying their learning rather than to help them carryout tasks by themselves (Shute, 2008; Gan et al., 2021).

4.4.2. Assessment assumptions

This study, further, aimed to explore participants' assumptions underlying current assessment practices in communicative English skills course, and also examined if difference in assumption exists between instructors and students and among different instructor groups. This section is devoted to discussing these two issues.

4.4.2.1. The participants' assumptions about assessment

Regarding students, taken together, the findings from both the qualitative and quantitative data consistently revealed their possession of assumptions which are deep-rooted in the traditional assessment culture. As to the instructors, the qualitative findings similarly revealed they primarily held assumptions which are incompatible with current learning and assessment theories. However, the quantitative results indicated their slight inclination to alternative assessment assumptions. Yet, their traditional assumption score that is found to be significantly higher than the hypothesized mean score of 3.0 signifies that instructors still have a traditional

assessment orientation. The fact that they do not reject or agree with traditional assumptions does not necessarily imply that they ignore them and possess alternative assumption. The quantitative result, thus, needs to be interpreted with caution, as it could mean that EFL instructors were either undecided in their assumptions, or possessed a combination of both traditional and alternative assumptions about assessment, coexisting in their belief systems.

Examining each theme separately, findings from both the qualitative and quantitative analysis indicated instructors' possession of more traditionalist assumptions regarding power and control and alternativist assumptions as to the content assessed theme. However, their assumptions concerning values and judgments in assessment were undetermined because of divergence between the qualitative and quantitative results.

Although the reasons for the divergence between the instructors' qualitative findings and the quantitative results are not clear, it could be related to the nature of questionnaire-based data. The very weakness of questionnaire is that it invites respondents to provide responses they perceive as socially desirable rather than their true thoughts, or to agree with positive statements regardless of their true beliefs (Dornyei, 2007; Cresswell, 2014). Perhaps, the long-time the questionnaire had been with the instructors, the decision made based on suggestion of the results of the pilot study, might have created opportunity for them to critically examine each item and respond in the way that others anticipate from them rather than their true assumptions. In addition, since interviews were used as the main instruments of data collection in this study, for the different reasons discussed under the third chapter, findings from the qualitative data would become more plausible than those from the quantitative.

Accordingly, overall, the findings suggest that basically traditionalist assessment assumptions underly instructors' and students' assessment practices and views in communicative English skills course. This is consistent with the available findings which revealed dominance of assessment assumptions incompatible with current theoretical approaches (Delandshere, 2001; Klenowski & Willis, 2011; Mitana, 2018). It also confirms general understanding in the literature that shows much language assessing still depends on assumptions embedded in behaviorist and measurement theories (Shepard, 2000; Teasdale & Leung, 2000; Serafini, 2002; Fulcher, 2014; Wellberg & Evans, 2022).

However, examining each theme separately gave additional insights regarding the participants' assessment assumptions. In terms of content assessed, the study indicated that students value assessments that focus on common core contents that promote correct reproduction of facts and knowledge. What is widely reflected among the students suggests the assumption that there are essential and fixed contents that students need to accumulate, and by focusing on these contents, assessment monitors their effective transmission, which also implies existence of knowledge outside the learner (Letina, 2015).

Contrastingly, through their focus on contents that consider sociocultural contexts and invite students to develop their own understanding by integrating and transforming existing knowledge, most instructors expressed the implicit assumption that knowledge is constructed, situated in context and holistic. The quantitative results similarly showed their inclination to assumptions underlying alternative assessment than the traditional, indicating their possession of assumptions productive in the current classroom assessment context regarding content assessed. Instructors in Latif and Wasim (2022) study also felt that assessment should emphasize application of learned skills and language production instead of language memorization. Such findings, that indicate instructors' assumption compatibility with current perspectives is interesting, especially considering Delandshere's (2001) and Mitana (2018) study, in which many assessment contents were predicated on the assumptions of universal and essential knowledge.

Regarding valid approach to assessing and judgement themes, although the quantitative results do not conform, the qualitative findings revealed that instructors primarily possess assumptions underlying traditional assessment. The instructors' and students' willingness to accept the notion of value-free assessments, where learning is expected to be measured objectively and consistently, underlines that their thinking about approach to assessing and interpretation of evidence of learning are hampered by misconceptions and traditions which are incompatible with current theories. Validity of such views raises question in the classroom context where assessments are mainly used to facilitate teaching and learning (Buhagiar, 2007; Carless, 2015). Since technical aspects such as neutrality and accuracy in pursuit of the intended reliability and objectivity are more emphasized, the facilitation role of assessment can be reduced, as 1) students are guarded from assessment issues and 2) one-size-fits-all approach is valued (Serafini, 2002; Buhagiar, 2007; Tan, 2012).

Furthermore, findings from both EFL instructors' and students' qualitative and quantitative data revealed their possessions of traditionalist assessment assumptions in terms of power in assessment. Both participant groups' justifications to different assessment decisions and practices reflected the assumption that much of the power should remain with the course instructor. Particularly, through their focus on providing students with rather than engaging them in assessment marking, scoring, feedback, and so on, the instructors expressed their demand to have powerful authority over students. This may suggest that assessment is a strong means of exercising power against students. This was also implied, though unnoticed in their accounts, in their over utilization of paper-based methods and selective item questions that restrict students from expressing their fuller views and preferences. This hidden instructor's privilege over students indicates that such approaches are taken for granted and are strongly embedded in their established practice-based assumptions.

Surprisingly, students were also positive for instructors to take on the dominant power. They accepted and expected unequal power relationship between them and their instructors in assessment. Tan (2012), with regard to this positive feeling of students for instructors to take on the dominant power, argues that it is naïve to assume that students want to be empowered and ready to exercise autonomy in assessment. Similarly, this study students regarded instructors responsible for every assessment activity including feedback provision and peer-assessment, justifying that they are knowledgeable. This withholds the view that the authority of instructor is considered superior to students' peers. This can also be explained by the concept of power distance in which higher status and more experienced individuals (instructors) are accorded more power and authority by those of lower status (students). In such high distance power settings, assessments are often seen as authoritative, with instructors exercising significant control and power throughout the assessment process, ultimately limiting students' voices and inputs.

This hierarchical power relationship in assessment between instructors and students also seemed to be influenced by policy documents. This is because, as also indicated by the participants, existing assessment-related documents often focused on what instructors have to do rather than highlighting the role and responsibility for students throughout the assessment processes. This, in turn, led instructors to accept their dominant power in assessment as a reasonable practice. Studies on general education teacher-student power relationship also reported teachers' power

control in assessment that resulted in unbalanced power relationship (e.g., Tan, 2012; Lau, 2013).

In spite of the fact, empowering students in assessment has recently been given prominence as a key venue to increase their independent learning and instructional decision-making power that potentially contributes to their English language learning (Price, et al., 2014; Wellberg & Evans, 2022). This finding of the current study, therefore, is disappointing especially given the existing assumptions adverse effect, for example, in limiting students' self-directed learning opportunity of the language by restricting their role in assessment process and creating powerful instructor authority. Generally, the study suggests that the participants' assumptions on instructor-student power relationships in assessment are often hampered by misconceptions and traditions that are incompatible with current theoretical approaches.

Overall, findings of the study strongly suggest that, unlike students, instructors are inconsistent in expressing their assumptions. While they held alternativist assessment assumptions on some factor, their assumptions on the other factors were in line with traditionalist assessment approach, suggesting that instructors may develop their assessment assumptions at different times between the different factors of assumptions.

This instructors' incoherence in expressing their assumptions sometimes results in semantically self-conflicting epistemic assumptions. For example, several instructors thought that assessment should focus on knowledge construction rather than reproduction. However, many of them were still unwilling to empower students in the assessment process upon which they develop their own understanding and build knowledge. Previous studies on assessment beliefs, conceptions and theories also revealed similar findings (e.g., Barnes et al, 2014; Latif & Wasim, 2022). For instance, in Latif & Wasim (2022), EFL instructors held divergent beliefs that validity and reliability and, on the other hand, authenticity and the assessment of language production comprise the main features of good assessment. Literature also indicates that instructors can simultaneously hold divergent and manifold assumptions of assessment (Anderson, 1998; Serafini, 2001; Remesal, 2011).

4.4.3. Assessment practice and assumption comparisons according to instructors' demographic variables

The study also examined whether there is statistically significant difference in EFL instructors' assessment assumptions and practices according to academic qualification and teaching experience. With regard to the academic qualification variable, results of the study showed statistically significant difference in alternative assumptions between instructors with third and second degree. In this regard, PhD degree holder instructors were found to have a higher alternative assumption than instructors with master's degree. This suggests the instructors' inclination to a more alternative assessment assumption as their academic level increases. Ph.D. degree instructors' additional post-graduate training is suspected to help them acquire a more contemporary understanding and expanded view of assessment in foreign language. This is because, the more instructors learn, the more they are exposed to new developments and theories, resulting in more developing contemporary views on assessment (Oo et al., 2023). So, they would show a higher alternative assumption than master's degree instructors.

However, this Ph.D. instructors' possession of contemporary understanding in foreign language assessment could not help them practice the innovative assessments more than instructors with master's degree. This is because statistically significant differences were not observed between the two groups in terms of their assessment practices (purposes, methods and feedback). Absence of significant differences in assessment practices between the two groups, despite the PhD holders' higher possession of alternative assumptions than the MA holders, may suggest existence of factors inhibiting instructors from interpreting their assumptions. Xu and Brown (2016) and Scarino (2013) also indicated that instructors may not implement assessment the way they think workable because of different contextual and institutional constraints they phase at their work context. This will be discussed in detail in the next section.

Another finding of interest is teaching experience which could not affect instructors' assessment assumptions. More experienced instructors reported significantly a lower alternative assessment assumption than the less experienced instructors. The less experienced instructors' possession of a higher alternative assumption about assessment than the more experienced instructors may indicate that the instructors' long experience did not help them acquire a more advanced assumption about assessment. The findings of this study seem to be consistent with previous studies conducted on EFL instructors' beliefs or conceptions of assessment (Margetic, 2014; Zheng & Borg, 2014).

Although the reasons for the differences in this study are unclear, they may be associated with ongoing learning. Keeping the role of professional development in enhancing instructors' assessment literacy (Brown, 2004; Nibret, 2013; Oo et al., 2023) in mind, one possibility might be due to the more experienced instructors' resistance to professional development than the younger instructors (Rodriquez and Mckay, 2010; Margetic, 2014). Obviously, instructors who are resistant to professional learning will have less exposure to current philosophy underlying effective practices. As a result, they would show a less sophisticated set of assumptions about assessment than the younger instructors, and the outdated assumptions, which they have already developed, remain most valid and what works for these instructors.

4.4.4. Factors inhibiting interpreting assumptions into practices

Asking EFL instructors to describe their reasoning behind their assessment practices also revealed the connection between instructors' underlying assumptions, practices, and contextual and institutional dynamics. Sometimes, instructors' reasons for their traditional assessment practices were not in ways that validate these practices. Rather, they specify instructors' criticality in using such assessments and their strong knowledge base on current assessment approach, despite their failure to employ the innovative assessments. Thus, some of the assumptions that have roots in the alternative assessment culture that the instructors expressed were found to be ideal world assumptions, as they were not reflected in practice. This suggests the influence of external factors on EFL instructors in translating their assumptions into practices.

Literature indicates that although instructors' assessment-related assumptions and beliefs play a strong role in inspiring their assessment practices, they cannot implement these assessments in a way they feel right since "they are employed within an immediate workplace community and larger social, political, and cultural contexts" (Xu & Brown, 2016: 157). In the current study, the few instructors' constrained approach towards assessment purpose, methods, and feedback, which found to be in contrast with their stated assumptions, reflects their compliance with institutional and contextual restraints. This stipulates how not only instructors' assumptions but also, even strongly, contextual and institutional dynamics dictate the practical choice of assessment in communicative English skills courses.

This further suggests that instructors' lack of using reform-oriented assessments does not always nullify the assumptions of alternative assessment on their part. In other words, what instructors actually do is not necessarily mean what they think and judge to be worthwhile, as their assumptions do not always translate into practice, which has important implications for instructor development and the practice context. It implies that simplifying about assessment practices by relying on assumptions and by overlooking the real complexity of instructors' practice contexts may result in ineffective interventions.

Latif and Wasim (2022) study also revealed that while tertiary EFL instructors held contemporary theories and beliefs about assessment, they were not reflected in their classroom practices because of contextual variables at micro and macro levels that affected their decision-making process in their assessment practices. This shows the complex nature of assessment influenced by distinct institutional, cultural and educational policy dynamics of the social context, which affect instructors' assessment assumptions, beliefs, and practices (McNamara, 2001; Remesal, 2011; Ergun, 2013; Fulmer et al., 2015; Latif and Wasim, 2022).

However, although the instructors attributed their reliance on traditional assessment practices to the established system that emphasize highly on documentation, this can be argued. For instance, they could have embedded the innovative assessments in their instruction to facilitate learning, not necessarily for grading purpose. Large class size was also an alarming reason. Although researchers mention that large class size is a problem for assessment, Carless (2015) argues that it should not be a barrier to implementing a variety of assessment methods and tries to convince its possibility using different examples. So, the minimal use of the recent assessment developments can bring into question whether instructors fully understand the role of the innovative assessments in enhancing EL learning.

CHAPTER FIVE

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

5.1. Summary

This study aimed at investigating assessment assumptions, practices and challenges in the enactment of a communicative English skills course in Ethiopian universities. Accordingly, the following findings were indicated. The study revealed that EFL instructors in the universities studied conducted assessment to measure students' achievement (of learning), inform instruction (for learning), and promote students' metacognition (as learning). However, while assessment of learning was the dominantly practiced purpose, assessment as learning was rarely utilized.

