Student Peer-Assessment of Learning in Some Selected Universities of Ethiopia: Practices and Challenges

Dame Abera[1]

Received: 3 November 2022; Accepted:22 May 2024

Abstract: *The major purpose of this study is to examine the extent to which student peer-assessment (SPA) of learning is practiced and the factors that affect its implementation. A mixed methods research design was employed in the study. Instructors, College Deans, department heads and undergraduate students were the major sources of data for this study. A total sample of 293 participants (258 questionnaire respondents, and 35 key interview informants) was involved in the study. Purposive and proportionate stratified random sampling techniques were employed for selecting the samples. Questionnaires and interviews were used to collect data. Both descriptive statistics (mean, percentage and standard deviation) and inferential statistics (one-sample t-test, independent samples t-test and one-way ANOVA) were used to analyze the quantitative data, while the thematic analysis technique was employed to analyze the qualitative data. The result of the study revealed that SPA is not aligned with the social constructivist principles and theories of learning, in the context of the three sample universities. The study also showed that the level of practising SPA of learning is very low in the context of the sample universities. The study also found that lack of adequate awareness, knowledge and skills among the students on how to apply peer-assessment techniques; lack of standard working document (peer assessment policy, guideline or directive); strong resistance from the instructors to maintain their power and control over assessment; threat from the side of instructors that* SPA *diminishes teachers' power in decision making and that sharing assessment with students lowers the standards; and lack of clear criteria for the peer-assessment tasks were reported by the study participants as the major bottle necks to the effective implementation of SPA. The study findings generally show that the assessment scheme in the context of higher learning institutions in Ethiopia has not yet transformed from the traditional teacher-centred notion to the modern student-centred conception; it is still driven and dominated by the teacher and did not create a window of opportunity for students to engage in the design and judgment of their own performance and assessment tasks.*

Key terms: Scaffolding, peer-review, peer feedback, student-centred assessment, cooperative learning strategies, and student peer-assessment

[1] Associate Professor of Psychology; School of Psychology, College of Education and Behavioral Studies, Addis Ababa University; Email: dame.abera@aau.edu.et

Introduction

The Contexts of the Study

Assessment as the systematic collection, analysis and interpretation of information (Olutosin & Gabriel, 2018) is classified as summative and formative (Clark, 2011). Summative assessment (assessment of learning) is judgmental (Huinker & Freckmann, 2009), while formative assessment [assessment for learning & assessment as learning] focuses on generating continuous feedback on student performance and improving learning (O'Neill, 2015). Assessment for learning is a teacher-dominated activity in which the teacher designs assessment tasks (Heritage, 2007), regularly monitors the learner's progress (Andrade, Huff & Brooke, 2012; Spiller, 2012), and provides ongoing feedback on what the learner can do (Clark, 2011; WNCP, 2006). Assessment for learning, as a traditional approach to assessment (Mc Sweeney, 2012; Olutosin & Gabriel, 2018)., is a teacher-centred/dominated activity that enables the teachers to monopolize power and leaves the learner passive (Li & Chen, 2016).

Research and theory suggest that such a traditional form of assessment is rooted in the behavioral approach to learning (Spiller, 2012) that highly dominated both the instructional and assessment activities until the 1970s (Stahl, 2003); that separated instruction from assessment (Donnelly & Fitzmaurice, 2005; Hodges, Eames & Coll, 2014); that considers students as passive (Gullo, 2005); and that suggests assessment should solely be conducted by the teacher (Lidz, 2003; Reece & walker, 2003).

Assessment as learning refers to a student-centered/dominated activity (O'Neill, 2015) that gives more power to the students in the assessment of learning (Juwah, et al., 2004; Ross, 2006; Scott, 2017; Siow, 2015; WNCP, 2006), promotes the development of meta-cognitive skills in students (Clark, 2011; Kearney, 2019; Topping, 1998), and provides opportunities for students to develop reflective practices (Mc Sweeney, 2012). Relatedly, peer-assessment is defined as a student-centered activity in which students make judgments about the quality of their

peers' work and provide feedback to each other using standard criteria for grading (Bozkurt, 2020; Costelloe & Egan, 2020; Li & Chen, 2016; Qu & Yang, 2010; Topping, 1998).

Peer assessment involves three interrelated processes: learning, assessment and feedback (Siarova, Sternadel & Mašidlauskaitė, 2017). As a learning process, peer assessment allows students to learn the skill and the content while assessing each other (Msiza, Zondi & Couch, 2020). As an assessment process, peer assessment promotes an instructional approach that allows learners to consider and evaluate the level, value or quality of the intellectual product of an equal-status learner, - a peer (Olutosin & Gabriel, 2018). As a feedback process, peer assessment emerged as a valuable way to provide participatory, active and social feedback to others (Double, McGrane & Hopfenbeck, 2020), and gain similar feedback from others (Topping, 1998). In peer-assessment, feedback is more immediate, timely, and individualized (Msiza, Zondi & Couch, 2020; Nortcliffe, 2012; Qu & Yang, 2010).

Moreover, in peer-assessment, students assume the roles of an assessor, assessee or both, coupled with a sense of accountability and responsibility (Lysaght, 2015). For instance, as an assessor, learners learn as receivers of the assessment (Rahman, 2015); and as assesses, learners learn from the given assessment feedback (Brown & Harris, 2014; Hanrahan & Isaacs, 2001; Topping, 1998). Not only this, but if it is properly designed, peer-assessment shifts a student's traditional conception of the *teacher as the only expert* and *sole source of feedback* (Clark, 2011; Sendziuk, 2010; Spiller, 2012; Trumbull & Lash, 2013) to place greater emphasis on learners' agentic engagement with the feedback processes on learners' active engagement (Andrade, Huff & Brooke, 2012; Bozkurt, 2020; Costelloe & Egan, 2020; Crowell, 2015; Dann, 2014; Nortcliffe, 2012).

The traditional curriculum and teacher-centred assessment practices that were rooted in the behavioural model (Stahl, 2003) started to be challenged beginning in the 1970s by the constructivist theorists that

advocate the student-centred curriculum and assessment (Hanrahan & Isaacs, 2001; Heritage, 2010; Mc Sweeney, 2012). Contrary to the behavioural model, the constructivist model strongly believes that learning is enhanced in the context of student-centred assessment (McMillan & Hearn, 2008; Trumbull & Lash, 2013). Constructivist generally believes that the teacher is not the sole assessor of students' work (Donnelly &Fitzmaurice, 2005) and that there is a strong alignment between the social construction of knowledge and student-centred peer assessment (Heritage, 2010).

In the context of higher learning institutions, the rise of student-centred assessment [the peer reflection phase), is based on the failure of the current traditional assessment practices to focus on developing students' ability to reflect on their learning (Double, McGrane & Hopfenbeck, 2020; Li & Chen, 2016). In recognition of this, most higher education institutions (HEIs) are now shifting their curriculum and assessment from teacher-centred to student-centred (O'Neill, 2015). Thus, this is exactly how peer-assessment is increasingly gaining attention in higher education institutions now a day (Msiza, Zondi & Couch, 2020).

Of course, a number of push and pull factors contributed to the current shift in the curriculum and assessment from the traditional approach to the modern approach in the context of HEIs. One push factor is a recently proposed paradigm: student-centred modularized curriculum and learning-focused assessment methods, in the context of HEIs across the globe (Murtagh & Webster, 2010). This new paradigm promotes independent and lifelong learning (Karami & Rezaei, 2015), active engagement of students in the learning process (McMillan, 2011), active involvement of students in the assessment (Donnelly & Fitzmaurice, 2005), and management of their learning (Juwah, et al., 2004). Relatedly, switching emphasis in university education from teaching to learning and from teacher management to student self-direction has also mounted interest in the rise of peer-assessment (Hanrahan & Isaacs, 2001). Moreover, two recent trends in education:

the design of competency-based curricula and the involvement of students in assessment urge HEIs to modify their educational practices (Sluijsmans & Prins, 2006). Consequently, peer-assessment is the outcome of changes in the curricula from teacher-centred to a logical interest in learner-centred curricula (Meihami & Razmjoo, 2016).

The second driving force is a changing context from the era of testing into an era of assessment (Dochy, Segers & Sluijsmans, 1999). As to these authors, the era of testing can be characterized by a complete separation of instruction from the testing activities, while the era of assessment promotes the integration of assessment and instruction, seeing the student as an active person who shares responsibility, reflects, collaborates and conducts a continuous dialogue with the teacher.

