

**TEACHERS' PRACTICES OF ASSESSMENT FOR LEARNING  
IN SCIENCE EDUCATION AT EAST GOJJAM PREPARATORY  
SCHOOLS, AMHARA REGIONAL STATE, ETHIOPIA**

**by**

**ASKALEMARIAM ADAMU DESSIE**

**Submitted in accordance with the requirements**

**for the degree of**

**DOCTOR OF EDUCATION**

**in the subject**

**DIDACTICS**

**at the**

**UNIVERSITY OF SOUTH AFRICA**

**SUPERVISOR: PROF. PJH HEERALAL**

November 2015

## DECLARATION

Name: Askalemariam Adamu Dessie

Student number: 49059939

Degree: Doctor of Education (DEd)

Title of Dissertation: “TEACHERS’ PRACTICES OF ASSESSMENT FOR LEARNING IN SCIENCE EDUCATION AT EAST GOJJAM PREPARATORY SCHOOLS, AMHARA REGIONAL STATE, ETHIOPIA”

I declare that the above dissertation/thesis is my own work and that all the sources that I have used or quoted have been indicated and acknowledged by means of complete references.



Signature

11/11/2015

Date

## ACKNOWLEDGMENT

I would like to express my deepest gratitude and appreciation to all individuals and organizations for their kind support and tireless guidance throughout this thesis.

First and foremost, I would like to express my heartfelt indebtedness and a deep sense of gratitude to my supervisor, Professor PJH Heeralal, for his professional and exemplary guidance and continuous encouragement throughout my thesis work. His expertise and insightful comments have been inspiring in completing this research. I would like to appreciate him very much. Professor PJH Heeralal is an intelligent, punctual, sympathetic and critical educator. Almighty God blesses him throughout his life.

I would like to express my gratitude towards the member of East Gojjam Education Department and second cycle secondary schools science teachers in East Gojjam Zone for their kind cooperation and encouragement which helped me in completion of this research.

It is my proud privilege to release the feeling of my thankfulness to all my friends and several persons who helped me directly or indirectly to conduct this research. I also thank my parents for their unceasing kind encouragement, support and attention which helped me in completion of this thesis. I will be always indebted to them.

I am also grateful to Dr. Haimanot Wassie, Mr. Minwagaw Temesgen, Dr. Wuhabie Birhan, Dr. Haile kassahun, Dr. Demeke Wollie, Mr. Melkamu Friew, Mrs. Yewulsew Melak, and Mr. Yeshambel Almaw, lecturers in Debre Markos University. My thanks and appreciations also go to Professor Sarah Howie, Director of Centre for Evaluation and Assessment at University of Pretoria and Professor Wynne Harlen, Visiting Professor at the University of Bristol, UK. In general, I am extremely thankful and indebted to those people who shared me their expertise and valuable guidance and encouragement.

## ABSTRACT

Empirical research evidences have confirmed the pedagogical power of formative assessment to improve students' learning, particularly in science education. Thus, this study investigated science teachers' practice of assessment for learning in second cycle secondary schools at East Gojjam Zone, Amhara Regional State. To meet this objective mixed method research design, particularly concurrent mixed method was used. The subjects of the study were all of science teachers in the randomly selected schools. Questionnaire, semi-structured interview, and observation were used to collect the data. From 186 science teachers, 153 teachers properly filled and returned the questionnaire. Moreover, 8 purposively selected teachers were included in the interview and observation sessions. To analyze the quantitative data, frequency, percentage, mean, standard deviation, one-sample t-test, multiple regression, MANOVA, and ANOVA were used. For the qualitative data content analysis was used. The results of the quantitative and qualitative data showed that the practice of assessment for learning in the selected schools was very low. Most science teachers administered tests, home works, assignments, and class works at the end of the lesson to consolidate what they taught and to collect marks, but they did not integrate different assessment for learning methods throughout their instruction for the sake of learning. Teachers mentioned lack of science resources, large class size, shortage of instructional time, inadequate school support, lack of appropriate professional development activities, lack of instructional materials, students' and teachers' negative perception on formative assessment, teachers' lack of knowledge and skill about formative assessment, and large content of courses as major factors for not implementing assessment for learning. Besides, this study revealed a significant relationship between teachers' perception and school supports with teachers' overall practice of assessment for learning. Teaching experience has also significant effect on the combined practice of assessment for learning, particularly teaching experience significantly affects the collection of learning evidences than other factors. However, class size, subject taught, and teaching load per week have no significant effect on the combined practice of assessment for learning. Moreover, the pre-service and in-service assessment trainings have no significant contributions to the practice of assessment for learning. Therefore, comprehensive and relevant assessment trainings should be given for science teachers on a regular basis to integrate assessment with daily instruction to improve learning.