It was also found that EFL instructors emphasized on paper and pencil-based tests, mid, and final exams methods to assess students' communicative English skills as data from interviews, questionnaire, and documents indicated. Other alternative methods such as peer and self-assessments, role-play, interview, and debate were found to be uncommonly employed.

Further, the study showed that instructors frequently used both written and oral feedbacks but over-dominated the feedback provisions. Their feedbacks focused on providing: information that describes whether or not results are correct (verification), explicit correct answers to the students' incorrect/incomplete answers (directive), and clues to students to support them work out a given task independently (scaffolding). However, while verification feedback was found to be the most frequently employed type, scaffolding feedback type was used to a lesser extent.

Regarding assessment assumptions, taken together, the students held a more traditional than alternative assumptions about assessment. They felt comfortable with assessments that value homogeneous and fragmented knowledge. They also held the assumptions that assessment should measure learning objectively and consistently regardless of contextual and individual differences, and that much of the power should remain with instructors.

As to instructors, they seem to held mixed traditional and alternative assessment assumptions. They expressed assessment should promote knowledge construction, be situated in context and assess skills holistically regarding content assessed. Yet, several traditional assessment

assumptions of objectivity, consistency and hierarchical power relationships were also found to highly drive their assessment-related decisions and practices.

Instructors with Ph.D. degree held significantly a higher alternative assessment assumption than instructors with master's degree. However, statistically significant difference was not observed in terms of the assessment practices between the two instructor groups. Further, less experienced EFL instructors reported higher alternative assessment assumptions than those with more years of teaching experience, which is found to be statistically significant.

Findings from the instructors' interview further revealed the interplay between their assessment assumptions, practices, and institutional and contextual factors. Multifaceted factors in relation to policy, students' language and competence, workload, and large class-size influenced them from interpreting their assumptions into practices.

5.2. Conclusions

Based on the study, the following conclusions are drawn.

1. Students' English language learning is more supported when instructors emphasize on the use of assessment as investigative tool to find out and fill gaps in the teaching and learning rather than on the accountability purpose. However, EFL instructors of this study focused more on the accountability purpose of assessment while limiting the learning purpose of assessment. Therefore, given the findings, the prospect of bringing significant changes in students' English language ability could be greatly diminished.
2. The conventional paper and pen methods, which dominated this study EFL instructors' assessment practices, have little role in preparing EL learners for real-life situations. For example, such methods may not accurately assess students' EL abilities since they are incapable to evaluate skills such as speaking and listening. Therefore, when this finding is considered with the overall trend of EFL instructors' limited utilization of alternative assessment methods, the current movement to improve the deteriorated students' English language proficiency would be challenged.

3. Feedback practices that are framed by active learning approaches are more effective in enhancing learning. However, transmission approach to teaching is highly instilled in EFL instructors' feedback practices. Such practices diminish students' opportunity of learning English by them-selves, as this is more promoted when they actively attend to and act upon feedback. Hence, particularly in light of the current learning theories, the existing feedback practices cannot be considered productive in enhancing students' English language learning.
4. On the other hand, although EFL instructors seemed to possess mixed assumptions about assessment, most of their assumptions were found to be those underlying traditional assessment approaches. On the other hand, the study strongly suggests that students were totally traditionalists in their assessment assumptions. Therefore, it can be concluded that assessment practices in the target course primarily rest on a set of outdated assumptions that interfere with the innovative assessment practices.
5. Further, the findings of this study highlighted the impact of studying PhD on EFL instructors' assessment assumptions. Instructors with third degree had a well-developed foundational understanding and sophisticated ideas about assessment than instructors with master's degree. So, increasing academic qualification by attending PhD program can allow EFL instructors to acquire a more modernist and productive assumptions about classroom assessment.
6. Assumptions underlying current assessment approaches held by EFL instructors often did not translate into practice. The instructors faced challenges such as the established assessment system, students' language proficiency and overall competence, large class size, and workload that acted as barriers to put their assumptions into practices. This led the instructors to keep practicing assessment traditionally regardless of their philosophy and interest. Therefore, external factors are also the other reason that constrained EFL instructors from implementing the recent assessment developments.

7. However, although the instructors criticized the established assessment policy that emphasizes on documentation for their limited utilization of alternative assessment methods, it appears that there was leeway to use such methods. The policy is real and can work for summative assessments used for accountability purposes. However, it does not forbid using alternative assessment methods to facilitate instruction and learning, though not for measuring achievement. Therefore, the findings lend support to the conclusion that the instructors' assessment knowledge gap could be potential reason to rely on the traditional assessment methods, notwithstanding the other institutional and contextual constraints.
8. Overall, EFL instructors' and students' assessment assumptions and practices were largely incongruent with the policy and curriculum intentions, and broadly, with the current theoretical approaches and recommended practices. Hence, existing assessment practices and assumptions in the target course appeared to be obstacles to curriculum innovation. Particularly, the practices are not in place to equip students with the necessary English language skills needed in the current academic and workplace settings.

5.3. Recommendations

Based on the study and the conclusions drawn, the following recommendations are made.

1. Given EFL instructors' heavy reliance on the accountability purpose of assessment and limited use of assessment for instructional and learning purposes, which are, however, central in the classroom context, efforts are required to change the existing practices. Thus, universities in Ethiopia in collaboration with Ministry of Education should provide short term trainings, workshops or seminars for EFL instructors on classroom assessment purposes, particularly focusing on strengths and weaknesses of each purpose.
2. Attention should also be given to the utilizations of a variety of assessment methods, as particular assessment methods were overwhelmingly used over the others. Thus, the above stakeholders are also recommended to provide relevant trainings on the area. Particularly, promoting instructors' awareness on how alternative assessment methods are

influential and beneficial in education and how they are underpinned and interrelated with many recently developed learning theories can be more helpful.

3. Further, based on the conclusion drawn that existing assessment feedback practices are less productive in supporting students' EL learning, universities in alliance with Ministry of Education need to provide learning programs for EFL instructors. Programs that emphasize on using feedbacks to support learning are more encouraged. In addition, a focus on students' engagement in assessment feedback and its impact in EL education needs consideration.
4. EFL instructors should also try to balance assessment alternatives, even, emphasizing more on assessment *for* and *as* learning rather than assessment *of* learning, as facilitating teaching and learning is the main purpose of classroom assessment. This requires radical shift from measuring students learning outcomes at the end of an instructional program to integrating assessment in the teaching-learning process. Hence, they should also incorporate alternative assessment methods into their instruction so that different learning domains and learners with various interests are addressed, and students are well prepared for real-life situations. Further, although there is no one best way to feedback provisions, a move away from instructor-led corrective feedback appears more beneficial for EL learners. Thus, the instructors need to focus on feedback strategies that engage students and scaffold their learning.
5. Changing and shifting the participants' assumptions about assessment by paying attention particularly to students, master's degree and more experienced instructors is also the other implication of this study. With this in mind, the concerned bodies need to offer professional development programs that focus on promoting the participants' awareness on assumptions underlying current classroom assessment approaches. They may focus on those incongruent with current learning theories when designing the training opportunities. Further, it will be more likely welcomed if participants are given the opportunity to reflect on their assumptions.

6. Besides attending the formal professional development programs offered by others, EFL instructors should also engage in their own self-directed learning. Instructors make more gains in PD when they are self-selected and run along informal lines, much like a book club. PD programs established under instructors' own initiatives and that follow their own interests are more effective in encouraging them to open-ended exploration of their ideas and to the questioning of their assumptions and practices. So, in such programs, it would be more fruitful if the group, first, identify their own assumptions and air them for debate. Then, they are challenged and examined within the group in light of current ideas of assessment and learning either to accept as compatible or reject as unjustified. Engagement in such critical reflections is the most effective way for instructors, when necessary, to change their own assumptions, that overall result in positive culture of assessment.
7. The instructors' possession of a more expanded view of assessment in foreign language as their academic level increases has also important implication for universities to enhance study opportunities. Universities in Ethiopia should provide more third-degree study opportunities for their TEFL academic staff. By doing so, they will be more benefited, as this helps the staff acquire contemporary assumptions about assessment that, in turn, potentially influence their practices, resulting in enhanced learning.
8. Based on the conclusion that EFL instructors could not translate their modernist assessment assumptions into practice due to contextual and institutional dynamics, there is urgent need to tackle the barriers. Universities should work on unraveling the barriers that hinder instructors from translating their assumptions to foster the implementation of the innovative assessments in communicative English skills classrooms.
9. This study presents an important step towards investigating EFL instructors' and students' assessment assumptions and practices in the context of communicative English skills course in universities located in western Ethiopia. It provides a starting point for complementary research. Hence, future researchers should conduct confirmatory studies

with participants teaching and learning the target course in the universities located in other areas of the country.

Differences in instructors' assessment assumptions inclination observed between the qualitative and quantitative findings also suggest the need for further confirmatory studies. Thus, not only to enrich the findings and provide more insights about the issue, but also to build a sound basis for decisions and to allow firmer conclusions, upcoming researchers are suggested to replicate the study. This is also to be conclusive about psychometric properties of assumptions about assessment questionnaire.

REFERENCES

- Abiy Yigzaw. 2013. High school English Teachers' and Students' perceptions, attitudes and actual practices of continuous assessment. *Educational Research and Reviews*, 8(16): 1489-1498. DOI:10.5897/ERR2013.1573
- Adair, J. 2009. *The Art of Creative Thinking*. London: Kogan Page Limited.
- Al-Bahlani, S. M. & Ecke, P. 2023. Assessment competence and practices including digital assessment literacy of postsecondary English language instructors in Oman. *Cogent Education*, 10:2, 2239535. <https://doi.org/10.1080/2331186X.2023.2239535>
- Alderson, J. C. 2005. *Diagnosing foreign language proficiency. The interface between learning and assessment*. Continuum, London, UK.
- Ali, A. F., Khaled, N. A, and Mohd, N. 2022. Online and Offline Assessment Methods in Higher Education: A Revisitation of EFL Instructors' Perceptions and Practices. *Journal of Language Teaching and Research*, 13(6): 1147-1155. <https://doi.org/10.17507/jltr.1306.02>
- Amlaku B. Eshetie .2010. Language policies and the role of English in Ethiopia. Presented at the IATEFLBESIG, (19-21), Bielefeld, Germany.
- Anderson, R. S. 1998. Why Talk About Different Ways to Grade? The Shift from Traditional Assessment to Alternative Assessment. *New Directions for Teaching and Learning. Database*, 74 (5-16). <http://www.jcu.edu/academic/planassess/pdf/>.
- Atkin, H. 2017. Investigating service user and staff assumptions about neurological rehabilitation practice, their influence on inclusion and examining conditions for change. Ph.D. Dissertation, Northumbria University, Newcastle.
- Bachman, L. F. 2007. Language assessment: Opportunities and challenges. Paper presented at the annual meeting of the American Association of Applied Linguistics, Costa Mesa, CA.
- Bedilu Teka. 2020. An investigation into suitability of English for academic purpose course in addressing learners' needs: the case of communicative English skills at Adama science and technology university. Ph.D. Dissertation, AAU, Ethiopia.
- Berhane Gerencheal .2019. Foreign Languages in Ethiopia: History and current status. *IJRAR*, 6(1): 1431-1439. <https://files.eric.ed.gov/fulltext/ED592409.pdf>
- Biggs, J. 1995. Assessing for learning: Some Dimensions Underlying New Approaches to Educational Assessment. *Alberta Journal of Educational Research*, 41: 1-18.

- Black, P., & William, D. 1998. Assessment and classroom learning. *Assessment in Education*, 5(1): 28-43. <http://dx.doi.org/10.1080/0969595980050102>
- Black, P., & Wiliam, D. 2009. Developing a Theory of Formative Assessment. In J. Gardner (Ed.), *Assessment and Learning* (pp. 81-100). London: SAGE Publication.
- Booke, J. and Willment, J. 2018. Teaching Assumptions within a University Faculty Development Program. *Transformative Dialogues: Teaching & Learning Journal*, 11(1).
- Borg, S. 2006. *Instructor cognition and language education: Research and practice*. Continuum: New York.
- Boud, D., and Falchikov, N. 2007. Assessment for the longer term, pp. 3-25. In D. Boud & N. Falchikov (Eds.), *Rethinking assessment in Higher Education: Learning for the longer term*. London and New York: Routledge.
- Braun, V. & Clarke, V. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2): 77-101. DOI:[10.1191/1478088706qp0630a](https://doi.org/10.1191/1478088706qp0630a)
- Brookfield, S. 1992. Uncovering Assumptions: The Key to Reflective Practice. *The Reflective Practitioner*, 3(4): 13-18. <https://doi.org/10.1177/104515959200300405>
- Brookfield, S. 1995. *Becoming and Critically Reflective Teacher*. San Francisco: Jossey-Bass.
- Brookfield, S. 2013. Teaching for Critical Thinking. *International Journal of Adult Vocational Education and Technology*, 4(1): 1-15. <http://dx.doi.org/10.4018/javet.2013010101>
- Brookhart, S. M. 2011. Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice*, 30(1): 3-12. <https://doi.org/10.1111/j.1745-3992.2010.00195.x>
- Brooks, C., Carroll, A., Gillies, R. M., & Hattie, J. 2019. A Matrix of Feedback for Learning. *Australian Journal of teacher Education* 44(4). <http://dx.doi.org/10.14221/ajte.2018v44n4.2>
- Brown, G. T. 2004. Teachers' conceptions of assessment: Implications for policy & professional development. *Assessment in Education: Principles, Policy & Practice*, 11(3): 301-31.
- Brown, H. D. 2003. *Language Assessment Principles and Classroom Practices*. Pearson Education, CA, USA.
- Brown, H. D., & Abeywickrama, P. 2010. *Language assessment: principles and practices*. Pearson, White Plains, NY.
- Bryman, A. 2012. *Social Research Methods (4th Ed.)*. SAGE Publications: London, England.