The third push factor is a growing demand for lifelong learners and reflective practitioners that stimulated a re-evaluation of the relationship between learning and its assessment and has influenced, to a large extent, the development of peer-assessment (Dochy, Segers & Sluijsmans, 1999). Research and theory show that traditional testing methods do not fit well with such goals as lifelong learning, reflective thinking, criticality, and problem-solving (Hanrahan & Isaacs, 2001).

The fourth push factor that contributed to the rise in the use of student peer-assessment is the emergence of information communication technology (ICT) as seen in the use of automated peer-assessment tools on e-learning platforms in HEIs (Msiza, Zondi & Couch, 2020). Moreover, as the climate in educational institutions moves steadily in the direction of increased efficiency and effectiveness, peer-assessment can be taken as an alternative way of assessing students' work either for feedback or for grading purposes while minimizing the cost in staff time (Hanrahan & Isaacs, 2001; Double, McGrane & Hopfenbeck, 2020).

Similarly, cooperative learning principles, scaffolding, and social constructivism of Vygotsky were proved to be the theoretical foundations

of peer-assessment (Dann, 2014; Sluijsmans & Prins, 2006). Social constructivism views learning as occurring within a social context (Hodges, Eames & Coll, 2014; McGarrigle & Carlow, 2013; Sardareha & Saad, 2012). Furthermore, social constructivism considers peer-assessment as an interactive process in which teachers and peers help learners use their zone of proximal development (ZPD) to progress to the next step in their learning (Sardareha & Saad, 2012). In the context of social constructivism, the teacher serves as a mediator between the student and the learning goal, providing scaffolding (i.e., learning support as a language translator) to aid the attainment of the goal (Ndoye, 2017; Olutosin & Gabriel, 2018).

Numerous sources (e.g., Kearney, 2019; McGarrigle & Carlow, 2013; Msiza, Zondi & Couch, 2020; Ndoye, 2017; Topping, 1998) show that peer-assessment, scaffolding, and cooperative learning are mutually interrelated and systematically complement one another. In support of this, Michael (2015) said that peer-assessment is a type of cooperative learning strategy in which students assess, provide feedback on the work of their peers and receive feedback from others. This implies that assessment is not a unidirectional activity, rather it is a reciprocal activity in which both teachers and students are involved as collaborative partners in generating feedback (Topping, 1998). Moreover, the Vygotskian concept of scaffolding learning (partially supported by a more competent other) is also involved in peer-assessment depending on whether the peer assessor merely identified weaknesses or strengths in the assessed work or suggested how the work could be improved (Trumbull & Lash, 2013). Scaffolding is generally believed to improve both the quality of peer assessment and increase the amount of feedback assessors provide (Dann, 2014; Double, McGrane & Hopfenbeck, 2020; Michael, 2015).

However, despite increased interest, peer assessment remains at the margin of assessment practices in higher education (Bozkurt, 2020; McMillan & Hearn, 2008). Several factors are associated with low levels of practicing student peer-assessment in the context of HEIs (Siarova,

Sternadel & Mašidlauskaitė, 2017). Such factors as lack of institutional peer-assessment standard manual, guideline or policy (Qu & Yang, 2010; Meihami & Razmjoo, 2016); failure of peer-assessment to addressing anonymity (Msiza, Zondi & Couch, 2020); learners' tendency to value their lecturers' feedback more often than that offered by their peers (Juwah, et al., 2004; McGarrigle & Carlow, 2013); lack of ownership of tasks by students (El-Koumy, 2010; Forsell, Frykedal & Chiriac, 2021; Long, 2000; Ross, 2006; Sendziuk, 2010; Shepard et.al., 2005); lack of adequate awareness, experiences, and skills of students on peer-assessment tasks (Reynolds, Miller, & Weiner, 2003); teachers' intention to retain all the ownership and power in the assessment process (Nicol, & Macfarlane-Dick, 2005; Siow, 2015; WNCP, 2006); conceptualizing feedback as a transmission process from the teacher to the learner (Juwah, et al., 2004); doubting the pedagogical values, quality, validity and reliability of peer-assessment data (Brown & Harris, 2014; Nortcliffe, 2012; Topping, 1998); teachers' lack of student-centered assessment literacy (Meihami & Razmjoo, 2016); and lack of clear and explicit peer-assessment criteria (Ross, 2006) are strongly associated with low level of practicing student peer-assessment in the context of the HEIs.

Statement of the Problem

Evidence shows that transformation has been made in the curriculum of higher education institutions (HEIs) from *teachin*g to *learning*, from *teacher-centred* to *student-centred* and from *traditional* to *modular*. However, the assessment scheme in HEIs has not been changed at the same pace and at the same level that the curriculum has been changed. At the same time, despite a strong claim in the literature that efforts have been made to apply the *principles of constructivist learning* in designing curriculum, these principles were not used in the design of the *assessment scheme*. This shows that students in HEIs are still highly exposed to the *traditional teacher-centred assessment*, where a window of opportunity was not left for the application of *student-centred assessment*. This traditional, teacher-dominated, assessment

authorizes the teacher to set the assessment criteria, evaluate learners' performance and provide feedback to learners to improve their learning. Such a traditional assessment approach is not in line with the current view of constructivist learning and assessment (Bozkurt, 2020). Moreover, though peer-assessment has theoretical roots in the social constructivism of Vygotsky, it is not still practically aligned with this notion of constructivism. In the social constructivist principle, students construct their knowledge and meaning through active interaction with others (Scott, 2017), though such a practice is missing in the context of the HEIs of Ethiopia.

In addition, though peer assessment practice is derived from developments in social learning theories and current understandings of feedback processes (Spiller, 2012), it remains a relatively uncommon assessment technique in the HEIs (Nicol, & Macfarlane-Dick, 2005).

At the same time, it is still the norm that academic teachers *retain all the ownership and power in the assessment process and make all the choices*, where which substantively limits the potential of learner development (Bozkurt, 2020). Similarly, despite an increased interest to use peer-assessment in higher learning environments, the assessment activity is still largely controlled by teachers (Nicol, & Macfarlane-Dick, 2005; Scott, 2017), leaving *little space* for the students to learn, construct knowledge and make meaning by being actively engaged in the assessment activities (Murtagh & Webster, 2010; Rust, 2002).

Despite strong advocacy in numerous local and national policy and program documents, including the New Education and Training Policy of Ethiopia (MoE, 2023) and various *strategic education documents* such as the Education Sector Development Program (ESDP V & VI), School Improvement Program (SIP) and Continuous Professional Development (CPD) that students should be involved in assessing their learning processes, so far nothing is known about the level of practising peer-assessment in the context of higher education institutions (HEIs) in Ethiopia.

On top of this, there is no comprehensive, adequate and up-to-date research-based data on the extent to which student peer-assessment of learning is practised; what factors hinder the level of practising SPA as well as whether the level of practising SPA varies across the Ethiopian higher learning institutions. Thus, the current study is an attempt to fill in such gaps; consolidate the existing data and add new shreds of evidence to the existing literature related to the practice and challenges of student-peer assessment of learning in the context of Ethiopian HEIs.

Purpose of the Study

The overall objective of this study is to assess the level of practice and challenges hindering the practice of student peer-assessment [SPA] of learning in some selected universities in Ethiopia.

Specifically, this study intends to:

- Examine the level of practising SPA in the sample universities
- Investigate if there is a statistically significant difference between teachers' and students' ratings on the level of practising SPA
- Investigate if there is a statistically significant variation in the level of practising SPA by the sample universities, colleges, gender as well as teaching experiences and academic rank of the university instructors
- Examine the extent to which the factors rated by instructors affect the level of practising SPA in the sample universities

Methodology

Research Design

This study employed a *mixed methods research approach* based on its appropriateness to the study's research questions (consisting of both closed and open-ended) and the generation of both quantitative and qualitative data in a single research (Bazeley, 2004; Bryman, 2006;

Greene, 2008; Johnson, & Onwuegbuzie, 2004). Specifically, the current study employed a *concurrent triangulation design* because this allows collecting of both the quantitative and qualitative data sets simultaneously or at the same time (Creswell, 2012).

Sources of Data

The study involved university instructors, college deans, senior undergraduate regular students, and department heads of the sample colleges and universities. In this study, the universities were selected as the main sources of data based on the Ministry of Science and Higher Education's (MoSHE, 2020c) scheme of differentiating Ethiopian public universities.