## **KEY WORDS**

- Assessment for learning
- Formative assessment
- Science education
- Preparatory school
- Sharing of learning objectives and success criteria
- Formative feedback
- Peer assessment
- Self-assessment
- Questioning

## TABLE OF CONTENTS

<b>Contents</b>	<b>Pages</b>
CHAPTER ONE.....	1
1. INTRODUCTION.....	1
1.1 Background to the study.....	1
1.2 The research problem.....	3
1.3 Significance of the study.....	7
1.4 Theoretical framework and literature review.....	8
1.4.1 Overview of assessment for learning theories.....	8
1.4.2 Overview of assessment for learning practice.....	11
1.5 Research methodology.....	13
1.5.1 Research design.....	13
1.5.2 Sampling.....	14
1.5.3 Data collection techniques.....	15
1.5.4 Reliability and validity.....	15
1.5.5 Method of data analysis.....	16
1.5.6 Ethical considerations.....	16
1.6 Definition of key concepts.....	17
1.7 Outline of chapters.....	17
1.8 Summary.....	18
CHAPTER TWO.....	19
2. THEORETICAL FRAMEWORK AND REVIEW OF RELATED LITERATURE.....	19
2.1 Introduction.....	19
2.2 Theories underlying assessment practices.....	20
2.2.1 Behaviourism.....	20
2.2.2 Constructivism.....	23
2.3 Putting assessment for learning into practice.....	30
2.3.1 The concept of assessment for learning.....	31
2.3.2 Research evidence on how formative assessment improves learning.....	46
2.3.3 Qualities of new modes of assessment.....	62

2.3.4	Inhibiting factors for effective practice of assessment for learning.....	71
2.4	Conclusion.....	86
2.5	Summary .....	87
CHAPTER THREE .....		88
3.	RESEARCH METHODOLOGY .....	88
3.1	Introduction .....	88
3.2	Research design.....	88
3.2.1	Quantitative approach .....	90
3.2.2	Qualitative approach .....	91
3.3	Sampling.....	92
3.4	Data collection techniques .....	94
3.4.1	Questionnaire .....	94
3.4.2	Observation .....	95
3.4.3	Semi-structured interview.....	96
3.5	Methodological norms.....	97
3.5.1	Validity .....	97
3.5.2	Reliability.....	98
3.6	Pilot study.....	99
3.7	Method of data analysis.....	104
3.7.1	Method of analysis for quantitative data.....	104
3.7.2	Method of analysis for qualitative data.....	110
3.8	Ethical considerations .....	111
3.9	Summary .....	112
CHAPTER FOUR.....		113
4.	RESULTS AND DISCUSSION .....	113
4.1	Introduction .....	113
4.2	Analysis of background data .....	113
4.3	Analysis of the practice of assessment for learning in science education.....	116
4.4	Discussion of Results .....	168
4.4.1	Teachers' practice of assessment for learning in science subjects .....	169

4.4.2	Factors that hinder the implementation of assessment for learning in science subjects .....	184
4.4.3	Teachers' perception on assessment for learning .....	188
4.4.4	Assessment trainings teachers received and its relevance to implement assessment for learning .....	191
4.4.5	School supports provided for teachers to implement assessment for learning .....	194
4.4.6	The relationship between teachers' perception and school supports with practice of assessment for learning.....	196
4.4.7	The impact of training and background characteristics of teachers' on the practice of assessment for learning .....	196
4.5	Summary .....	201
CHAPTER FIVE .....		202
5. FINDINGS, CONCLUSIONS AND RECOMMENDATIONS .....		202
5.1	Introduction .....	202
5.2	Findings and conclusions .....	203
5.3	Contributions of the study .....	214
5.4	Recommendation.....	215
5.4.1	Recommendation for improving assessment for learning practice.....	215
5.4.2	Recommendation for future research.....	218
5.5	Limitation of the study .....	219
5.6	Summary of the study .....	220
REFERENCE.....		223
APPENDIX A: ETHICAL CLEARANCE CERTIFICATE .....		240
APPENDIX B: PERMISSION LETTER TO CONDUCT THE RESEARCH.....		240
APPENDIX C: INFORMED CONSENT LETTER.....		243
APPENDIX D: QUESTIONNAIRE TO TEACHER.....		246
APPENDIX E: INTERVIEW GUIDE FOR TEACHERS .....		253
APPENDIX F: SAMPLE INTERVIEW TRANSCRIPTS .....		254