- Buhagiar, M. A. 2007. Classroom assessment within the alternative assessment paradigm: Revisiting the territory. *Curriculum Journal*, 18(1):39-56. <http://dx.doi.org/10.1080/09585170701292174>
- Carless, D. 2015. Exploring Learning Oriented Assessment Processes. *Higher Education*, 69: 963-976. <https://doi.org/10.1007/s10734-014-9816-z>
- Carson J. 2005. A Philosophical Analysis of Objectivist Education. Ph.D. Dissertation, The University of Texas, Austin.
- Cheng, L., Rogers, T. & Hu, H. 2004. ESL/EFL instructors' classroom practices: Purposes, methods, and procedures. *Language Testing*, 21(3): 360-389. <https://doi.org/10.1191/0265532204lt288oa>
- Cheng, L. & Wang, X. 2007. Grading, Feedback, and Reporting in ESL/EFL Classrooms. *Language assessment quarterly*, 4(1): 85-107. <https://doi.org/10.1080/15434300701348409>
- Cheng, L., Rogers, T. & Wang, X. 2008. Assessment purposes and procedures in ESL/EFL classrooms. *Assessment & Evaluation in Higher Education*, 33(1): 9-32. <https://doi.org/10.1080/02602930601122555>
- Cohen, L., Manion, L. & Morrison, K. 2007. *Research Methods in Education*. Routledge, London and New York.
- Creswell, J. 2012. *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. Pearson, London, England.
- Creswell, J. 2014. *Research design: Qualitative, quantitative, and mixed methods approaches (4th ed.)*. SAGE Publications Inc, Los Angeles, USA.
- Cumming, A. 2001. ESL/EFL instructors' practices for writing assessment: Specific or general purposes? *Language Testing*, 18(2): 207-224. <https://doi.org/10.1177/026553220101800206>
- Dagne Tiruneh. 2009. The Exploration of the Practices of Continuous Assessment by EFL Instructors: The Case of Jimma Instructors College. MA Thesis, AAU, Ethiopia.
- Dashti, S. 2019. EFL Teachers' Beliefs and Practices about Classroom Assessment: A Multiple Case Study in The Context of Kuwait. Ph.D. Dissertation, York University, Toronto, Ontario.

- Delandshere, G. 2001. Implicit theories, unexamined assumptions & status quo of educational assessment. *Assessment in Education: Principles, Policy & Practice*, 8(2): 113-133. DOI:[10.1080/09695940120062610](https://doi.org/10.1080/09695940120062610).
- Denscombe, M. 2007. *The Good Research Guide for Small-scale Social Research Projects (3rd ed.)*. Open University Press, England.
- Denscombe, M. 2008. Communities of Practice: A Research Paradigm for the Mixed Methods Approach. *Journal of mixed methods research*, 2: 270-283.
- Denzen, N. K. and Lincoln, Y.S. 2005. *The sage handbook of qualitative research*, 3rd edition. Thousand Oaks, Calif.
- Dornyei, Z. 2003. *Questionnaire in Second Language Research: Construction, Administration and Processing*. Lawrence Erlbaum: Mahwah, NJ.
- Dornyei, Z. 2007. *Research Methods in Applied Linguistics: Qualitative, Quantitative and Mixed Methodologies*. Oxford University Press, Oxford.
- Earl, L. 2003. *Assessment as Learning: Using Classroom Assessment to Maximize Student Learning*. Corwin Press, Thousand Oaks, CA.
- Elshaw, N. R., Abdullah, A. N., and Rashid, S. M. 2017. Malaysian Instructors' Assessment Beliefs in Tertiary ESL Classrooms. *International Journal of Education & Literacy Studies*, 5(2): 29-46. <http://dx.doi.org/10.7575/aiac.ijels.v.5n.2p.29>
- Elton, L. and Johnston, B. 2002. *Assessment in Universities: A Critical Review of Research*. York: Higher Education Academy.
- Ergün, R. 2013. The Significance of Assumptions Underlying School Culture in the Process of Change. *Inter. J. Educat. Res. Technol.*, 4(2): 43-48. <https://soeagra.com/ijert/ijertjune2013/7.pdf>
- Ethiopian Education and Training Authority (EETA). 2023. Standards for Higher Education Institutional Accreditation. Ministry of Education, Addis Ababa, Ethiopia.
- Ezzy, D. 2002. *Qualitative analysis: Practice and innovation*. Routledge, London, England.
- Federal Democratic Republic of Ethiopia (FDRE). (2009). Higher Education Proclamation: Proclamation No. 650/2009. Federal Negarit Gazeta, pp. 4976-5044.
- Filer, B. 2017. An Evaluation of the Approaches to Feedback in ELT. <https://api.semanticscholar.org/CorpusID:221780880>

- Flynn, E. and Schachter R. 2017. Teaching for tomorrow: An exploratory study of pre-kindergarten instructors' underlying assumptions about how children learn. *Journal of Early Childhood Instructor Education* 38(2):182-208. DOI:10.1080/10901027.2017.1280862.
- Fulcher, G. 2014. Philosophy and Language Testing. In: A. J. Kunnan (Ed), *The Companion to Language Assessment*, (First Ed.). John Wiley & Sons, Inc.
- Fulcher, G. 2010. *Practical language testing*. Hodder Education, London, UK.
- Gabbitas, B. W. 2009. Critical Thinking and Analysing Assumptions in Instructional Technology. MA theses, Brigham Young University, Provo.
- Gan, Z., An, Z. and Liu, F. 2021. Teacher feedback practices, student feedback motivation, and feedback behaviour: how are they associated with learning outcomes? *Front. Psychol.*, 12(697045). <https://doi.org/10.3389/fpsyg.2021.697045>.
- Geberew Tulu. 2014. EFL classroom assessment: Teachers' practice and teaching techniques adjustment in Ethiopia. *Educational Research and Reviews*, 9(20): 1071-1089. <http://www.academicjournals.org/ERR>
- Geremew Lemu. 1999. A study of the requirements in writing for academic purposes at Addis Ababa University: four departments in focus. Ph.D. Dissertation, AAU, Ethiopia.
- Gipps, C. 1999. Sociocultural Aspects of Assessment. *Review of Research in Education*, 24: 355-392. <https://api.semanticscholar.org/CorpusID:143770598>
- Gonzales, R. and Aliponga, J. 2012. Classroom Assessment Preferences of Japanese Language Instructors in the Philippines and English Language Instructors in Japan. *MEXTESOL Journal*, 36(1). <https://doi.org/10.2139/ssrn.2746287>
- Gray, D. E. 2004. *Doing research in the real world*. Sage publications, Thousand Oaks, California.
- Guba, G. and Lincoln, S. 2005. *Fourth Generation*. Sage Publications, London, England.
- Guo, W. 2020. Grade-level differences in teacher feedback and students' self-regulated learning. *Front. Psychol.*, 11(783). <https://doi.org/10.3389/fpsyg.2020.00783>
- Harlen, W. 2005. Instructors' summative practices and assessment for learning- tensions and synergies *Curriculum Journal*, 16(2):207-223. <https://doi.org/10.1080/09585170500136093>

- Harmer, J. 2007. *The Practice of English Language Teaching* (4th ed.). Harlow: Pearson Education Limited.
- Hattie, J. and Timperley, H. 2007. The power of feedback. *Rev. Educ. Res.* 77: 81-112. <https://doi.org/10.3102/003465430298487>.
- Hill, K. and McNamara, T. 2012. Developing a comprehensive, empirically based research framework for classroom-based assessment. *Language Testing*, 29(3), 395-420. <https://doi.org/10.1177/0265532211428317>.
- Holroyd, C. 2003. Changing assessment in higher education: policy, practice and professionalism. Ph.D. Dissertation, The University of Stirling, UK.
- Holzweiss, P. C., Bustamante, R. and Fuller, M. B. 2016. Institutional Cultures of Assessment: A Qualitative Study of Administrator Perspectives. *Journal of Assessment and Institutional Effectiveness*, 6(1): 1-27. <https://www.jstor.org/stable/10.5325/jasseinsteffe.6.1.0001>
- Hughes, A. 2003. *Testing for language instructors*. Cambridge University Press, Cambridge.
- Iezzoni, L. I. 2006. Make no assumptions: communication between persons with disabilities and clinicians. *Assistive Technology*, 18(2): 212-219.
- Irons, A. 2008. *Enhancing Learning Through Formative Assessment and Feedback*. Routledge, New York.
- Janisch, C., Liu, X. and Akrofi, A. 2007. Implementing alternative assessment: Opportunities and obstacles. *The Educational Forum*, 71(3): 221-230. <https://doi.org/10.1080/00131720709335007>
- Katz, A., and Gottlieb, M. 2013. Assessment in the classroom. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics*. Blackwell, Singapore.
- Kezar, A. 2012. Institutionalizing Student Outcomes Assessment: The Need for Better Research to Inform Practice. *Innov High Educ*, 38:189-206. DOI:10.1007/s10755-012-9237-9.
- Kitaw Yoseph. 2017. Active Learning in Teaching English Language Support Courses to First-Year Students in Some Ethiopian Universities. Ph.D. Dissertation, University of South Africa, Ethiopia.
- Klenowski, V. and Willis, J. 2011. Challenging teachers' assumptions in an era of curriculum and assessment change. *The Primary and Middle Years Educator*, 9(1):1-9. <https://eprints.qut.edu.au/43662/>

- Latif, M. and Wasim, A. 2022. Teacher beliefs, personal theories and conceptions of assessment literacy- a tertiary EFL instructors' perspective. *Language testing in Asia*, 12(11). <https://doi.org/10.1186/s40468-022-00158-5>
- Lau, A. M. 2013. Assessment for learning in higher education. Ph.D. Dissertation, The university of south Wales.
- Leung, C. 2005. Classroom teacher assessment of second language development: Construct as practice. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 869-893). Lawrence Erlbaum, Mahwah, NJ.
- Lund, A. 2008. Assessment Made Visible: Individual and Collective Practices. *Mind, Culture, and Activity*, 15: 32-51.
- Maguire, M. and Delahunt, B. 2017. Doing a Thematic Analysis: A Practical, Step-by-Step Guide for Learning and Teaching Scholars. *AISHE-J* 8(3): 3351-3364. <http://www.ojs.aishe.org/index.php/aishe-j/article/view/335>.
- Manning, A. 2013. EAP teacher Assessment Literacy. Ph.D. Dissertation, University of Leicester, USA.
- McNamara, T. 2001. Language assessment as social practice: Challenges for research. *Language Testing*, 18(4): 333-349. <https://doi.org/10.1177/026553220101800402>
- Mebea Fetene. 2008. The Challenges of Implementing Continuous Oral Assessment in EFL Classes with Reference to Debrebirhan Teachers' College. MA Thesis, AAU, Ethiopia.
- Mesfin Abera. 2016. The Students' Level of English Language Proficiency in Ensuring Quality Education with Particular Reference to Hawassa University. *Research Journal of English Language and Literature*, 4(1): 374-385.
- Miles, M. & Huberman, A. 1994. *Qualitative Data Analysis (2nd ed.)*. Sage Publications Thousand Oaks, CA, USA.
- Ministry of Education. 1994. Education and Training Policy. Addis Ababa, Ethiopia.
- _____. 2002. The Education and Training Policy and Its Implementation. Ministry of Education, Addis Ababa, Ethiopia.
- _____. 2012. Academic Policy of Ethiopian Public Higher Education Institutions. Ministry of Education, Addis Ababa, Ethiopia.
- _____. 2013. Nationally Harmonized Modular Curriculum for Undergraduate English Language and Literature. Ministry of Education, Addis Ababa, Ethiopia.

- _____. 2018. Ethiopian Education Development Roadmap (2018-30). Ministry of Education (Education Strategy Center), Addis Ababa, Ethiopia.
- Ministry of Science and Higher Education. 2019. Ethiopian Higher Education Curriculum Framework. Bishoftu, Ethiopia.
- Mitana, J. M. 2018. Philosophical Assumptions of Educational Assessment in Primary Schools in Uganda: Case of Kampala and Kabale Districts. Ph.D. Dissertation, Makerere University, Uganda.
- Morgan, D. 2007. Paradigms lost and pragmatism regained: Methodological implications of combining qualitative and quantitative methods. *Journal of Mixed Methods Research*, 1(1): 48-76.
- Moss, G. 2003. Critical self-reflective narrative of portfolio assessment in instructor preparation. *Scholar-Practitioner Quarterly*, 2(1): 45-60.
- Motuma Hirpassa & Chaka, C. 2023. Implementing Alternative Assessment in a Communicative English Skills Course at Three Ethiopian Universities. *MEXTESOL Journal*, 47(2).
- Nasab, F. G. 2015. Alternative versus traditional assessment. *Journal of Applied Linguistics and Language Research*, 2(6), 165-178.
<http://www.jallr.com/index.php/JALLR/article/download/136/pdf136>
- National Research Council. 2001. Knowing What Students Know: The Science and Design of Educational Assessment. Washington, DC: The National Academies Press.
<https://doi.org/10.17226/10019>
- Nguyen, T. and Walker, M. 2016. Sustainable assessment for lifelong learning. *Assessment and Evaluation in Higher Education*, 41(1): 97-11.
<https://doi.org/10.1080/02602938.2014.985632>
- Nibret Asrade. 2013. The Effect of Modifying EFL Teachers' Assessment on Students' Integrated Approach to Learning English. Ph.D. Dissertation, AAU, Ethiopia.
- Nkwake, A. 2013. *Working with Assumptions in International Development Program Evaluation*. Springer, New York.
- Norton, L. 2009. Assessing student learning, Pp. 132-149. In: H. Fry, S. Kitteridge and S. Marshall (eds.), *A Handbook for Teaching and Learning in Higher Education: Enhancing Academic Practice (3rd ed.)*. Routledge, London and New York.