Sample and Sampling Techniques

In this study, a multi-stage sampling technique was employed to select the study sites, samples and study participants. In the first stage, MoSHE's (2020c) scheme of *differentiation* of Ethiopian public universities as research applied, and specialized institutions was taken as is (purposively). As MoSHE's (2020c) scheme of classification, 8 universities (Addis Ababa, Arba Minch, Bahir Dar, Gondar, Haramaya, Hawassa, Jimma and Mekelle) were categorized under research; 15 under applied (Arsi, Asossa, Aksum, Dila, Dire Dawa, Debre Birhan, Debre Markos, Kotebe, Jigjiga, Wolita Sodo, Semera, Ambo, Wellega, Welkite and Wollo universities), and 3 under specialized educational institutions (Addis Ababa and Adama Science and Technology universities as well as Federal TEVT). Then, from each of MoSHE's three categories/clusters of public universities, one university was selected using a simple random sampling technique (lottery method). Accordingly, Addis Ababa University from the research cluster, Adama Science and Technology University from the specialized institutions' cluster, and Arsi University from the applied cluster were selected as sites for the present study.

In the second stage, colleges were randomly selected as samples of the study from among the existing and actively operative colleges in the three sampled universities. Accordingly, three colleges, namely, the College of Education and Behavioral Studies, College of Business and Economics, and College of Social Sciences from Addis Ababa University; Two Schools namely the School of Civil Engineering and the School of Applied Natural Sciences from Adama Science and Technology University; and four colleges namely, the College of Education and Behavioral Sciences, College of Agriculture, College of Business and Economics, and College of Social Sciences from Arsi University were selected at random (through lottery method).

In the third stage, departments were randomly selected from each of the sampled colleges. Accordingly, two departments from the College of Business and Economics, namely; Accounting and Economics, four departments from the College of Education and Behavioral Studies - Psychology, Curriculum and Instruction, Special Needs Education, and Educational Planning and Management; two departments from the College of Social Sciences: Social work and Sociology from Addis Ababa University; four departments from the School of Applied Natural Sciences -Chemistry, Biology, Physics and Mathematics; two departments from the School of Civil and Architectural Engineering: Architecture and Water Resource Management from Adama Science and Technology University as well as two departments from College of Agriculture: Natural Resource Management and Agro-economics; three departments from College of Business and Economics: Accounting, Economics and Management; four departments from the College of Social Sciences: Geodesy, Geography, English and Civics, and four departments from the College of Education and Behavioral Sciences: Educational Planning and Management, Curriculum and Instruction, Psychology and Special needs education from Arsi University were randomly selected as samples of the current study.

In the fourth stage, for the quantitative part of the study, actual participants: instructors and senior students (3^{rd} & 4^{th} batches due to

their extensive exposure to the university assessment practices) were selected from each of the sampled departments using stratified random sampling technique, where the strata are student batches as well as the experiences, gender, and ranks of instructors. Stratified sampling was employed in this study for it creates a window of opportunity to fix the sample proportions and then draw the study participants proportionate to the relative size of the target population in each stratum. Similarly, a simple random sampling technique (lottery method) was used concerning stratification for it gives an equal chance/opportunity for every target population in each stratum to be selected as samples of the study. Generally, for the quantitative part, the current study employed *30%* of the target population as samples based on Neuman's (1997) guideline. As suggested by Neuman, for the study population of *1000* or *under*, a sample *ratio of about 30%* or more is an acceptable representative sample for quantitative studies. Consequently, based on the data obtained from the respective departments of the three universities the population size of instructors in the overall sampled departments was 326 while that of students was 533. As a result, 30% of 326 approximately yielded *98* sample size teachers while 30% of 533 yielded *160* sample size of students.

For the qualitative part of the study, *eight* deans from the sampled colleges and *27* department heads of the sampled departments were selected as samples of this study using a *purposive sampling technique*. The selection of samples for the qualitative part was mainly based on their experiences, relevance to an issue and position or responsibility in the existing structure to monitor and guide the implementation of the assessment scheme. In supporting this view, the existing literature (e.g., Bazeley, 2004; Onwuegbuzie & Collins, 2007) suggests that in *mixed methods studies*, because of the complexities of data it generates, samples for qualitative investigations tend to be smaller and drawn purposively.

Instruments

The study involved key informant interviews and *questionnaires*. An interview guide that consists of 13 items was developed by the researcher to capture information on the level of practising SPA and challenges to its effective implementation. The key informants of the interview were college deans and department heads running the under-graduate programs and offering modularized curriculum/modules.

Similarly, the researcher developed a questionnaire consisting of *three* parts. Part I of the questionnaire contains seven socio-demographic questions. Part II of the questionnaire consists of 26 questions that are intended to measure the level of practising SPA. Part III of the questionnaire consists of 22 questions that are aimed at assessing factors that affect the level of practising SPA. In developing both the KII and questionnaire, an extensive review of related literature was made by the researcher. Moreover, to establish the *content validity* of the instruments (KII and questionnaire), a panel of six subject matter experts (SMEs), who are practitioners and have expertise in the area of psycho-metrics (assessment, testing, and measurement) were invited to examine the content adequacy of the tools (the original English version since the study participants were instructors and senior 3rd and 4th batch university students). Consequently, the SMEs were requested to engage in judging the appropriateness, adequacy, relevance, quality and proper wordings of each item in measuring the specified constructs, where their comments were thoroughly identified, examined and then incorporated to refine the data collection tool for the main study. Moreover, the reliability of the quantitative part [questionnaire instrument] in terms of Cronbach alpha was computed from the pilot study data captured by administering a questionnaire to 30 senior undergraduate students and 40 instructors (a total of 75 participants) that have similar characteristics in terms of batches, and experiences with the main study samples. Consequently, reliability in terms of Cronbach alpha was computed to be .78, which is within an acceptable range for utilizing the questionnaire as an assessment and research tool.

Moreover, the respondents of the questionnaire were informed to respond to each item on a scale ranging from 1 (strongly disagree) to 5 (strongly agree). At the same time, a composite score was computed for each of the two sub-scales separately and a *high score* on each sub-scale indicates that the student peer-assessment (SPA) is widely implemented; and that the factors rated by instructors with high scores are found to significantly affect the level of practising SPA.

Procedures of Data Collection

At first, having secured ethical approval for the project from the Research Ethics Committee of the School of Psychology at Addis Ababa University, the respective college deans, department heads and program coordinators of the three sampled universities were consulted to discuss how to contact the actual study participants and facilitate the data collection process. Then, program coordinators of the respective departments facilitated the contact to be made with the study participants to get their free oral consent or willingness to participate in the study or not, explain the purpose of the study and fix the schedule for actual data collection. Accordingly, the questionnaire tool was administered to the sample students in a classroom and at a time prearranged with them at their convenience, where the researcher was personally present at all levels of the questionnaire administration. The instructors, after briefly orienting on how to fill in the questionnaire, were given the questionnaire tool to fill in at their office and return it to the researcher within the scheduled and agreed-upon time frame. At the same time, all the key informant interviews were conducted by the researcher at the informants' convenient time and place, where an attempt was also made to record the responses using both *field-notes* and *audio/video recordings*. On average, the interview took 55 minutes to an hour. Generally, both the data gathering tools (KII and questionnaire) were administered in a face-to-face approach, where the researcher was available for the clarification of any potential respondents' queries or concerns.

Data Analysis

The thematic analysis method was used to analyze qualitative data captured through the key informant interview. Before qualitative data analysis, the informants' responses to each question of the interview guide were transcribed verbatim to ensure that the informants' messages, opinions or views were properly captured. The rationale for using the *thematic analysis method* in the present study was that it helps code and organize data into categories and then identify, analyze and report patterns (themes) within the data (Bazeley, 2004; Braun & Clarke, 2006). Moreover, to secure confidentiality both in the transcription and analysis of the qualitative data, codes (such as Participant$_1$, P$_2$, P$_3$ etc.) were used to designate the study participants rather than their real names, where analysis of the qualitative data was conducted manually (no qualitative software was employed).

Descriptive statistics (such as mean, percentage, and standard deviation) and inferential statistics (such as one-sample t-test, independent samples t-test and one-way ANOVA) were used to analyze quantitative data. One sample t-test was used to examine the level of practising SPA by comparing the observed sample mean against the hypothetical mean obtained from the scale used for assessing the practice of SPA. An independent t-test was used to examine variations in the level of practising SPA against the gender of instructors while one-way ANOVA was used to examine variation in the level of practising SPA against types of the universities as well as experience and academic rank of instructors. Before the quantitative data analysis, careful data entry using SPSS and data checking were thoroughly made.

Data screening, sorting out or cleaning was made to figure out any missing data or outliers and, then take corrective measures such as mean or median replacement method to prevent the observed outliers. In addition, before using the above parametric tests in data analysis, their model assumptions (such as normality, homogeneity and independence) were tested, and where result of the test confirmed that

the three assumptions were met to carry out the required analysis using the proposed statistical tools.