- Onwuegbuzie, A. J. and Collins, K. M. 2007. A Typology of Mixed Methods Sampling Designs in Social Science Research. *The Qualitative Report*, 12(2): 281-316. <https://nsuworks.nova.edu/tqr/vol12/iss2/9>
- Oo, C. Z., Alonzo, D., and Davison, C. (2023). Using a needs-based professional development program to enhance pre-service teacher assessment for learning literacy. *International Journal of Instruction*, 16(1): 781-800. <https://doi.org/10.29333/iji.2023.16144a>
- Oxford Dictionary. 2015. 'Belief'. <https://en.oxforddictionaries.com/definition/belief>
- Patton, M. Q. 2002. *Qualitative Research and Evaluation Methods (3rd ed.)*. Sage Publications, CA, USA.
- P. Dawson, M. Henderson, P. Mahoney, M. Phillips, T. Ryan, D. Boud & E. Molloy. 2019. What makes for effective feedback: staff and student perspectives. *Assessment & Evaluation in Higher Education*, 44(1):25-36. <https://doi.org/10.1080/02602938.2018.1467877>
- Pellegrino, J. and Chudowsky, N. 2003. The Foundations of Assessment. *Measurement: Interdisciplinary Research and Perspectives*, 1(2): 103-148.
- Peterson, S. & Bainbridge, J. 2012. Questioning the Underlying Assumptions of Practices in Literacy Education. *Language and Literacy*, 4(2). <https://doi.org/10.20360/G2601Q>
- Phuong, H. Y. & Nguyen, T. V. S. 2019. Using alternative assessment to improve EFL learners' learning achievement: From theory to practice. Nova Science.
- Popham, W. 2009. Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice*, 48(1): 4-11. <https://doi.org/10.1080/00405840802577536>.
- Price, J.K., Light, D. & Pierson, E. 2014. Classroom assessment: a key component to support education transformation. In Huang, R., Kinshuk, & Price, J. K. (Eds.). *ICT in Education in Global Context: Emerging Trends Report*. Springer, New York.
- Pring, R. 2014. *Philosophy of Educational Research (3rd ed.)*. Bloomsbury, London, England.
- Reis, K. 2012. *Be a change Master: Twelve Coaching Strategies for Leading Professional and Personal Change*. Corwin Press, Thousand Oaks, CA.
- Remesal, A. 2011. Primary and secondary teachers' conceptions of assessment: A qualitative study. *Teaching and teacher Education*, 27(2): 472-482. <https://doi.org/10.1016/j.tate.2010.09.017>
- Rodriquez, A. & McKay, S. 2010. Professional Development for Experienced Teachers Working with Adult English Language Learners. <https://files.eric.ed.gov/fulltext/ED540598.pdf>

- Russ, R. S., Coffey, J., Hammer, D. and Hutchison, P. 2008. Making classroom assessment more accountable to scientific reasoning: a case for attending a mechanistic thinking. *Science studies and science education*, 93(5): 875-891. <https://doi.org/10.1002/sce.20320>
- Rust, C., O'Donovan, B., and Price., M. 2005. A social constructivist assessment process model: How the research literature shows us this could be best practice. *Assessment & Evaluation in Higher Education*, 30(3): 231-240. DOI:[10.1080/02602930500063819](https://doi.org/10.1080/02602930500063819)
- Saefurrohman, P. D., and Balinas, E. S. 2016. English teachers' classroom assessment practices. *Int. J. Eval. Res. Educ.* 5:82–92. <http://doi.org/10.11591/ijere.v5i1.4526>
- Saldana, J. 2009. *The Coding Manual for Qualitative Researchers*. SAGE Publications, London, England.
- Sara, M. 2019. Assessment Literacy: A Study of EFL Teachers' Assessment Knowledge, Perspectives, and Classroom Behaviors. Ph.D. Dissertation, University of Arizona, USA.
- Scarino, A. 2013. Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and in teacher learning. *Language Testing*, 30: 309-327.
- Schraw, G. J. and Olafson, L. J. 2008. Assessing Teachers' Epistemological and Ontological Worldviews, Pp. 25-45. In M. S. Khine (Ed.), *Knowing, Knowledge and Beliefs: Epistemological Studies across Diverse Cultures*. Springer, Perth.
- Serafini, F. 2001. Three Paradigms of Assessment: Measurement, Procedure, and Inquiry. *The Reading Instructor*, 54(4): 384-393. <http://www.jstor.org/stable/20204924>
- Serafini, F. 2002. Dismantling the factory model of assessment. *Reading and Writing Quarterly*, 18(1): 67-85. DOI:[10.1080/105735602753386342](https://doi.org/10.1080/105735602753386342)
- Shohamy, E., Inbar-Lourie, O. & Poehner, M. 2008. Investigating assessment perceptions and practices in the advanced foreign language classroom. Report n, 1108.
- Sheehan, S. and Munro, S. 2017. Assessment: attitudes, practices and needs. *ELT Research Papers* 17.08, British Council.
- Shepard, L.A. 2000. The role of assessment in a learning culture. *Educational researcher*, 29(7): 4-14. <https://doi.org/10.3102/0013189X029007004>
- Shute, V. J. 2008. Focus on formative feedback. *Rev. Educ. Res.* 78, 153-189. <https://doi.org/10.3102/0034654307313795>

- Siarova, H., Sternadel, D., & Mašidlauskaitė, R. 2017. Assessment practices for 21st century learning: Review of evidence, NESET II report. Publications Office of the European Union, Luxembourg. <https://doi.org/10.2766/71491>
- Stabile, C. 2014. A Culture of Teaching and Learning Excellence Starts with an Examination of Assumptions: Influences of General Semantics on Faculty Development. *A Review of General Semantics*, 71(3): 220-226. <https://www.jstor.org/stable/24761902>.
- Stiggins, R. 2002. Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, 83(10): 758-765. <https://doi.org/10.1177/003172170208301010>
- Stiggins, R., and Chappuis, S. 2004. Putting testing in to perspective: It's for learning. *Principal Leadership*, 6: 16–20.
- Struyven, K., Dochy, F., Janssens, S., Schelfhout, W. and Gielen, S. 2006. The overall effects of end-of-course assessment on student performance: A comparison between multiple choice testing, peer-assessment and portfolio assessment. *Studies in Educational Evaluation*, 32: 202-222. DOI:[10.1016/j.stueduc.2006.08.002](https://doi.org/10.1016/j.stueduc.2006.08.002)
- Tan, K. H. K. 2012. How teachers understand and use power in alternative assessment. *Education Research International*, 2012(2012). <http://dx.doi.org/10.1155/2012/382465>.
- Tasdemir, M & Yalcin, A. 2018. Feedback preferences of EFL learners with respect to their learning styles. *Cogent Education*, 5:1. <https://doi.org/10.1080/2331186X.2018.1481560>
- Teasdale, A. A., & Leung, C. C. 2000. Teacher assessment and psychometric theory: a case of paradigm crossing? *Language Testing*, 17(2), 163-184. DOI:[10.1177/026553220001700204](https://doi.org/10.1177/026553220001700204)
- Tefera Tadesse, Manathunga, C. E. & Gillies, R. M. 2018. Making sense of quality teaching and learning in higher education in Ethiopia: Unfolding existing realities for future promises. *Journal of University Teaching & Learning Practice*, 15(1).
- The Institute of International Education. 2012. Enhancing the Quality of English Language Education in Ethiopia. In collaboration with the Ministry of Education of the Government of the FDRE, the Institute of International Education and Ambo University.
- Tsushima, R. 2015. Methodological Diversity in Language Assessment Research: The Role of Mixed Methods in Classroom-Based Language Assessment Studies. *International Journals of Qualitative Methods*, 14(2): 104-121. <https://doi.org/10.1177/160940691501400202>

- Wafubwa, R. N., and Csíkos, C. 2022. Impact of formative assessment instructional approach on students' mathematics achievement and their metacognitive awareness. *International Journal of Instruction*, 15(2): 119-138. <https://doi.org/10.29333/iji.2022.1527a>.
- Wellberg, S. and Evans, C. 2022. Assumptions underlying performance assessment reforms intended to improve instructional practices: A research-based framework. *Practical Assessment, Research, & Evaluation*, 27(23). <https://scholarworks.umass.edu/pare/vol27/iss1/23>.
- World Bank. 2011. Learning for all: Investing in people's knowledge and skills to promote development. Washington, DC: International Bank for Reconstruction and Development/ World Bank.
- Xu, Y. & Brown, G. T. 2016. Teacher Assessment Literacy in Practice: A Reconceptualization. *Teaching and Teacher Education*, 58: 149-162. <https://doi.org/10.1016/j.tate.2016.05.001>
- Yitayal Damtew and Amare Ayalew. 2023. Language Assessment Literacy Level of Ethiopian EFL Instructors: A Focus on South West Ethiopian Public Universities. *Middle Eastern Journal of Research in Education and Social Sciences*, 4(2). <https://doi.org/10.47631/mejress.v4i2.615>
- Zelalem, B. & Emily, J. 2018. EFL Instructors' Beliefs and Practices of Formative Assessment in Teaching Writing. *Journal of Language Teaching and Research*, 9(1): 42-50. <http://dx.doi.org/10.17507/jltr.0901.06>.

APPENDICES

Appendix A: Questionnaire for instructors

Dear Instructors,

I am Dagim Endale, a doctorate candidate in the program of TEFL at Haramaya University, Ethiopia. Now I am conducting research on how assessment in the communicative English language skills course is practiced and the assumptions that underly the practice. Thus, a questionnaire with three parts is deigned, and you are kindly requested to genuinely respond to each of the items so that the research boosts its confidence on the collected data. All information will be kept confidential.

Thank you in advance!

Part I: Assumptions about Assessment

This questionnaire is designed to gather information on your assumptions about assessment in the target course. Please, read each statement and decide whether you agree or disagree with the statements by putting a tick (✓) mark against your choice from the given scales. The scales are:

Strongly Disagree (SD)= 1

Disagree (D)= 2

Undecided (U)= 3

Agree (A)= 4

Strongly Agree (SA)= 5

	Items	SD (1)	D (2)	U (3)	A (4)	SA (5)
1	Assessments should be differentiated across student populations and contexts.					
2	Assessment which results in verbal data is illuminating than which result in numeric data.					
3	In order to effectively assess, the language needs to be broken into its components/skills.					
4	The focus of assessment should be on learner's ability to correctly reproduce what he/she has been taught.					
5	Assessments only give a rough indication of students' English language ability.					
6	Assessment needs to focus on certain specified contents which are similar across contexts.					
7	Assessment is a joint process between instructor and students.					
8	Students should have prior knowledge of what is going to be in					

	the assessment.					
9	Students should be shared power to decide on their assessments of the course.					
10	There are acknowledged correct responses to the course questions asked.					
11	The course instructor is sole authority to decide on assessment.					
12	Instructors should control assessment for the fact that involving students in the assessment processes makes the results inherently invalid.					
13	Much of the power in assessment needs to remain with the course instructor.					
14	Assessment should be designed by integrating English language skills than focusing on isolated skills.					
15	Students' English language learning can and should be assessed objectively.					
16	Assessment should require students to apply their own understanding of concepts, theories, etc. to real world problems.					
17	Consistency in assessment measures is the most important element of assessment practices in the course.					
18	Assessment should engage students in the knowledge construction process rather than in repeating the established one.					
19	A question can have multiple justifiable correct answers.					
20	The assessments are accurate indicators of students' English language learning					

Part II: Assessment Practices

This part of the questionnaire is to gather information on how assessment is practiced in the target course. Please, read the items carefully and indicate your response to each of the items by putting a tick (✓) mark against your choice from the given scales. The scales are:

Never (N)= 1

Rarely (R)= 2

Sometimes (ST)= 3

Often (O)= 4

Always (A)= 5

Items	N (1)	R (2)	ST (3)	O (4)	A (5)
<i>I use the assessment in the communicative English language skills course to:</i>					
plan instruction for the course					
obtain information on students' course progress					
group students for the course instruction purposes in the class					
diagnose strengths and weaknesses in my own teaching of the course					
provide feedback to learners as they progress through the course					
measure extent of English language learning at the end					
rank students based on their class performance					
determine the final grades for my students					
provide information to the central administration (e.g., depart., college)					
create an environment where it is helpful for students to complete a given task on their own					
facilitate students to become independent learners					
help students in recognizing what aspects of their own work in the course need to improve					
help students to examine their own learning					
assist students to assess their own performance in class					
<i>How often do you assess your students' Communicative English skills using the following methods?</i>					
Asking oral questions (informal)					
Student self-Assessment					
Peer Assessment					
Written assignment					
Portfolio					
Project work					
Oral Presentation					
Role Play					
Interview					
Debate					
Observation					
Quizzes/tests					
Midterm exam					
Final exam					
<i>Thinking of how feedback of assessment in the course is provided, please rate how often the following occur</i>					

Written comments on students' work					
Oral feedback on students' work					
Peer-feedback: fellow students commenting on each other's work					
Self-feedback: students' own evaluation of their work					
You provide a score/grade on students' assessments					
You indicate whether a student's answers in the class are correct					
You point out where a student is right and wrong in his/her work					
You point out the places that are not completely correct or incomplete in students' homework					
You tell a student what knowledge points he/she has mastered					
You directly tell a student the correct answer when he/she gets an answer wrong in the classroom					
When a student gives an incomplete answer, you directly supplement it for him/her					
When a student makes mistakes in his/her homework, you directly correct it for him/her					
You directly tell a student how to solve difficult problems in his/her homework					
You directly tell students the correct answers to the test questions					
You help a student solve problems by offering some hints or cues					
You help students to complete difficult problems by breaking down the problem or reducing the difficulty					
To help them do their homework, you remind students of some learning methods or techniques					
To help them do their homework, you remind students of the task requirements or evaluation criteria					

Part III: Background Information

Please, circle on the options that best represents you.

- Gender: a. Male b. Female
- Level of education:
 - B.A/B.Ed.
 - M.A./M.Ed.
 - Ph.D.
- Field of specialization:
 - TEFL
 - Non-TEFL

4. Years of teaching experience:

- a. ≤ 5 years b. 6-10 years c. 11-20 years d. >20 years

5. Credit hours you are teaching this semester:

- a. <6 b. 6-12 c. 12.5-20 d. >20

6. Average number of students in the class you are teaching:

- a. <40 b. 40-60 c. 61-80 d. >80

Appendix B: Questionnaire for Students

Dear Students,

My name is Dagim Endale. I am a doctorate candidate in the program of TEFL at Haramaya University, Ethiopia. Now I am conducting research on how assessment in the communicative English language skills course is practiced and the assumptions that underly the practice. The purpose of this three-part questionnaire is, thus, to collect data on assessment practices and assumptions. There is no right or wrong answer to the questions, so you are requested to genuinely respond to each of the items to help the research boosts its confidence on the collected data. All information will be kept confidential.

Thank you in advance!

Part I: Assumptions about Assessment

This part is designed to gather information on your assumptions about assessment. Please, read each statement and decide whether you agree or disagree with the statements by putting a tick (✓) mark against your choice from the given scales. The scales are:

Strongly Disagree (SD)= 1

Disagree (D)= 2

Undecided (U)= 3

Agree (A)= 4

Strongly Agree (SA)= 5

	Items	SD (1)	D (2)	U (3)	A (4)	SA (5)
1	Assessments should be differentiated across student populations and contexts.					
2	Assessment practice that uses qualitative description of students' learning is informative than which uses numeric data.					
3	In order to effectively assess, the language needs to be broken into its components/skills.					
4	The focus of assessment should be on learner's ability to correctly reproduce what he/she has been taught.					
5	Assessments only give a rough indication of students' English language ability.					
6	Assessment needs to focus on certain specified contents which are similar across contexts.					
7	Assessment is a joint process between instructor and students.					

8	Students should have prior knowledge of what is going to be in the assessment.					
9	Students should be shared power to decide on their assessments of the course.					
10	There are ready-made acknowledged correct responses to the course questions asked.					
11	The course instructor is sole authority to decide on assessment.					
12	The course instructor should control assessment for the fact that involving students in the assessment processes makes the results inherently invalid.					
13	Much of the power in assessment needs to remain with the course instructor.					
14	Assessment should be designed by integrating English language skills than focusing on isolated skills.					
15	Students' English language learning should be assessed objectively.					
16	Assessment should require students to apply their own understanding of concepts, theories, etc. to real world problems.					
17	Consistency in assessment measures is the most important element of assessment practices in the course.					
18	Assessment should engage students in knowledge construction process rather than in repeating the established one.					
19	A question can have multiple justifiable correct answers.					
20	Assessments are accurate indicators of students' English language learning.					

Part II: Assessment Practices

This part is to collect information on how assessment in the target course is practiced. Please, read the items in each dimension carefully and indicate the extent to which they are used by putting a tick (✓) mark against your choice from the scales provided. The scales are:

Never (N)= 1

Rarely (R)= 2

Sometimes (ST)= 3

Often (O)= 4

Always (A)= 5

Items	N (1)	R (2)	ST (3)	O (4)	A (5)
<i>Thinking of why you are assessed in communicative English skills classes, please rate how often the course instructor uses assessment to:</i>					
plan instruction for the course					
obtain information on students' learning progress					
group students for the course instruction purposes in the class					
diagnose his/her own strengths and weaknesses of teaching the course					
provide feedback to learners as they progress through the course					
measure extent of English language learning at the end					
rank students based on their class performance					
determine the final grades for students					
provide information to the central administration (e.g., depart., college)					
create an environment where it is helpful for students to complete a given task on their own					
facilitate students to become independent learners					
help students recognize what aspects of their own work in the course need to improve					
help students to examine their own learning					
assist students to assess their own performance in class					
<i>How often are the following methods used by the course instructor?</i>					
Asking oral questions (informal)					
Student self-assessment					
Peer-assessment					
Written assignment					
Portfolio					
Project work					
Oral Presentation					
Role Play					
Interview					
Debate					
Observation					
Quizzes/tests					
Midterm exam					
Final exam					
<i>Thinking of how assessment feedback is provided, please rate how often the following occur</i>					

instructor provides written comments on your assessments					
instructor provides oral feedback on your assessments					
Peer-feedback: fellow students commenting on each other's work					
Self-feedback: students' own evaluation of their work					
instructor gives a score or grade on your quiz/test					
instructor indicates whether your answers in the class are correct					
instructor points out where you are right or wrong in your assessment					
instructor points out the places that are not completely correct or incomplete in your assessment					
instructor tells you what knowledge points you have mastered					
instructor directly tells you the correct answer when you get an answer wrong in the class					
When you give an incomplete answer, your instructor directly supplements it for you					
When you make mistakes in your work, instructor directly corrects it for you					
instructor directly tells you how to solve difficult problems in your homework					
instructor directly tells you the correct answers to the test questions					
instructor helps you solve problems by offering some hints/cues					
instructor helps you to complete difficult problems by breaking down the problem or reducing the difficulty					
to help you do your assignment, instructor reminds you of some learning methods or techniques					
to help you do your assignment, instructor reminds you of the task requirements or evaluation criteria					

Part III: Background Information

Please, circle on the options that best represents you.

- Sex: a. male b. female
- Age: a. 14-16 b. 17-19 c. 20-22 d. ≥ 23
- Field of study: a. natural science b. social science

Appendix C: Interview Guide for Instructors

I. Assessment practices and underlying reasons

1. Demographic Questions
 - i. Educational qualification
 - ii. Field of specialization
 - iii. Length of teaching experience
 - iv. Number of credit hours you hold for this semester
 - v. average number of students in a class
2. For what purposes do you regularly assess your students' communicative English language skills? In your opinion, why is assessment conducted, what do you gain? Please, justify.
3. How often do you assess your students in one class? How many assessments? Why?
4. What are the methods and activities you regularly use to assess your students? Can you explain why you follow such methods and activities?
5. How is the course assessment feedback provided?
 - a. Who provides the assessment feedback?
 - b. What mode/channels of provisions are used?
 - c. What functions feedback provided serve? Or the focus of the feedback?
 - d. Rationales behind the way feedback is practiced? E.g., Why students are allowed/denied to provide feedback?

II. General interview questions to elicit instructors' assumptions about assessment

1. How do you decide on what to assess and what not to assess in this course? what factors influence your decisions?
2. What do you expect from your students' work or in terms of their responses to the questions/tasks you set? As you think, what are the differences between a good answer/response and a bad one?
3. How do you describe instructor's and students' roles in relation to assessment? Should students take part in the assessment process? Why (not)?
4. How should be your relation to students with regard to assessment? please, explain.

5. Should assessment contents and measures be uniform, consistent or differentiated across students' populations and contexts? Why do you think so? Please, justify!
6. Did you experience any good assessment? Did you experience any bad assessment? please, elaborate (or what is your view about good and bad assessment?).
7. Is there anything else about classroom assessment you would like to add?

Thank you for your willingness to participate in this interview!

Appendix D: Interview Guide for Students

1. Seenaa duubaa: saala, umurii, gosa barnootaa
2. Koorsii ingliliffaa kana waliin walqabatee, faayidaan assessmenti inni guddaan maal ta'uu qaba jettee yaadda? Maaliif?
3. Asseesmenti keessatti ga'ee barataa fi ga'ee barsiisaa akkamin ibsita? Barattoonni adeemsa assessmentii keessatti hirmaachuu qabumoo hinqaban jetta? Maaliif? Akkamin?
4. Assessmenti waliin wal-qabatee, walitti dhufeenyi barataa fi barsiisaawoo akkam ta'uu qaba jettee yaadda?
5. Koorsicha keessatti qabiyyee isin qoramuu qabdani fi qoramuu hin qabne akkamiin murteessuun danda'ama? waantoota qormaataarratti dhufuu qabani fi hinqabne maaltu (haal-dureewwan akkamiitu) murteessa?
6. Yeroo hojii-darees ta'e hojii-manaa hojjetee galchutti, barataa tokkorraa maaltu eegama? Akka gaariitti ykn sirriitti hojjete kan jechisiisu maali? Garaagarummaan deebii/hojii gaarii fi deebii yaraa maali?
7. Asseesmentin waalta'aa/walfakkaataa ta'uu qabamoo, haala qabatamaa irratti hundaa'ee baratadhaa barataatti akkasumas bakkaa bakkatti adda ta'uu qaba jettee yaadda? Maaliif? Akkamitti?
8. Walumaagala, Asseesmenti sirriin/gaariin simudatee beekaa? Asseesmenti sirrii/gaarii hin taane simudatee beekaa? Mee bal'inaan ibsi. (koorsii kanaan walqabatee asseesmenti ati gaarii/sirrii dha fi gaarii/sirrii miti jettu kan akkamiiti? akkamiin ibsita?)
9. Dhumarratti, asseesmenti korsii kanaan walqabatee yaada ati hafeera jettu yoo jiraate?

Gaaffii fi deebii kanarratti hirmaachuukeef galatoomi!

Appendix E: Sample transcription of EFL instructor interview

Date 21/02/2022

Instructor 4

Sex: Male

Place: Office

Ok, first of all you're welcome to this interview session. As I already tried to inform you, I am going to ask you some questions related to assessments in communicative English language skills course. I kindly request you to answer the questions.

Q1. Shall we begin with your background information? Tell me about your education qualification, teaching experience, credit hour you hold and average number of students in your class.

R. Ok, I learnt English language, I mean teaching of English as foreign language all throughout my first degree, second degree and my PhD. It has been three years since I have conducted my PhD and now, I am teaching English at _____ university. I have been teaching English at college and university levels for the past eighteen years. Currently, because of office position I have only six credit hours for this semester. And regarding number of students, approximately there are around 45 students.

Q2. For what purposes do you regularly assess your students' communicative English language skills? In your opinion, why is assessment conducted, what do you gain? Please, justify.

R. Frankly speaking, most of the time we assess our students to measure understanding, check their achievements of the course, communicative English skills course. For instance, after teaching them some particular skills, I sometimes check whether they are able to apply that particular skill by engaging them in some activities. Assessment also helps me know the strengths and weaknesses in my instruction. For me, assessment is more of a checkpoint to make sure that everybody has acquired the skills taught. If students do well on the lesson, we push through. But if not, it informs me to go back and review a little bit, or to do something to get those skills up. But the existing system doesn't support you to do this. More than we use assessment for instructional activities, our concern is on what we report to department, college or registrar. Our emphasis is on determining students' grades, direction of the university.

There is one more additional purpose I forget to mention that I conduct assessments in my classroom. Although it doesn't occur regularly, sometimes I give them activities that they do it

alone regardless of my interference. For example, I order students to review, edit and correct each other's writing tasks that I gave them. This time I allow them total freedom to form small group based on their interest and to select peer editor. The purpose of my assessment in this case is to help them learn by themselves without my interference, develop their English skills by relying on themselves by engaging in the independent editing. Learners must develop how to learn by themselves. This is in line with the current teaching approach in which learners share the major responsibility for their learning. In this type of practice, obviously, there might be students are not perfect and their comments can have full of errors. Due to this, marks are not given.

***Q3.** How often do you assess your students in one class? How many assessments? Why?*

R. This basis on the guidelines we follow, that is centrally monitored by ministry of education. For example, if you take the number of assessments required in the past and in present time, it is different. In the past we used to conduct five continuous assessments that covers fifty percent and the other would be from final exam, but this time I am required to give one assessment, mid exam and final exam. In principle this is what is working in our university. I don't know what others are doing. However, what you have to know: there are also informal assessments that I usually carryout; so, I can say that, if both the formal and informal assessment is considered, I can say that I frequently assess my students, several times. I believe that this way of assessing is important to more develop students' language. Despite the problems and factors influencing, I try all the best I can to improve their English skills by frequently providing different activities.

***Q4.** What are the methods and activities you regularly use to assess your students? Can you explain why you follow such methods and activities?*

R. As I have tried to point out previously, because of different reasons most of the time I use the older and the commonly known assessment methods. These reasons related to the institutional policy that we depend on; it guides us. Emphasis is given for grades. It is also because of some constraints like material or resources, huge number of students in most classes, and also unwillingness of students. Most students don't like when you change your assessment methods. I don't know but they prefer and are pro of those old methods; do not want to take part in, for example, in the authentic assessment activities.