Results

Demographic Data

Table 1:

Students by their University, College, Department and Gender

University	***Colleges***	***Departments***				
ASTU	SoCEA		Batch	Number		
				M	F	T
			3rd	9	2	11
			4th	15	3	18
		Architecture				
Arsi	Agriculture	NRM	3rd	5	1	6
			4th	4		4
		Agro-Economics	3rd	6		6
	CoEBS	EDPM	3rd	8	12	20
AAU	CoEBS	Psychology	3rd	5	3	8
			4th	4	2	6
		SNE	3rd	3	2	5
			4th	2	3	5
		EDPM	3rd	4	1	5
			4th	5		5
	CoBE	Accounting	3rd	6	4	10
			4th	2	4	6
		Economics	3rd	7	3	10
			4th	3	3	6
	CSS	Social Work	3rd	7	2	9
			4th	5	2	7
		Sociology	3rd	3	2	5
			4th	4	3	8
Grand Total						***160***

Key: C & I = Curriculum & Instruction; EDPM = Educational Planning & Management; SNE = Special Needs Education; NRM = Natural Resource Management; WRM = Water Resource Management; SoCEA = School of Civil Engineering & Architecture; SoANS = School of Applied Natural Sciences; CoEBS = College of Education & Behavioral Studies;

CoSS = College of Social Sciences; CoBE = College of Business & Economics.

Table 2. Instructors by their University, College and Department and Gender

University	***College***				***Department***			
Arsi	CoEBS	M	F	T	Psychology	M	F	T
		9		9		4		4
					Curriculum	1		1
					SNE	1		1
					EDPM	3		3
	Agriculture	5		5	Agro-Economics	2		2
					NRM	3		3
	CoBE	11		11	Dev't Economics	4		4
					Accounting	5		5
					Management	2		2
	CoSS	7		7	Geodesy	1		1
					English	2		2
					Geography	2		2
					Civics	2		2
ASTU	SoCEA	19	1	20	Architecture	15	1	16
					WRM	4		4
	SoANS	11		11	Chemistry	1		1
					Physics	5		5
					Mathematics	4		4
					Biology	1		1
AAU	CoEBS	12	2	14	Psychology	3	2	5
					C & I	3		3
					Curriculum	4		4
					SNE	2		2
	CoBE	9	1	10	Accounting	5	1	6
					Economics	4		4
	CoSS	12		12	Social Work	7		7
					Sociology	5		5
Grand Total				**98**	**Grand Total**			***98***

The socio-demographic data summarized in Table 1 and Table 2 show that both the study sites (universities, colleges and departments) and study participants (deans, department heads, students and instructors) have drawn sufficient samples from the target population; that they are capable of describing the study variables; that they are representative of the population; that they are eligible as data sources; and that the results

obtained from these representative samples can be generalized to the study population.

Quantitative and Qualitative Data Analysis

Practice of SPA

Table 3. Level of Practicing Student Peer-Assessment (SPA)

Variable	*N*	*M*	*SD*	*df*	*Test value*	*t*	*P*
PSPA	98	66.3	11.9	97	78	-9.71	.000

PSPA = Practice of student peer-assessment; SD = Standard Deviation; Test Value = the hypothetical population mean that was computed from the rating scale (1 to 5) of the 26 items [1+2+3+4+5 = 15/5x26 = ***78***]

The result of a one-sample t-test, as presented in Table 3, revealed that there is a statistically significant difference between the observed mean obtained from a sample of instructors' ratings and the hypothetical mean (test value) calculated from the scale [t (97) = -9.71, P = .000]. The mean difference of 11.69 between the observed sample mean (*M* = 66.3) and the test value (*M* = 78) indicates that the level of practising SPA is *very low* in the context of the three sample universities (AAU, ASTU and AU).

Likewise, in response to an interview question, *to what extent is SPA practised in your course/class?,* informants said that SPA is not being practised in the context of their universities as there is no adequate awareness among teachers and students on the one hand and, no standard working documents (policy, guidelines or directive) on how to implement it on the other hand.

In line with this, a dean from one college said:

> *As a dean, I always inform, remind and push my instructors to make use of continuous assessment (consisting of individual and group assignments, projects, quizzes, and formal tests) in addition to the mid and final exams in assessing their students' learning and performance. This is exactly what is suggested in the national modular courses and curriculum materials prepared by the MoE and implemented by universities. So far, we have not been communicated by any authority (MoE or the university management) to apply the student peer assessment techniques either as an alternative tool or as a supplement to the existing teacher-centred assessment. In short, even the idea of student peer assessment is new to me and my instructors as well.*

The main message to be conveyed in the contents of the above quotation is that let alone involving students in the design of peer-assessment, choice of assessment tasks and setting of assessment criteria, the very concept of student peer-assessment is not yet introduced and made familiar to the university community.

Similarly, a dean from another college said:

> *Truly speaking, our instructors did not even properly understand and apply continuous assessment techniques, let alone practice the student peer assessment in their courses. To most of us, continuous assessment means continuous testing and adding the scores to the students' grades. We have tried to organize several training sessions for our instructors on the what, the advantages and how to apply continuous assessment into their respective courses, but nothing has been changed so far,*

> *significantly. Generally, in my college, no single instructor is practising the student peer-assessment as there is no well-established culture in the University for involving students in the assessment of their peers.*

The message conveyed in the contents of the above quotation shows that the assessment of student learning is still designed, derived and centred toward the teacher; that students are only the subject of assessment (still tested and assessed by the teacher), where they do not have any opportunity to appraise their peers' performance; that students have no opportunity to develop through assessment tasks, self-regulation and self-criticism; and that there is no enabling environment for learners to evaluate their learning progress.

At the same time, a head of one department said:

> *So far, there is no clear direction, policy, or guideline concerning applying the SPA in our university. Of course, starting from the ministry level to the smallest unit of the university (department level), every document (including the course outlines) advocates the use of continuous assessment in assessing student learning. As everyone knows, continuous assessment is still designed, planned & driven by the instructors, where the students' stake is doing the assignments, projects and tests that are solely marked by the instructors. Even, there is high resistance to applying continuous assessment on the side of the teacher claiming that students copy assignments from one another (wanting to gain grades in a shortcut) and that examinations help identify the level of students' understanding compared to continuous assessment. In short, the culture of involving students in designing assessment*

> *tasks and appraising the work of their peers is not yet established in our university*.

A critical examination of the content of the above quotation shows that at present, there is not any opportunity for the students to exercise peer-assessment (SPA) in the sample universities; the instructors do not have adequate awareness for applying SPA, and are not committed to serve as *role models* by demonstrating peer-assessment tasks to their students; that learners lack adequate awareness on the value of SPA in promoting self-development; and that there is no clear policy, guideline or criteria in the sample universities regarding the application of student peer-assessment.

Likewise, a head from another department said:

> *The entire university community (the university management, academic staff and students) are not well informed about the student peer-assessment, which, as I am hearing from you now, requires the instructors to involve students in the assessment of their peers' learning. Even, the existing culture (in which students copy assignments from others, instead of producing their original work, teachers do not want to lose their authority over the assessment, students lack motivation to learn, and most of them want to get grades without adequate effort) does not allow the practice of student peer-assessment. In the present situation, even if the students are let to judge their peers' work, they may give the same mark to all their friends, which, in turn, may lead to mistrusting the reliability and validity of its information.*

An important message contained within the contents of the above quotation is that SPA is not still institutionalized and well-established in the sample universities; that SPA is not currently practised in the sample

universities since a *classroom culture* that supports its practice is not created on the ground (peer assessment policy and guidelines were not developed, teachers and students were not trained on the procedures of student-centred assessment, learners give more attention to getting grades in a short-cut instead of learning to master the subject matter, gain knowledge, and understand learning tasks).

Variations in the level of practising SPA

Table 4. Variation in the Ratings of Teachers and Students Regarding the Level of Practicing SPA

DV	IV	*N*	*Mean*	*SD*	*df*	*t*	*P*
PSPA	Teachers' score	98	66.3	11.9	256	1.85	. 42
	Students' score	160	64. 2	10.8			

PSPA = Practice of Student Peer-Assessment; DV = Dependent Variable;

IV = Independent Variable; SD = Standard Deviation

As summarized in Table 4, the result of an independent t-test revealed that there is no statistically significant disparity between the ratings of teachers and students regarding the extent to which SPA is practised in the sample universities [t (156) = 1.85, p= . 42]. This means that both the students and the teachers rated the level of practising SPA as *very low* since the two sample means obtained from the ratings of teachers (*M* = 66.3) and students (*M* = 64.2) are far below the hypothetical population mean (Test value = 78).

Table 5. Variation in the Level of Practicing SPA by Gender of Instructors

DV	IV	*N*	*Mean*	*SD*	*df*	*t*	*P*
PSPA	Male	94	65.9	11.9	96	-1.68	. 097
	Female	4	76	7.7			

PSPA = Practice of student peer-assessment

As presented in Table 5, the result of an independent t-test revealed that there is no statistically significant variation in the level of practising SPA in the context of the three sample universities by gender [t (96) = -1.68, P = .097]. This means that both male and female instructors of the sample universities practice SPA at the same *low level* as the observed sample means for males (M = 65.9) and for females (M = 76) are below the hypothetical population mean (Test value = 78).

Table 6. Variations in Level of Practicing SPA by Instructors' University, College, Academic Rank, and Service Years

DV	*IV*		*N*	*M*	*SD*	*F*	*P*
PSPA	University	AAU	36	63.39	10.54	2.11	.126
		ASTU	30	69.33	12.59		
		Arsi U	32	66.75	12.34		
	College	Education	45	64.53	10.86	2.23	.056
		Natural Sciences	11	76.45	9.91		
		Social Sciences	7	60.71	60.71		
		Engineering	19	65.42	65.42		
		Agriculture	3	67.40	67.40		
		Business & Economics	11	68.00	67.00		
	Service year	1-10	40	66.73	13.51	.143	.934
		11-20	29	66.69	11.31		
		21-30	22	65.86	10.35		
		31-45	7	63.71	11.37		
	Academic rank	Lecturer	51	66.98	13.88	. 36	.78
		Ass. Prof.	32	64.84	10.82		
		Assoc. Prof.	12	68.08	5.74		
		Professor	3	63.33	2.08		

PSPA = Practice of student peer-assessment

As summarized in Table 6, the result of One-way ANOVA revealed that there is no statistically significant difference in the level of practising SPA by instructors' University [F (2, 95) =2 .11, P = .126]; Service years [F (3, 94) = .143, P = .934]; College [F (5, 92) =2 .23, P = .056], and academic rank [F (3, 94) = .36, P = .78]. This shows that across all the attributes of teachers considered in this study (e.g., university type, college type, service year & academic rank), the level of practising SPA is very low in

the sample universities since all the observed sample means are quite below the ideal or hypothetical population mean (*Test value* = 78).

Factors Affecting the Level of Practicing SPA

Table 7. Instructors' rating of factors that affect the level of practising SPA

Variable	*N*	*M*	*SD*	*df*	*Test value*	*t*	*P*
Instructors' rating	98	86.9	8	97	66	25.7	.000

Test value = Hypothetical population mean computed from the item rating scale of 1 to 5 for 22 items: 1+2+3+4+5 = 15/5 = 3x22 = ***66***

As presented in Table 7, the result of a one-sample t-test revealed that there is a statistically significant difference between the sample mean obtained from instructors' ratings on a *five-point rating scale* of factors affecting the level of practising SPA and the hypothetical mean or test value [t (97) = 25.7, P = .000]. This result clearly showed that the observed sample mean (*M* = 86.9) is significantly higher than the test value (*M* = 66), where the mean difference of 20.9 between the observed value and the test value means that the higher the score on the scale, the higher the agreement that the factors rated by instructors significantly contribute to or serve as major bottlenecks to the lower level of practising SPA in the sample universities.

Likewise, according to the instructors' ratings, both the teachers and students do not have adequate awareness, knowledge and skills on the assumptions and principles underlying SPA; teachers strongly doubt that students inflate their accomplishments or distort evaluative information to favor their peers or friends; there is a doubt to trust the reliability and validity of SPA information; there is lack of the classroom culture that supports the practice of SPA; there is fear among instructors that SPA diminishes teachers' power in decision making and that sharing assessment with students lowers the standards; there is no explicitly

stated SPA criteria in the curriculum materials; the students do not have adequate experience and training on how to use SPA; assessment is not constructively aligned with specific learning outcomes; there is no clear policy, standard manual or guideline on how to apply SPA; and the instructors consistently struggle to maintain their power and control over the assessment.

Relatedly, in response to the interview question, *what major factors do you think affect the practice of SPA in your university*, informants said that lack of clear peer-assessment policy and guideline in the university; lack of a well-established culture of involving students in assessing their own learning progress; lack of motivation and self-responsibility on the part of students for their own learning; lack of adequate experience and awareness on the part of students to evaluate the work of their peers; the fear that students either under or overestimate their peers' performance; teachers' resistance to maintain their tradition or status quo (*doing things the way they were taught*) and authority in the assessment tasks are the major challenges that lower the level of practicing SPA in the sample universities.

In connection to this, a dean from one college said:

> *Our administration and instructors are not well informed and experienced with the SPA scheme. The assessment tools our instructors are most familiar with include examinations and continuous assessment (such as projects, individual and group assignments, quizzes, regular tests). Such student-centered or learner-driven assessment scheme is not even explicitly mentioned in the new curriculum materials (such as even in the most recent freshman harmonized modules) for teachers and students. So, where such very important foundation was not laid down and procedures were not put in place both for teachers and students in the existing modular or course materials, how one expects the SPA to be endorsed in the university setting.*

The important message conveyed in the above quotation is that the sample universities as educational institutions, their academic staff and students are not psychologically and technically ready to apply the SPA techniques since nothing is explicitly stated both in the *instructional materials* (courses) and overall *curriculum materials* (syllabi) about the student-centered or *learner-driven assessment*. The quotation also indicated that starting from the policy level to the grassroots implementation level as well as from the highest structure of educational administration (MoE) to all levels of the university administration, the use of *student-centered assessment* (e.g., student peer-assessment-SPA) as alternative to or parallel to (as a complement to) the *teacher-centered assessment* did not get adequate attention that it deserves.

Similarly, another college dean said:

> *There are many bottlenecks to the proper application of SPA in the context of our university. One is lack of adequate knowledge, skills and experiences on peer assessment among the instructors and students. The other is a belief on the side of instructors that when the student-centered assessment scheme is fully utilized it will diminish their powers and influences on the students. Still the other is a fear on the side of instructors that peer assessment data may not be valid and reliable as there are many evidences showing that some students even copy assignment from peers, inflate self-scores and that of their peers, as well as they do not value their peers' assessment scores*.

The central point embedded in the above quotation is that the instructors of the sample universities still want to work in the existing '*comfort zone*' (claiming the continuity of the teacher-driven assessment scheme) as this maintains their control, authority, and influential power over the assessment tasks and resist the introduction of the student-centered

assessment at the front-line. It also conveys that there are no enforcing laws (policy documents, guidelines, or directives) put in place for the essential stakeholders and practitioners in the profession so as to introduce this alternative student-centered assessment scheme at the classroom level.

Relatedly, a head of one department said:

> *Many factors hinder the practice of SPA in the context of universities. One is that there is no standard manual or procedure for its use. The other one is that there is no classroom culture that permits the practice of student peer-assessment in HEIs. Still the other one is that the instructional and curricular materials themselves do not explicitly encourage the peer reviews, self-regulated learning, self-reflection, and self-criticism practices. Generally, SPA is not institutionalized in the university culture due to these and other factors.*

The central theme of the above quotation is that SPA is not given adequate attention both in the existing educational documents (even in the freshman modular courses, syllabi) and by the necessary stakeholders (policy developers, curriculum designers, university officials, assessment experts, the staff) and, as a result, its practice is still at the inception or *infancy stage*.

Discussion

Level of Practising SPA

Result of the present study showed that the level of practicing SPA is very low in the context of the three sampled public universities (AAU, ASTU & AU). This means that peer assessment as one vital aspect of the student-centered assessment schemes is not given due attention as

well as being practiced as an alternative tool or as a supplement to the teacher-centered assessment of students learning progress in the real context of the Ethiopian public Universities. The fact that peer assessment as a component of the student-centered assessment techniques is at the inception or infancy stage clearly indicates that the assessment tasks being practiced in the universities are still teacher-centered, teacher-dominated and teacher-derived. It means it is the classroom teacher that solely designs the assessment tasks, sets the assessment criteria as well as judges, marks and grades the works or performance of students.

Obviously, when the *assessment activity* (task design, choice of assessment criteria, or marking) and practice are solely owned or controlled by teachers, students remain passive and the fate of students totally depends on what their teachers offer them. Again, in such assessment platform, teachers monopolize all the assessment tasks (starting from the design stage of assessment tasks to the grading stage); students do not have any specific roles in assessing their own learning; teachers and students are not viewed as collaborators or partners in the assessment tasks (meaning their relationship is not flat, has clear command of chain, is viewed as hierarchical or top-down); students are considered as the subject of assessment [being tested and assessed]; and students do not feel a sense of ownership in their own learning. Similarly, in such an assessment system, there will be little or no opportunity for students to learn, construct knowledge, and make-meaning by assessing, judging or appraising the work of their peers as well as by providing or receiving feedback. The other very important implication of the current finding is that the students are still over-burdened with both the *summative assessment* that primarily focuses on grading students' level of achievement at the end of the course/module and the *teacher-dominated formative assessment* in which the teacher regularly designs assessment tasks, sets assessment criteria, conducts the assessments, marks the students' performance on the assessments, and provides feedback to students.