What are the old assessment methods you mentioned? Can you list them? Yes, by this, I am referring to the paper-based methods that we always used like tests and examinations to assess students' understanding of the course objectives. Also include selective item questions including true-false, choice, and matching and the others. I call them old approaches because they are no more important and usable in this time when many alternative techniques that are more effective in teaching learning exist. Especially, I don't like to assess students on multiple choice questions, true-false, matching item questions since they foster memorization; just recall of facts. And especially in this course, it has little role to play in preparing students to use the language.

Actually, these are not the only methods I use to assess. Others such as class activities, homework, assignment are also used. Even if rarely, I also use the assessment methods that exactly reflect real-life situations. If you take, for example, I assess the students through role plays at least one time. I give them activities on functional languages such as greetings, introducing, and giving advice and ask them to role-play and demonstrate in front of their classmates by organizing the class into some small groups. I, then, first allow the group itself to comment and mark their own performance. But this mark is not something I consider while grading. Then, the class is made to provide comments and feedback on the group's role-play. However, most of the students are inaccurate and are not honest in giving marks for their groups. So, side-by-side, I also teach them how to evaluate and correct oneself work. I think such methods of assessment are good in motivating students. They become more eager to learn. And that is what I want in my classrooms since the activities help students to develop self-confidence.

Q5. How is the assessment feedback provided: Who provides it? What mode/channels of provisions are used? What is/are the focus/es of feedback you provide (functions feedback provided serve)? Please, justify your answers.

R. For your first question, regarding who provides feedback, usually it's me that gives them. I think I have clarified the reason above: since most students are reluctant to in giving feedback for their peers because of may be fear of using the language. But I believe this is mostly because of knowledge gap. As you know many students who come to university in the present time lack the basic skill and knowledge. Because of this, I am forced to give the feedback rather than asking students to give comments. Nevertheless, on activities I thought are simple, yes, students also took part in feedback provisions. They have to do this before my feedback. Mine comes usually

at end whenever they have class activities. I sometimes come with anonymously written sentences and paragraphs and ask them to review, comment and correct the grammar errors and arrangement problems committed. Before doing this, however, I first list for them the criteria on which he or she evaluate his/her own work. I do this to help learners identify their own weaknesses and strengths by themselves.

I always try my best to provide both written and oral feedback. But most of the time my feedback is more of the written one. It emphasized more on written comments because of students' linguistic difficulty in comprehending oral feedback. For example, for the questions with short answers, I simply write down the correct answers on the board so that they cross-check with their answers and determine and also proof the marks provided on the paper. But still I believe that the two ways of feedback provisions are a must since some students prefer oral and other want written feedback.

Whenever I give them comments or feedback on their assessments, my first purpose is to let them know the questions they correctly answered and the questions they got wrong. You have to indicate and tell them the correct answer of the question. This help them not to repeat the same mistake in the future. Unfortunately, they can't remember the right answer you give them, and you may expect the same mistake from the students the next day. You tell them now and forget it tomorrow. For example, when I assign them to assignment that is may be submitted after a month, I always try to assist them in different ways in order that they come with quality work, and the learners have opportunity to consult with me and I give them hints; tell some techniques that may guide how they need to pursue their work. However, I hardly find satisfactory work and appropriately prepared assignment despite all these efforts. This remained a challenge for all of us.

Q6. How do you decide on what to assess and what not to assess in this course? what factors influence your decisions?

R. Oh! This is a bit tough question. I need to respond it in two different ways. What should be known is, some assessments are prepared by individual instructors but in the case of mid and final exams, uniform exams are prepared by group. Look, most of the time we prepare exams together with other instructors, and we determine together what should be included in the exams.

So, what to include and not in the assessments is not something you define individually. If you take, for example, there is committee formed from few members assigned for this semester to prepare the mid exam and final exam, and this committee governs the process including setting questions. The questions are prepared based on objectives that are already set. That is clearly listed in the module. Of course, what is commonly accustomed is selecting and preparing some questions which are used in the previous years, or sometimes you come across when the same exam used in the past is repeated without even altering the order of the questions.

However, when preparing, say it test or quiz, there are some factors like students' ability, my work load, available resources and also number of students that determine me what I have to assess. But when I say this, I am not saying that the objectives in the module we commonly use across universities is not taken into consideration. They are the blue print based on that I set questions.

Q7. What do you expect from your students in terms of their work or responses to the tasks you assign? As you think, what is the difference between a good answer/response and a bad one?

R. I think you are wondering to know what type of responses for the task I gave I want my students to come with. Look, from the beginning, I see quality or appropriate responses from the point of view of its originality. For example, I sometimes give them either individual or group assignment, and I feel inappropriate work if students come with an assignment copied from different textbooks or other materials. You can easily detect whether the students' answers are their own efforts or something directly copied from somewhere else. Surprisingly, it is common to see when several students submit similar papers, others coping from each other. Of course, I don't have tolerance for such works. It is my stand that at least they have to assimilate it to their own knowledge and write using their own understanding of the topic and own words. In addition, if their piece of work is shallow in depth and copy of what I taught them in the classroom, I also consider it poor work. Because we instructors must be careful and should have responsibility to discouraging cheating or coping; our objective should be to help our learners to depend on themselves. This is a bad culture to be corrected, and I believe that this is the root cause for the current disgusting situations observed in academics and also others.

***Q8.** How do you describe instructor and students' roles in relation to assessment? Should students take part in the assessment process? Why (not)?*

R. As I told you before, I think that assessment is not only the role of instructors. Even though majority of the work of assessment belongs to the instructor, I don't think that students have no roles and responsibilities in the assessment. Students should also be part of the assessment process. As I indicated above, directly or indirectly students have roles in my classroom assessment and feedback practices. While doing this, it basis on my understanding of the important role involving students in the teaching and learning bears. Students more benefit when they have roles in learning, when they dig out something by themselves than when the instructor make spoon feeding. So, I believe that students have to take part in the assessment and feedback, and I am doing this, for example, by giving them chance to involve in the peer-feedback, self-feedback because learners are capable to learn by themselves if they are engaged. May be you may not allow them during scoring and marking since I observed that students are not accurate enough and cannot be honest.

***Q9.** How should be your relation to students with regard to assessment? please, explain.*

R. First of all, there is something I have to point out here. Instructors and students are co-learners and co-instructors, having similar intention to achieve the same goal they both share, and as instructors have a number of things to learn from students like the students do. It is when they mutually work together that the students' learning can be succeeded. So, the more instructors and students come and work together, the better learning will come. I feel that my relationship with students should not be something repelling like in the earlier education system where a instructor alone carries out everything. Today, students are given high, important position in the education to play positive effect on the teaching learning. So, we have friendly approach in the classroom and also outside they ask me any ambiguity that they don't understand in the class. I approach them like their friends and allow them to communicate with me their problems, and they know this of my behavior, and because of this they ask me any questions they have without any fear. This, I think smooths and facilitate learning and make more learning to occur. We have good relationship in assessment too. Similarly, in the context of assessment, I think the view that since instructors are the most knowledgeable and experienced, they should control assessment issues is not right. Students should be part of decisions and implementations of assessment.

***Q10.** Should assessment contents and measures be uniform, consistent or differentiated across students' populations and contexts? Why do you think so? Please, justify!*

R. Ok, this is what we actually do in this university. I think I have tried to illustrate earlier somewhere. The common practice here is both the mid exam and final exams are centrally prepared and students of any field are required to take these exams. When I say this, uniform questions are prepared and all students taking this course sit for the same exam regardless of their differences. This is the strict principle that no one should violate. But according to my understanding, it is very important if assessments are prepared according to students' needs and learning styles. For example, for some students, as they want to learn through demonstrations, they also prefer to be assessed the same way. There are also others which prefer paper and pencil tests. So, in this case, by using the same testing method, we are favoring some students over the others. So, I personally disagree and do not accept the existing practice, although I don't have the power to change it, and I believe that students should be given alternative ways of expressing their understanding and views because it provides fair opportunity for learners of various interests, learning styles, abilities.

I also have objections with regard to the similar assessment contents students of different fields are forced to sit for. It is not right to assess students of different areas of study on the same contents. Of course, not only the assessments, if you take the modules and curriculum, they are not designed according to students' specific fields of study. However, what I personally believe is that the assessments we conduct will be more effective in developing students' English language learning when the questions designed take students' areas of study such as medicine, management, law and etc. into considerations. In other words, unless assessment tasks are designed in accordance with learners' areas of study, I mean with their field- like medicine or management or any other, assessment practices will not be fair. Assessment should be mock of what they study and will do.

Of course, I understand why we are critically required in our university to make the exams so similar and uniform across the students. It is to maintain reliability of the questions and to avoid discrimination by using the same questions for all students learning this course. But, for me, its problem outweighs since we lose many important things we get by differentiating these assessment methods and contents according to the students. In line with this, in other word, it is

believed that this help to eliminate errors, to be objective, which is also controversial idea. Personally, I don't think that it is possible to be objective in assessment. But majority of us are still struggling to achieve objectivity, and often forced to achieve this through quantifications, and you do since you have to. The relation between grade a student holds and his actual English ability tells you much how ineffective procedure we are made to follow, which led us to crisis and why I don't trust number-based judgements.

Q11. Did you experience any good assessment? Did you experience any bad assessment? please, elaborate (what is your view about good and bad assessment?).

R. Let me ask you one question. Can one hand alone clamp? Ih? I have reasons to raise this question. You see, in any language in the world one skill cannot work independent of the other skills. These language skills are interdependent, and it is when students develop, let's say listening skills, that they simultaneously get good skill of writing or speaking, and the like. The second point of my argument is that it relate to the current communicative approach of teaching and learning language. If our ultimate purpose is developing the students' communicative skills, I am sure this cannot be achieved by focusing on teaching and assessing of the skills one after the other separately. Despite these facts, commonly we often teach and evaluate each skills distinctly. I don't know where the practice emerged, but I see it malpractice that goes against the nature of the language itself that naturally works as a complete system. However, I strongly believe that language assessment must be in ways language is used in real life. It needs to work on it, to disconnect from practices that treat skills independently.

Let me come to my point. The way we are assessing our students today is quite far from the principle and science of how language is acquired since we are obsessed on treating each language skill, writing, reading, speaking, listening, grammar, vocabulary separately. For your surprise, the module, the communicative English language skills module was also designed in such manner. You also know this. You will be amused by people who prepared the modules. Any ways this is what I identify as a bad assessment culture that clings to my mind and I always wanted to be corrected that many didn't paid attention to.

Thank you for your willingness to participate in this interview!

Appendix F: Sample transcription of the interview conducted with students

Date 26/02/2022

Student 10

Sex: Female

Place: Classroom

Duraan dursa, gara gaaffii fi deebii kanaatti dhufuu keef galatoomi. Akkuman siif caqasuu yaaletti gaaffiiwwan muraasa asseesmantii koorsii communicative English language skills wajjiin walqabatan irratti sigaafachuufan deema. Akka ati naaf deebistus kabajanan sigaafadha.

Q1. Mee jalqabarratti waayee umuriikee fi gosa barnootaa ati barattuu haa dubbannu.

R. Tole, anis gaaffii fi deebii kanarratti nahirmaachisuu keessanif galatoomaa jedha. Ani umuriin koo amma wagga digdama, kanan baradhummo first year muummee naachuraal saayinsiit.

Q2. Koorsii ingliliffaa kana waliin walqabatee, faayidaan asseesmenti inni guddaan maal ta'uu qaba jettee yaadda? Maaliif?

R. Ani akkan yaadutti asseesmentiin faayidaa guddaa qaba. Koorsii kana keessati asseesmentin ammam nuti barattoonni akka hubanne adda baasuuf nufayyada. Sababni isaas barsiisaan erga barsiisee booda maal akka qabanneef maal akka irranfanne yokan daganne adda baasee beekuu qaba. Kana gochuuf immoo asseesmentin dirqama. Barataa isa kamtu akka cimaa ta'eefi isa kamtu dadhabaa akka ta'es kanumaan adda baasa. Kana qofas miti. Akka ani yaadutti gireedi keenya nuuf hojjechuufillee barsiisaan qormaata nuqoruu qaba. Namni hundi wal-qixa miti; asseesmentin immoo gireeditti jijjiiramun eenyuuf maaltu mala isa jedhuuf deebii laataa jechuudha. Asseesmentin hin jiru taanan kun hundi hin ta'u; kanaaf asseesmentin dirqama koorsi kamiyyuu keessatti jedheen yaada.

Q3. Asseesmenti keessatti ga'ee barataa fi ga'ee barsiisaa akkamin ibsita? Barattoonni adeemsa assessmentii keessatti hirmaachuu qabumoo hinqaban jetta? Maaliif? Akkamin?

R. Akka ani yaadutti, asseesmentin harka caalaa hojii barsiisaat. Barattoonni ga'ee hin qaban ituu hin taane, isaanis ga'ee mataa isaanii qabu. Garuu, asseesmentidha waan ta'eef saayinsii matasaa qaba. Kanammoo kan beeku barsiisaadha malee nu miti. Akkaataa itti qoru, akka itti qormaanni ba'u, maal akka baasuu qabu, yeroo itti qoramnu seera matasaa qaba kan barsiisaan beekuu fi kan nuti hinbeekne. Sababa kanaaf asseesmentin ga'ee barsiisaat jedheen yaada. Gaaffii baasuu, qoruu, soroorsee deebii sirrii ta'e kan laachuu qabu barsiisaadha malee barataa

ta'uu hin qabu natti fakaata. Maaliifi? ituu amma carraan sororsuu kun barataaf kenname dogoggoratu baayyata. Kanaaf sirrii miti. Ani akka barataa tokkootti qormaanni koo dogoggora tokko malee akka naaf sororamee deebii walhinfalmisiisne waliin akka naaf deebi'un barbaada. Kanaaf barsiisaa qofti akka qormaata koo sororu barbaada. Kanammoo barataan hinbarbaanne jira jedhee hin amanu. Sababnisaammoo, qabxichi akka itti amanamuuf hojiin keenya hundi haala walfakkaatun barsiisaa keenyaan sororamu fi madaalamuu qaban jedha. Barataanimmoo kan irraa eegamu ofqopheessuu qaba; sirriitti dubbisee akka gaaffii hunda deebisutti dhiyaachuu qaba. Akkas jedheen yaada ani.