Such a teacher-centered, derived and dominated assessment activity is, in fact, contrary to what is stated in the philosophy of student-driven or student-centered peer assessment activity as well as the principles of constructivist theories that advocate the notion that learning is best enhanced in the context of student-centered assessment, meaning when students engage in assessing their own or their peers' work (Heritage, 2010). This also implies that too much reliance on the traditional teacher-driven assessment scheme leaves little or no opportunity for students to exercise their full potential, academic tasks and responsibilities through various assessment activities or tasks. At the same time, the observed teacher-monopolized assessment scheme confirms the existence of a huge gap between the theory and practice of SPA, where this is clearly a timely *warning sign* for the concerned stakeholders to exert concerted efforts to bridge the existing gaps so as to improve the prevailing situation.

Generally, the tendency of the university instructors to move forth and back in such a *vicious circle* (maintaining the continuity of such old traditions as: the teacher is the *sole knower*, it is the teacher who should control over the assessment tasks, the students are passive or over value their performance; resisting the introduction of modern alternative assessment schemes) seems to be driven by such a myth as *we teach the way we were taught*. In fact, holding and driven by such a myth may significantly affect teachers' current performance, thinking, attitude and ways of doing things (including the assessment practices). Not only this, being driven by such a myth may also allow the *old status quo*, customary practices, teachings, memories, and traditions to flash back and forth, and limit *one's professional life* to solely operate within the already established *comfort zone* or *known box.* In fact, thinking within the *prevailing box* let the university instructors to be detached from contemporary assessment paradigms, developments as well as up to date knowledge, thinking, and perspectives in assessment. It also hinders the teaching force from regularly enriching or updating themselves with innovative and recent developments in the assessment

arena (theories, principles, policies, digital technology, and research findings) and modernity.

In a similar way, the determination of university teachers to work within the present *comfort zone* leads them to promote the continuity of old and deep-rooted tradition of the teacher-centered assessment, at the expense of the alternative student-centered assessment, without making significant improvement, change or development in their professional career as well as classroom practices (teaching, assessment, and class management). Of course, demanding the continuity of the existing assessment tradition has long been enforced by or derived from the behavioral theories of learning that assume that the *child is passive* (Gullo, 2005); that learning and assessment should be conducted by the *knower* - the well-trained and knowledgeable teacher (Lidz, 2003); and that students learn better by observing and listening to what the teacher is doing in the class (Reece & walker, 2003).

On the contrary, constructivists and developmental theorists strongly advocate the necessity of student-centered assessment, believing that the child is active creator and processor of information; that the child's learning is facilitated when it is designed and provided based on the child's natural developmental order; that children learn by constructing their own knowledge through making *meaningful interactions* with others (peers, teachers); and that student-centered peer assessment promotes *knowledge construction* (Stahl, 2003). Generally, what all the above points makes clear is that the university instructors need to go beyond the comfort zone and think outside of the existing box so as to properly address the demands of the 21st century skills and the era of a digital technology.

In sum, the present finding is consistent with the reviewed literature. For instance, Spiller (2012) reported that though tremendous efforts have been made to design classroom learning opportunities that reflect the principles of constructivist learning, this principle is mostly ignored in the design and implementation of the student-centered or peer assessment.

Similarly, Scott (2017) reported that despite an increased interest to use the student-centered-peer assessment in the context of higher education learning environments, the assessment activity is still largely controlled, monopolized, monitored and driven by the classroom teachers.

Variations in the Level of Practicing SPA by some Socio-demographic Variables

The finding of the current study revealed that there is no substantive difference in the level of practicing SPA by the type of university, type of College, service years, gender and academic rank of the study participants. This means that the level at which SPA is being practiced is the same across all the study sites and sample characteristics considered in the present study. As also confirmed by the present study data, SPA is being practiced at low level or at an infancy stage in all the three sample public universities, and among the participants of the current study. In fact, in a situation where there are *no standard national level working documents* (policy, manual, guideline or directive) on the procedures and applications of the student-centered peer assessment, it seems not logical to expect high level of practicing SPA at this junction.

Not only this, the apparent lack of variation in the level of practicing SPA particularly among the university instructors who are less experienced and well experiences as well as those who have high academic ranks and low academic ranks may not be a big surprise. This is because, in a situation where both the public university instructors and students have no adequate orientation, awareness, skills and knowledge on how to design and carry out SPA, it is again illogical to expect them to practice SPA at a high level. Of course, the important lesson to be learned from this scenario is that the university community needs to have adequate awareness, skills and knowledge as well as needs to have clear criteria, policy and guideline on how to properly carry out student peer assessment tasks.

Factors Affecting the Level of Practicing SPA

In the current study, instructors rated that the level of practicing SPA is very low in the sample Universities and that many factors were rated by them to contribute significantly to this low level of practice. As to the rating of these university instructors such factors as: students' lack of knowledge and skills for assessing their peers' work; the fear that students distort evaluative information to favor their peers or friends; teachers' doubt on the reliability and validity of SPA information; lack of the classroom culture that supports the practice of SPA; the threat that SPA diminishes teachers' power in decision making; the fear that sharing assessment with students lowers the standards; lack of explicitly stated SPA criteria in the instructional (course modules) and curricular (syllabi) materials; lack of adequate experience and training for students on how to use SPA; tendency of the instructors to maintain their power and control over the assessment tasks; and lack of *standard working document* (policy, guideline, directive) on the procedures of student peer-assessment were figured out as severe bottleneck to effectively practicing SPA at the three sample public universities in Ethiopia.

Basically, the prevalence and severity of these challenges seem to consistently promote the long-lasting continuity of the traditional teacher-centered assessment scheme (which detaches students from learning by their own initiatives through *constructing their knowledge* and making-meaning using either the *inner resources* (critical thinking, self-reflection, self-criticism, self-regulation) or *external resources* (interacting meaningfully and purposefully with other peers), while blocking the opportunity to introduce the progressive and modern-day student-centered approach like peer-assessment that encourages the growth and development of students through the assessment activity or exercise. Generally, the main implication of the current finding is that a concerted effort should be made to properly orient and bring all the essential stakeholders (education experts, education policy makers, university management, and academic staff) to be on the *same page* or *on equal footing* to lend their ears to properly design and implement such

an authentic assessment scheme. This finding is certainly an indication, a *warning sign* or a call particularly for the education policy makers and education experts to develop and communicate a standard working document (such as the peer assessment policy framework or guideline) detailing the procedures to be followed in its implementation.

In fact, the present finding is in line with the existing body of knowledge. For instance, according to Miller and Weiner (2003), the low level of practicing student peer assessment can be linked to lack of adequate experiences, skills and focus on peer assessment by teachers and students. At the same time, Juwah et al. (2004) reported that the low utilization of SPA is related to the tendency to conceptualize feedback as a transmission process from the teacher to the learner. Moreover, Brown and Harris (2014) reported that the tendency of peer-assessors to have inflated perceptions of their accomplishments can also undermine the use of SPA. Besides, according to McGarrigle & Carlow (2013), controversy in the pedagogical values, reliability and validity of peer assessment is a challenge to its practice. Furthermore, learners' and teachers' lack of assessment literacy were the most important challenges of implementing student peer assessment in the context of HEIs (Meihami & Razmjoo, 2016; Nortcliffe, 2012).

Conclusion

Based on the major findings of the study and discussion above, the following conclusions can be made:

The study revealed that, though it is theoretically believed student-centered peer-assessment is strongly aligned with the principles of constructivist learning theories (social constructivism-which is the foundation for peer-assessment) and modular instruction; enhances deep learning; and can potentially lead to improved academic achievement and motivation of the students, its level of practicing at the classroom level is *very low* among the sampled universities in Ethiopia.

The finding of the current study also revealed that there is no statistically significant variation in the practice of student-peer assessment of learning by the sample Universities, sample colleges, gender, teaching experiences, and academic ranks of instructors.

The current study also concludes that lack of adequate awareness; knowledge and skills among students and instructors as well as absence of *standard working documents* (SPA policy, guideline) are the major bottle necks to effectively practice SPA among the sampled universities in Ethiopia.