***Q4.** Assessmenti waliin wal-qabatee, walitti dhufeenyi barataa fi barsiisaawoo akkam ta'uu qaba jettee yaadda?*

R. Akka ilaalcha kootti, barataa fi barsiisaan walitti dhufeenya gaarii ta'e ta'e yoo qabaatan ani rakkina qaba jedhee hin yaadu. Kuniyyuu kaninni ta'uu qabu yeroo baruu fi barsiisut malee qormaataa wajjiin wal-qabachuu hin qabu. Yoo gaaffii qabatee, waan isaa hingalle yoo jiraates kutaa keessattis ta'e alatti, eessattu akkaataa seera qabeessa ta'een itti siqee gaafachuu danda'a. Garuu waan qormaataa wajjiin walqabatee, akka ani yaadutti barsiisaan barataa kamiyyuu ofitti butuu hinqabu. Sababnisaas, kun hojii barsiisaati malee kan barataa miti. Akka nama qoruu fi akka nama qoramu tokkootti walilaaluu qabu. Qormaata waan ta'eef barsiisaan icitiidhaan dhoksaatti qabuuyyuu qaba.

Hariiroon isaani inni guddaan barataan barsiisaa isaa akka barsiisaatti ilaaluu qaba; akkasumas barsiisaan barataa isaa akka barataatti ilaaluu qaba. Barattoonni tokko tokko barsiisaa isaanii waliin walitti dhufeenya hinmalle qabaachuun qabxii isaanif hinmallelee warri argatan jiru. Kunimmoo tokko fayyaduun barattoota kan immoo miidhuu ta'a. Hariiroon akkasimmoo barattoota gidduuttis shakki uuma. Tokko itti du'ee qayyabatee dhufa, kaanimmoo utuu hin dhama'iin qabxiin kennamaaf taanaan inni dubbisu akka jabaatee hin dubbisne hamilee cabsa. Kanaaf barsiisonni barattoonni akka jala fiigan eeyyamu hin qabu. Maalifi, kun qormaata, kan hundi keenya ija tokkoon ituu walhincaalsisin irratti evaluate ta'uu qabnu dha.

***Q5.** Koorsicha keessatti qabiyyee isin qoramuu qabdanifi hin qabne akkamiin murteessuun danda'ama? Waantoota qormaataarratti dhufuu qabani fi hinqabne maaltu (haal-dureewwan akkamiitu) murteessa?*

R. I think, kana kan murteessuu danda’u barsiisaa natti fakkaata. Akkuman gubbaatti ibsuu yaaletti kan beekumsasaa qabu barsiisaa waan ta’eef waan nuti qoramuu qabnuuf hinqabne isatu beeka. Barsiisaan immoo akka beekumsa isaatti qabiyyee qormaata irratti dhufuu qaban adda baasee beeka. Kanammoo barataniiru jedheen amana.

Tole, barsisaanumtiyyuu haal-dureewwan akkamiirratti hundaa’ee qabiyyee qormaataa kana murteessa?

Haal-dureewwan barsiisaan irratti hundaa’ee qabiyyee qormaataa murteessuu, akkan yaadutti, fakkeenyaaf, waan nu barsiisee fi hinbarsiisin adda baasee kan nuti barannee darbinee ala qormaata irratti fiduu hin qabu jedheen yaada. Fakkeenyaaf, moojuliin koorsii kanaa jira miti kan barattoonni hundi keenya koorsii kana fudhachaa jirru irraa barannu. Qabiyyeewwan achi keessatti hammataman jiru waan ta’eef barsiisaan isaa ala ba’uu hin qabuu jechuudha.

Q6. *Yeroo hojii-darees ta’e hojii-manaa hojjetee galchutti, barataa tokkorraa maaltu eegama? Akka gaariitti ykn sirriitti hojjetee kan jechisiisu maali? Garaagarummaan deebii/hojii gaarii fi deebii yaraa maali?*

R. Tokkoffaa, waan gaafatame sirriitti xiinxaluutu irraa eegama. Jalqaba gaafficha sirriitti hubachuu gaafata. Erga gaaffii kennaman sirriitti hubatee booda, material sanaaf ta’u laaybrarii seenee adda baafachuu qaba. Isa booda erga sirriitti dubbiseen booda noottii qabatee barreessee galchuu qaba. Waan kutaa keessatti barsiisaan barsiisees irra deebi’ani ilaaluun akka dogongora hin umneef ni fayyada. Gaaffiwwan dhihaataniif baratan deebii isaa kan kutaatti baraterras ta’u kitaabota dubbiserra funaannachun ituu hin dogongoriin barreessee yoo galche, kun hojii gaariidha jedheen yaada ani. Walumaagala, inni guddaan ati deebii sirrii ta’e barbaadii argadhu; isa booda dogoggora ituu hin uumiin seeraan barreesun nama hindhibu.

Lammaffaan, instruction barsiisaan kennu hordofuu dha. Fedhii barsiisichaa adda baasuun akka barsiisichi barbaadetti hojjetanii argamuun gaarii ta’a. Barsiisaan tokko tokko assaayimantii kenne bal’inaan akka ati ibsiteef hojjetu barbaada. Tokko tokkommoo gabaabsitee peejjii xiqqooratti yoo xumurte itti tola. Sababa kanaaf, barataan yeroo assaayimantii hojjetee galchutti akka fedhii barsiisichaatti yoo deebii isaa qindeeffatee barreesse, hojii bareedaadha waan ta’eef barsiisaanis itti gammada; qabxii gaariis isaaf kennisiisa. Inni kan biraan immoo assaayimantii kenname yeroosaa ituu hindabarsiin, yeroo kenname keessatti xumuruu barbaachisa.

***Q7.** Asseesmentin waalta'aa/walfakkaataa ta'uu qabamoo, haala qabatamaa irratti hundaa'ee baratadhaa barataatti akkasumas bakkaa bakkatti adda ta'uu qaba jettee yaadda? Maaliif? Akkamitti?*

Kanumti koorsii kanamoo?

Eeyyee, barattoota hunda walfaana koorsii walfakkaataa fudhatan gidduutti.

R. Koorsii kana yeroo tokkotti walfaana fudhachaa jirra taanaan akkamitti asseesmantii addaa addaa qoramnaree. Jechuun, ituu amma barsiisan keenya daree keenyaafi daree kan biraa barsiisuuf qormaata addaaddaa baasa jechuun sirrii natti hinfakkaatu. Maaliifi, gosti koorsiis tokkuma barsiisaa tokkotu nubarsiisaa jiras taanaan asseesmantii gosuma tokko nuuf baasuu qaba malee, daree keenyaaf kan bira daree kaanif kan biraa baasa taanaan loogiitu uumama. Kun maaliifi, ulfaatinni fi laafinni qormaataa lammani kun walcaaluu danda'a, qixa ta'a jedhee hin yaadu. Yookan kan keenya yookanimmo kan daree biraa caalaa ulfaachuu danda'a. Karaa biraammoo, hunda keenyaaf akkaataa gireeddin itti hojjetamu haala walfakkaataadha. Kanaaf, amma beekumsakootti, daree tokko keessatti barataadha barataattis ta'u, dareedhaa dareetti koorsii tokko fudhachaa jirra waan ta'eef hundi keenyayyuu qormaata gosa tokkorratti madaalamuu qabna. Waldorgommiin gireeddii waan jiruuf qormaata walfakkaataa hundaaf dhiheessuun hundumtuu akka dandeettisaatti gireeddii isaaf malu dirree walfakkaataa irratti dorgomee argachuu qaba; kan kanaa achii feerii hin ta'u.

***Q8.** Walumaagala, Asseesmenti sirriin/gaariin simudatee beekaa? Asseesmenti sirrii/gaarii hin taane simudatee beekaa? Mee bal'inaan ibsi. (koorsii kanaan walqabatee asseesmenti ati gaarii/sirrii dha fi gaarii/sirrii miti jettu kan akkamiiti?)*

R. Akkakootti anaan kan natti tolu gaafa barsiisaan bakkeedhaa gaaffii baasuu dhiisee waan baranne keessaa qofa nuqorudha. Garummoo, barsiisonni tokko tokko waan kutaa keessatti nu hinbarsiisne fidu. Kun akka kootti asseesmantii sirrii miti. Sababni isaas akkan yaadutti asseesmantii jechuuniyyuu baratan tokko waan barate tokko ammam anderistandi godhe isa jedhu adda baafachuufidha kan barsiisaan barataa qoruu qabu.

Koorsii kana keessatti asseesmantiin ani sirriidha jedhee itti amanus jira. Fakkeenyaaf, yeroo barsiisaan qormaata nu qore sirriitti dogoggora tokko malee soroorsee ituu ofharka hintursiin

riisaltii keenya akka ilaallannuuf gaafa nuuf deebisutu jira. Kun yeroo hunda ta'uu baatullee waan akkasi barataa baayyee gammachiisa. Barsiisaan yoo akkasitti qormaata keenya nu agarsiisu, nutis hammam akka arganne waan beeknuf dhumarratti qabxii keenya waliigalaa irratti dogoggorri yoo uumamellee dafnee sirreeffachuu dandeenya jechuudha. Kanaaf barsiisaan yeroo akkas dafee dafee qormaata nuuf deebisu iftoomina uumaa deema. Kunimmoo barataa fi barsiisaa gidduutti shakkiin akka hin uumamneefillee ni fayyada. Kanaaf hojiin akkasii jajjabeeffamuu qaba.

Gaaffii fi deebii kanarratti hirmaachuukeef galatoomi!

Appendix G: Sample course outlines used

3.3 Strategies for improving English grammar knowledge 3.4 Reading: The Awramba community Unit 4: Wild Animals 4.1 Listening: Human-wildlife interaction 4.2 Reading: Africa's wild animals 4.3 Vocabulary: Denotative and connotative meanings 6. Teaching Methods ➤ Short lectures, Pair and group work, Discussions, Presentations. ➤ Independent learning (e.g. independent reading) ➤ Self-Study and portfolio compilation on: • Present tenses, perfect tenses, conditionals, active and passive constructions. 7. Assessment • Active Participation = 5% • Individual Assignment (portfolio) = 10% • Group Assignment = 15% References Alfassi, M. 2004. Reading to learn: Effects of combined strategy instruction on high school students. <i>Journal of Educational Research</i>, 97(4):171-184. Anderson, N. 1999. <i>Exploring second language reading: Issues and strategies</i>. Toronto: Heinle & Heinle Publisher. Bade, M. 2008. Grammar and good language learners. In C. Griffiths (Eds.), <i>Lessons from good language learners</i> (pp. 174-184). Cambridge University Press. https://doi.org/10.1017/CBO9780511497667.016 Bouchard, M. 2005. <i>Reading comprehension strategies for English language learners: 30 research-based reading strategies that help students read, understand and really learn content from their textbooks and other nonfiction materials</i>. New York: Scholastic. Cameron, L. 2001. <i>Teaching languages to young learners</i>. Cambridge: Cambridge University Press. Chamot, A. G. 1987. The learning strategies of ESL students. In A. Wenden & J. Rubin (Eds.), <i>Learner strategies in language learning</i> (PP 71-85). Prentice-Hall, Hemel Hempstead. Gairns, R. & Redman, S. 1986. <i>Working with words: A guide to teaching and learning vocabulary</i>.	4.4 Grammar focus: Conditionals revised Unit 5: Population 5.1 Listening: Population density 5.2 Reading: Population pyramid 5.3 Vocabulary: Collocation 5.4 Grammar focus: Voice
--	--

4.4 Grammar focus: Conditionals revised Unit 5: Population 5.1 Listening: Population density 5. Teaching Methods ➤ Short lectures, Pair and group work, Discussions, Presentations, Independent learning (e.g. independent reading) ➤ Self-Study and portfolio compilation on: • Present tenses, perfect tenses, conditionals, active and passive constructions. 6. Assessment • Individual Assignment(portfolio)= 10% • Test/Group Assignment = 10% • Mid exam= 30% • Final Examination = 50% 7. References	5.2 Reading: Population pyramid 5.3 Vocabulary: Collocation 5.4 Grammar focus: Voice
---	---

Appendix H: Sample assignment corrected

short paragraph, at first day when I join university

My first day at university ~~was~~ ^{SS} a day, I ~~will~~ ^{SS} never forget. I ~~was~~ very nervous, because I ~~didn't~~ know any one. However my first day at university ~~was~~ memorable.

At first I ~~was~~ ^{asked} confused & lost. I ~~didn't~~ know where to go. I ~~had~~ ^{asked} a stranger, where the first year students ~~were going~~ to meet. This person ~~was~~ very polite & ~~showed~~ me the way. I ~~walked~~ through some big buildings & when I ~~arrived~~, I ~~saw~~ many students & I felt scared. I ~~was~~ beginning to feel that, I ~~had~~ become a university student.

Finally, we ~~began~~ our classes. However, all our classmates ~~were~~ quiet. The teacher talked for a while & then we ~~introduced~~ ourselves. I think it ~~is~~ natural that on our first day at university we ~~feel~~ nervous. That's why I ~~say~~ that my first day at the university ~~was~~ very exciting & full of memories.