Recommendations

The researcher hopes that the finding of the present study will have both theoretical and practical values for policy designers, curriculum (module) developers as well as the university administration, instructors and students. It will have theoretical values for these key stakeholders to gain adequate insights on the importance of shifting the *curriculum* from traditional to modular, the *assessment* from teacher-centered to student-centered, and the *instruction* from teaching to learning so as to create a window of opportunity for introducing peer-assessment and accommodating students in the assessment of their own work or performance and that of their peers. Such a paradigm shift will unequivocally help the essential stakeholders (education experts, education policy makers, university management, and academic staff) align the assessment activity with the principles of constructivist learning theories and modular instruction.

The finding of the present study will have also practical values for designing appropriate assessment policy or guideline; and in preparing *instructional* (modular courses) and *curricular* (syllabi) materials, in which a detailed procedure on how to implement the student-centered peer-assessment is explicitly explained.

Moreover, the finding of the current study will have tangible values for the instructors to view students as active rather than passive learners as well as learning as a *self-monitored* rather than a *teacher monitored* activity. Not only this, as the concept of modularization is embraced in it the philosophy of developing self-monitoring, self-regulation, self-reflection and self-criticism skills by directly involving students into the design and evaluation of their own work and that of their peers, the result of the present study will have also significant values for the university instructors to design a *student-centered peer-assessment scheme* that consistently fosters these meta-cognitive skills in students. In fact, this requires teachers to possess adequate *pedagogical content knowledge* (including knowledge of peer-assessment) in addition to the domain/subject matter knowledge and knowledge of students' previous learning. In other words, more *professional development* is needed for teachers to become comfortable and knowledgeable about peer-assessment. Of course, this can be enhanced through building the *professional capacity of teachers* (knowledge-base, skills, attitudes and experiences) via need-based training, seminars, and experience-sharing forum. It is strongly believed that teachers' increased knowledge about peer-assessment will help optimize the active participation of students in various authentic assessment methods and feedback as well.

Besides, the present study advises the university administration and academics to *create a classroom culture* that supports the practice of SPA, such as setting a standard for the *class size*, because as the class sizes increase, there is a tendency to utilize examinations and tests more often, compromising what the specific learning outcomes require to be assessed or measured at a time. The present study also advises the assessment policy makers and university academics to make a *balance* between the *continuity of old tradition* (teacher-centered assessment scheme) and *modernity* (student-centered peer-assessment scheme) since such a middle ground enhances a smooth transition; promotes collaboration, sense of partnership, positive relationship or interaction between students and teachers; gradually increases both the teachers and students' trust in the pedagogical values, reliability, and validity of

SPA; increases teachers' opportunity to model the students through demonstration of the assessment tasks; increases students' ability to slowly learn by imitating their teachers as models; and increases students sense of ownership, and self-confidence in the peer-assessment activities.

The current finding also suggests that learners should get adequate awareness on the procedures of peer assessment; be given the opportunity to develop adequate experiences and skills in judging their peers' work as well as in practicing the peer assessment tasks from the initial stages; get explicitly defined peer assessment criteria (such as rubrics, checklists and rating scales) that enable them to assess their peers' work objectively, realistically and fairly; and that teachers should pass evaluative responsibilities to their students through *scaffolding* and *modeling* goal setting, evaluation, good feedback practice, reflection and quality work that reflect competency-based curriculum outcomes.

Moreover, the university management can use the findings of the current study as a corner-step, input or baseline (for it suggests using peer-assessment as a supplement or alternative to the existing teacher-dominated assessment scheme so as to promote self-reflection and self-regulated lifelong learning among the students) to design and implement appropriate interventions (e.g., developing clear policy/guideline on the *student-centered peer-assessment* that reflects the *principles of constructivist learning theories*, that is aligned with *performance-based curriculum* and in line with *modular approach to instruction*.

Finally, the current study encourages future research to extend the scope of this study in terms of the study sites, study variables, conceptual framework and methodological issues so as to ensure more generalizability and capture better insights on the student peer assessment scheme.

References

Andrade, H., Huff, K., and Brooke, G. (2012). *Assessment in the context of Student-centered Learning: The Students at the Center Series*. Jobs for the Future: The Nellie Mae Education Foundation

Andrade, H. L. (2019). A Critical Review of Research on Student Self-Assessment. *Frontier Education, 4*(87): 1-13. Doi: 10.3389/feduc.2019.00087.

Bazeley, P. (2004). Issues in Mixing Qualitative and Quantitative Approaches to Research. In R. Buber, J. Gardner, & L. Richards (Eds.), *Applying Qualitative Methods to Marketing Management Research* (pp.141-156). UK: Palgrave Macmillan.

Bineyam Taye (2014). Online discussion for block teaching in postgraduate health professionals' curriculum: the Ethiopian experience. *BMC Medical Education, 14* (29): 1-6. http://www.biomedcentral.com/1472-6920/14/29

Bozkurt, F. (2020). Teacher Candidates' Views on Self and Peer Assessment as a Tool for Student Development. *Australian Journal of Teacher Education, 45*(1): 47-60. http://dx.doi.org/10.14221/ajte.2020v45n1.4

Brown, G.T.L., and Harris, L.R. (2014). The future of self-assessment in classroom practice: Reframing self-assessment as a core competency. *Frontline Learning Research 3*: 22-30. ISSN 2295-3159

Bryman, A. (2006). Integrating Quantitative and Qualitative Research: How is it done? *Qualitative Research, 6*(1) 97–113. DOI: 10.1177/1468794106058877.

Clark I. (2011). Formative Assessment: Policy, Perspectives and Practice. *Florida Journal of Educational Administration & Policy, 4*(2): 158-180.

Costelloe, L., and Egan, A. (2020). Using self- and peer-assessment to develop Higher Education teachers' skills in assessment and feedback. Higher Education Advances (HEAd'20): 501-508. DOI: http://dx.doi.org/10.4995/HEAd20.2020.11096.

Creswell, J. (2012). Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research (4th ed.). Upper Saddle River, NJ: Pearson Education.

Crowell, T.L. (2015). Student Self Grading: Perception vs. Reality. *American Journal of Educational Research, 3* (4): 450-455.DOI:10.12691/education-3-4-10.

Dann, R. (2014). Assessment as learning: blurring the boundaries of assessment and learning for theory, policy and practice. *Assessment in Education: Principles, Policy & Practice, 21*(2): 149-166. DOI: 10.1080/0969594X.2014.898128.

Donnelly, R., and Fitzmaurice, M. (2005). Designing Modules for Learning. In O'Neill, G., Moore, S., McMullin, B. (Eds.), *Emerging Issues in the Practice of University Learning and Teaching*, Dublin: AISHE

Dochy, F., Segers, M., & Sluijsmans, D. (1999): The use of self-, peer and co-assessment in higher education: A review. *Studies in Higher Education, 24*(3): 331-350.

Double, K. S., McGrane, J. A., & Hopfenbeck, T. N. (2020). The Impact of Peer Assessment on Academic Performance: A Meta-analysis of Control Group Studies. *Educational Psychology Review, 32*: 481–509. https://doi.org/10.1007/s10648-019-09510-3.

El-Koumy, A.S.A.K. (2010). *Student self-assessment in higher education: Alone or plus? Paper accepted for presentation at the CPLA Conference*. Lebanese American University, Lebanon

Forsell, J., Frykedal, K. F., & Chiriac, E. H. (2021). Teachers' perceived challenges in group work assessment. *Cogent Education, 8*(1): 1-16. DOI: 10.1080/2331186X.2021.1886474.

Gizat Mekonnen (2014). *An Assessment of the Practices and Challenges of Modular Master's Program of Addis Ababa University*. MA Thesis.

Greene, J. C. (2008). Is mixed methods social inquiry a distinctive methodology? *Journal of Mixed Methods Research, 2*(1), 7-22.

Gullo, D. F. (2005). *Understanding Assessment and Evaluation in Early Childhood Education* (2nd). New York: Teachers College press.

Hanrahan, S. J., and Isaacs, G. (2001). Assessing Self- and Peer-assessment: the students' views. *Higher Education Research & Development, 20*(1): 53-70.

Heritage M. (2007). Formative Assessment: What Do Teachers Need to Know and Do? *Phi Delta Kappa International, 89*(2):1-8.

Heritage M. (2010). *Formative Assessment and Next-Generation Assessment Systems: Are We Losing an Opportunity?* Paper prepared for the Council of Chief State School Officers: Washington, DC.

Hodges, D., Eames, C., and Coll, R. K. (2014). Theoretical perspectives on assessment in co-operative education placements. *Asia-Pacific Journal of Cooperative Education, Special Issue, 15*(3): 189-207.

Huinker, D., & Freckmann, J. (2009). Linking principles of formative assessment to classroom practice. *Wisconsin Teacher of Mathematics, 60*(2), 6-11.

Ibabe, I., and Jauregizar, J. (2009). Online self-assessment with feedback and meta-cognitive knowledge. *Higher Education*: 1-16. DOI 10.1007/s10734-009-9245-6

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*(7), 14–26.

Juwah, C., Macfarlane-Dick, D., Matthew, B., Nicol, D., Ross, D., and Smith, B. (2004).*Enhancing student learning through effective formative feedback*. The Higher Education Academy.

Karami, A., and Rezaei, A. (2015). An Overview of Peer-Assessment: The Benefits and Importance. *Journal for the Study of English Linguistics, 3*(1): 93-100. ISSN 2329-7034

Kearney, S. (2019). Transforming the first-year experience through self and peer assessment. *Journal of University Teaching & Learning Practice, 16*(5): 3-17. https://doi.org/10.14453/jutlp.v16i5.3.

Lidz, C. S. (2003). *Early Childhood Assessment*. New Jersey: John Wiley & Sons, Inc.

Li, Y., and Chen, L. (2016). Peer- and self-assessment: A Case Study to Improve the Students' Learning Ability. *Journal of Language Teaching and Research, 7*(4): 780-787. DOI: http://dx.doi.org/10.17507/jltr.0704.20.

Long, M. (2000). *The Psychology of Education*. New York: Routledge Falmer

Lysaght, Z. (2015). Assessment for Learning and for Self-Regulation. *The International Journal of Emotional Education, 7* (1): 20-34.

McGarrigle, J.G., and Carlow, I. (2013). What Students Think of Peer Social Constructivist Paradigm. *All Ireland Journal of Teaching and Learning in Higher Education, 5*(2): 1-15.

McMillan, J.H., and Hearn, J. (2008). *Student Self-Assessment: The Key to Stronger Student Motivation and Higher Achievement*. Educational Horizons.

Mc Sweeney, K. (2012). *Assessment for Learning: from Theory to Practice: Final Report.*

Michael Mogessie (2015). Peer-Assessment in Higher Education – 21st Century Practices, Challenges, and the Way Forward. In *Assessment & Evaluation in Higher Education*. Trento, Italy. DOI: 10.1080/02602938.2015.1100711

MoE (2023). *Ethiopian Education and Training Policy*. Unpublished Policy Document.Addis Ababa, Ethiopia.

MoSHE (2020c). *Report on Differentiation Strategies of Ethiopian Public Universities*. Addis Ababa.

Msiza, V., Zondi, T., and Couch, L. (2020). The use of peer assessment at a time of massification: Lecturers' perceptions in a teacher education institution. *Journal of Education, 79*: 37-64. doi: http://dx.doi.org/10.17159/2520-9868/i79a04.

Murtagh, L., and Webster, M. (2010). Scaffolding teaching, learning and assessment in Higher Education. *Tean Journal 1* (2): 1-20 [Online].

Ndoye, A. (2017). Peer / Self-Assessment and Student Learning. *International Journal of Teaching and Learning in Higher Education, 29*(2): 255-269.

Neuman, W.L. (1997). *Social Research Methods: Qualitative and Quantitative Approaches* (3rded). Boston: Allyn and Bacon.

Nicol, D.J., and Macfarlane-Dick, D. (2005). *Formative Assessment and Self-regulated Learning: A Model and Seven Principles of Good Feedback Practice*. Studies in Higher Education.

Nortcliffe, A. (2012). Can students assess themselves and their peers?: A five year study. *Student Engagement and Experience Journal, 1* (2): 1-17.

Olutosin, A. O., and Gabriel, A. (2018). Influence of Two Assessment Techniques on Attitude to and Achievement in Economics among Adolescents in Ilorin South Local Government Area of Kwara State. *International Journal of Educational Research, 5*(1): 33-46.

O'Neill, G. (2015). *Curriculum Design in Higher Education: Theory to Practice*, Dublin: UCD Teaching & Learning. ISBN 9781905254989

Onwuegbuzie, A. J. & Collins, K. M. T. (2007). Typology of mixed methods sampling designs in social science research. *The Qualitative Report, 12*(2), 281-316.

Panadero, E., & Alonso-Tapia, J. (2013). Self-assessment: Theoretical and Practical Connotations. When it Happens, How is it Acquired and what to do to develop it in our Students? *Electronic Journal of Research in Educational Psychology, 11*(2): 551-576. ISSN: 1696-2095, no. 30 - 551 – http://dx.doi.org/10.14204/ejrep.30.12200.

Pound, L. (2011). *Influencing Early Childhood Education: Key Figures, Philosophies and Ideas*. New York: Open University Press

Qu, W., and Yang, S. (2010). A Peer and Self-assessment Project Implemented in Practical Group Work. *Journal of Language Teaching and Research, 1*(6): 776-781. doi:10.4304/jltr.1.6.776-781.

Rahman, M.H. (2015). Learning Assessment in a Self -Learning Material. *International Journal on New Trends in Education and Their Implications, 6* (3), 10:95-101.

Reece, I., and Walker, S. (2003). *Training, Teaching and learning: A practical guide (5thed.)*.Oxford: Business Education Publishers Limited,.

Reynolds, W. M., Miller, G. E., & Weiner. I. B. (2003). *Educational Psychology.* New Jersey: John Wiley & Sons, Inc.

Ross, J.A. (2006). The Reliability, Validity and Utility of Self-Assessment. *Practical Assessment Research & Evaluation, 11* (10): 1-13

Rust C. (2002). The Impact of Assessment on Student Learning: How Can the Research Literature Practically Help to Inform the Development of Departmental Assessment Strategies and Learner-Centered Assessment Practices? *Active Learning in Higher Education, 3(2)*: 145-158. DOI: 10.1177/1469787402003002004

Sardareha, S. A., and Saad, M. R. M. (2012). A Sociocultural Perspective on Assessment for Learning: The Case of a Malaysian Primary School ESL Context. Procedia- Social and Behavioral Sciences, 66: 343 – 353.

Scott, G.W. (2017). Active engagement with assessment and feedback can improve group-work outcomes and boost student confidence. *Higher Education Pedagogies, 2*(1): 1-13. DOI: 10.1080/23752696.2017.1307692

Sendziuk, P. (2010). Sink or Swim? Improving Student Learning through Feedback and Self-Assessment. *International Journal of Teaching and Learning in Higher Education, 22* (3): 320-330.

Shonkoff, J. P., and Phillips, D. A. (2000). *From Neurons to Neighborhoods: The Science of Early Childhood Development*. Washington, D.C: National Academy press.

Siarova, H.; Sternadel, D.; and Mašidlauskaitė, R., (2017). Assessment practices for 21st century learning: review of evidence', NESET II report, Luxembourg: Publications Office of the European Union. Doi: 10.2766/71491.

Siow, L.F. (2015). Students' Perceptions on Self- and Peer-Assessment in Enhancing Learning Experience. *Malaysian Online Journal of Educational Science, 3* (2): 21-35.

Sluijsmans, D., & Prins, F. (2006). A Theoretical Framework for Integrating Peer Assessment in Teacher Education. *Studies in Educational Evaluation, 32*(1), 6-22. https://doi.org/10. 10 16/j.stueduc.2006.01.005.

Spiller, D. (2012). *Assessment Matters: Self-Assessment and Peer Assessment*. New Zealand: Teaching Development unit

Stahl, B. (2003). *The information-constructivist model of school learning: Implications for teaching, instructional design and testing*. Paper presented at the meeting of the American Educational Research Association, Chicago, IL.

Stiggins et al. (2004). *Assess How? Designing Assessments to Do What You Want*. Portland, Oregon: Assessment Training Institute.

Struyven, K., Dochy, F., and Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: A review. *Assessment & Evaluation in Higher Education, 30* (4): 331–347.DOI: 10.1080/0260293042000318091

Thomas, G., Martin, D., and Pleasants, K. (2011). Using self- and peer-assessment to enhance students' future-learning in higher education. *Journal of University Teaching & Learning Practice, 8* (1) 5: 1-17.

Topping, K. (1998). Peer Assessment between Students in Colleges and Universities. *Review of Educational Research, 68*(3): 249- 276.

Trumbull, E., & Lash, A. (2013). *Understanding formative assessment: Insights from learning theory and measurement theory*. San Francisco: West Ed.

Western and Northern Canadian Protocol for Collaboration in Education (WNCP, 2006).*Rethinking Classroom Assessment with Purpose in Mind: Assessment for Learning, Assessment as Learning and Assessment of Learning*.