The types of tenses exist in this paragraph listed as follow:-

<u>present simple</u> • we feel • I say • is natural	<u>past simple</u> • I was ... , I saw ... , were • Univ. was ... , I didn't ... , scared • I arrived ... , I introduced • talked ... , began • showed, walked <u>simple future</u> • I will	<u>present perfect</u> • had become
<u>past continuous</u> • were going • was beginning		<u>past perfect</u> • I had ask • I had become

Appendix I: Sample test/quiz corrected

Read that provide an appropriate relative
Pronoun (who, which, whose, whom)

I had a wonderful time in your city, and I
specially liked your family who treated me
like their own son. I made friends with which I
went out every weekend. During the weekdays I
had to study, and the librarian who office was near
to our residence allowed me to borrow ^{possession} as many books
as I wanted. ^{I had good} But it ~~is~~ not time learning about the
city and its history. And I was not the only strange
face in this school. I became friends with an American
guy whom ^{verb} came last year. The party in which I met
this guy was a traditional affair, and it was inter-
esting ~~to~~ to join people in singing and dancing
around the bonfire. I really enjoyed my time here
and I always cherish the time ~~where~~ which I spent
here with you.

5/10

Appendix J: Sample major exam corrected

the number of dynamic and creative people who can contribute to the development of their country. Likewise, with more entrepreneurs taking their investments abroad, developing countries are missing an opportunity of wealth creation. This has also negative consequences on tax revenue and employment.

Most of the measures taken so far have not had any success in **alleviating** the problem of brain drain. A more global view must take into consideration the provision of adequate working and living conditions in the sending countries. Another option should involve encouraging the expatriates to contribute their skill to the development of their countries without necessarily physically relocating.

I. Based on the above passage, choose and put your answer on the spaces.

- What is the attitude of the writer towards the topic of discussion? He/she is:
 - Against brain drain
 - Supporting brain drain
 - Neutral
- What could be the most possible title of the passage?
 - Migration
 - Advantages of brain drain
 - Educated People
 - Brain Drain
- paragraph 1, line 3, "heading" means?
 - leaving
 - coming
 - eating
 - no answer
- Paragraph 4, line 1 "alleviating" means?
 - decreasing
 - increasing
 - bringing
 - supplying

Answer: 1. ☒ A 2. ☒ C 3. ☒ A 4. ☒ A

II. Give short answer depending on the passage you read.

- "these" paragraph 2, line 7, refers to: push factor and pull factor
- What is the main idea of paragraph three? Brain drain
- Write at least three causes of brain drain.
 - Social unrest
 - Political conflict
 - War

18

III. Grammar section: Choose and circle on your best answer

- If you _____ the exam, I would have bought you a new shoe.
 - passed
 - had passed
 - pass
 - passing
- If she worked harder, she _____ better marks.
 - will get
 - would gotten
 - would get
 - b and c
- We _____ the lecture if the taxi had delayed.
 - Will miss
 - Will missed
 - Would miss
 - Would have missed
- If we _____ money this year, we will not be able to buy the new book released.
 - did not save
 - do not save
 - have not saved
 - had not save

Answer: 8. ☒ B 9. ☒ C 10. ☒ C 11. ☒ D

Complete the sentences with the correct form of **present perfect** or **simple past** verbs given in bracket

- The carpenter was builded (build) the house yesterday.
- Fenet is eats (eat) her lunch. She looks satisfied.
- The students in my classroom have beginning (just, begin) the study for communicative English course, but I had finished (finish) the study three days ago.

IV. Speaking section: choose and circle the correct answer from the alternatives given

15. Buli: Hi, my name is Buli.

Chaltu: _____

- I know you.
- Why do you tell me?
- Hi, I'm chaltu.
- Didn't I tell you that I am chaltu?

16. Dr. Gemechu: Mr. Belay, let me introduce sister Sichale. She works in Nekemte Hospital.

Mr. Belay: I'm pleased to meet you sister, I'm Belay.

Sichale: _____

- Nice to meet you, too.
- I'm pleased to meet you.
- I'm pleased to meet you, too.
- How do you do Mr. Belay.

17. A: _____

B: Hello, this is Maria.

- What's your name?
- It's nice to meet you Tizita.
- Hello, my name is Toran.
- How're you, Tizita?

Answer: 15. ☒ C 16. ☒ C 17. ☒ C

Appendix K: Course outlines analysis

Document Analysis: techniques used to determine assessment methods used

Methods	Course outlines								%
	1	2	3	4	5	6	7	8	
Oral questioning									0
Self-assessment									0
Peer-assessment				*			*	*	7.1
Individual/group written assignment	*	*	*	*	*	*	*	*	19
Portfolio		*							2.4
Class activity	*		*						4.8
Project work	*		*						4.8
Oral Presentation		*					*	*	7.1
Role Play									0
Interview									0
Debate									0
Quiz/test	*	*	*	*	*	*	*	*	19
Written midterm exam	*		*	*	*	*	*	*	16.7
Written final exam	*	*	*	*	*	*	*	*	19

The sign * indicates that the method in the row exists in that particular course outline in the column

NB. Although active participation and attendance were also included in some course outlines, they are excluded from this chart since they are not considered as methods of assessing in this study

Appendix L: Instructors' assessment packages analysis

Document analysis: criteria used to determine the focuses of instructors' assessment feedback

Instructor	Docs	Verification: affirm the correct and incorrect answers; indicate the places that are not completely correct or incomplete; provide scores	Directive: directly provide the correct answer when it is wrong; supplement the correct answer when it is incomplete	Scaffolding: drive learner towards solving questions through hints, cues, questions; encourages further learning
I1	T	18	3	2
	E	41	9	4
I2	T	11	6	-
	A	1	6	6
	E	45	15	5
I3	T	14	8	1
	A	7	2	4
	E	51	4	5
I4	T	23	7	2
	A	1	4	6
	E	31	3	1
I5	T	16	3	2
	E	36	-	1
I6	T	23	4	3
	A	3	1	5
	E	37	15	3
I7	T	13	11	-
	A	2	-	4
	E	31	9	2
I8	T	21	5	1
	A	1	2	3
I9	T	16	8	1
	A	3	1	4
	E	34	11	-
Total		479/70.3%	137/20.1%	65/9.6%

Key: I1-I9 represents instructor 1 to instructor 9

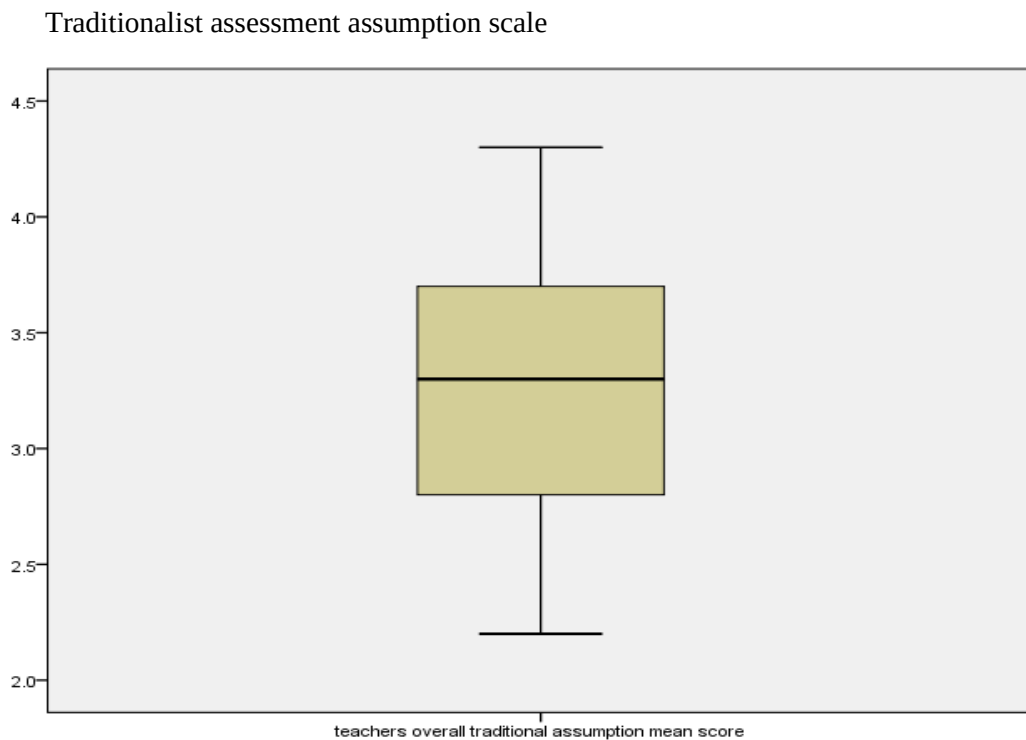
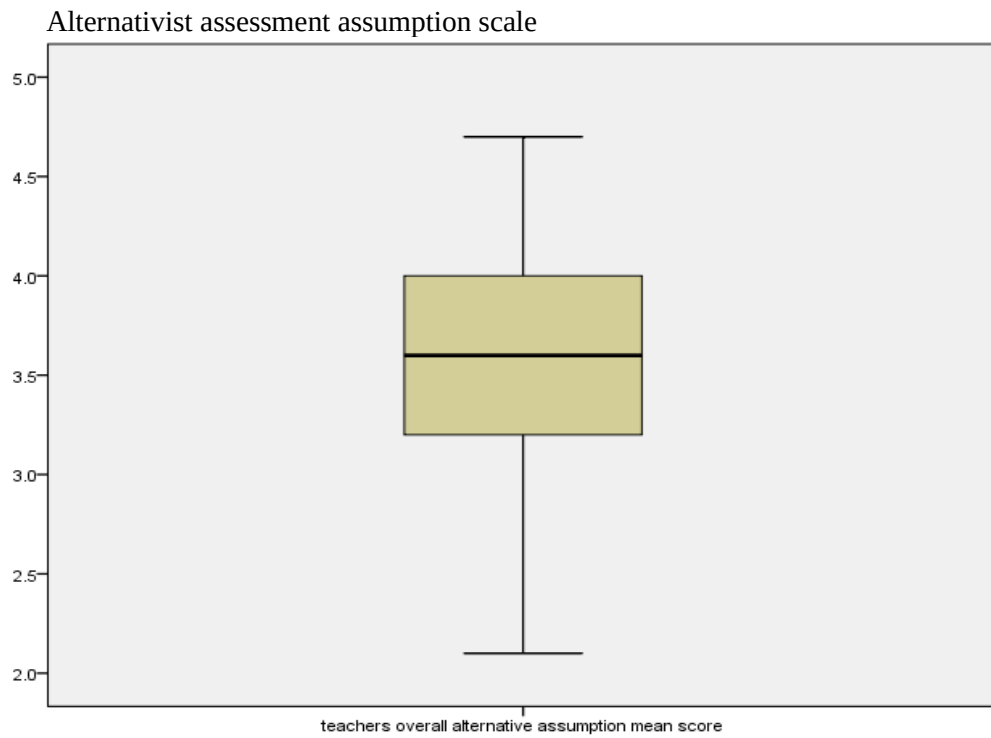
T: test/quiz

A: assignment

E: major exam

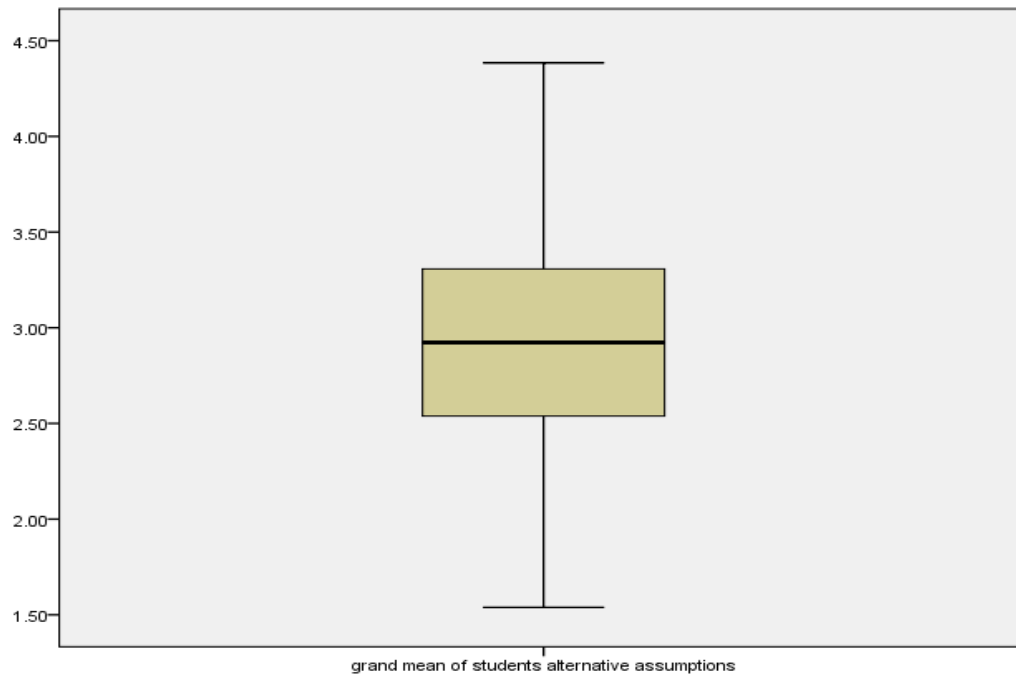
- No evidence in the document

Appendix M: Box plot representing instructors' assessment assumption data distribution of normality



Appendix N: Normality test for students' assessment assumption data using box plot

Alternative assumption scale



Traditional assumption scale