## LIST OF TABLES

Table 3.1	Number of science teachers in the selected second cycle secondary schools .....	93
Table 3.2	Factor structure of teachers' practice of assessment for learning sub-scale .....	102
Table 3.3	Factor structure of teachers' perception of assessment for learning sub-scale.....	103
Table 3.4	Factor structure of the possible factor sub-scale.....	104
Table 4.1	Respondents' background data .....	114
Table 4.2	Factor grouping of practice of assessment for learning items with factor loadings, mean, and percentage of teachers' response .....	126
Table 4.3	Factor groupings of the possible factor items with factor loadings, mean, and percentage of teachers' response .....	135
Table 4.4	Factor groupings of perception items with factor loadings, mean, and percentage of teachers' response .....	142
Table 4.5	Science Teachers' Professional Learning and its' relevance to implement formative assessment.....	148
Table 4.6	School support provided for science teachers to implement assessment for learning strategies in science subjects in the selected schools.....	150
Table 4.7	The multiple linear regressions model summary and analysis of variance .....	152
Table 4.8	Significance of the predictor variable .....	152
Table 4.9	Inter-correlation among the six dependent variables .....	153
Table 4.10	Mean scores and standard deviation for the combined dependent variables as a function of pre-service training.....	155
Table 4.11	MANOVA for the combined dependent variable as a function of pre-service assessment training .....	155
Table 4.12	Mean scores and standard deviation for the combined dependent variables as a function of in-service training.....	156

Table 4.13 Multivariate analysis of variance for the combined practice as a function of in-service training.....	157
Table 4.14 Mean scores and standard deviation for a combined dependent variable as a function of teaching experience .....	159
Table 4.15 Multivariate analysis of variance for the combined dependent variable by teaching experience .....	159
Table 4.16 Univariate analysis of variance for the six dependent variables across teaching experience .....	160
Table 4.17 Multiple comparison post hoc test (Tukey) for collection of learning evidences ...	161
Table 4.18 Mean scores and standard deviation for the combined dependent variables as a function of class size.....	162
Table 4.19 Multivariate analysis of variance for the combined practice by class size .....	163
Table 4.20 Mean scores and standard deviation for the combined dependent variables by teaching load per week.....	164
Table 4.21 Multivariate analysis of variance for the combined dependent variable across teaching load per week.....	165
Table 4.22 Mean scores and standard deviation for the combined dependent variables by subject taught.....	167
Table 4.23 Multivariate analysis of variance for the combined practice across subject taught.	167

LIST OF FIGURES

*Figure 1: Equal weight concurrent mixed design*.....89

# CHAPTER ONE

## 1. INTRODUCTION

### 1.1 Background to the study

Currently, the education system in Ethiopia is going through reforms and the long term vision of the reform is to transform Ethiopia into a middle income country within 15 years (Eshetu, Dilamo, Tsfaye, & Zinabu, 2009:2; Ministry of Education (MoE), 2010:11). The specific skills, concepts and knowledge students acquire in schools, colleges and universities may no longer be relevant once they are employed in the changing world. Rather, learners should be problem solvers and creative-thinkers to achieve such long-term vision. Thus, developing these lifelong skills is vital to meet the ever changing needs and demands of the globalized society (Centre for Educational Research and Innovation (CERI), 2008:1-3; MoE, 2010: 10-11).

Due to this fact, the Ministry of Education in Ethiopia made reforms in the education system at all levels and placed a great emphasis on the importance of science and mathematics education (MoE, 2010: 10-11). This is because highly qualified and competent professionals in the field of science and technology are vital tools to achieve rapid economic growth in the country (Eshetu et al., 2009:2). Accordingly, the government of Ethiopia recently planned 70% of the higher education intake for science and technology (i.e. 40% of newly enrolling students will be placed in various engineering fields, 30% for the natural science streams), whereas the remaining 30% is left for social sciences and humanities (MoE, 2010: 11). The basic assumption of this strategy is that science and technology is the heart for the economic growth of the country. It is, therefore, critically important to give greater attention for secondary school science and mathematics subjects to realize the strategy and, as a result, to produce scientifically literate citizens who can make the country competitive in the increasingly knowledge-based global economy (Eshetu et al., 2009:2).

However, most secondary schools in Ethiopia did not prepare their students adequately to fulfill the policy intention (Nega, 2009: 2-10; Shibeshi et al., 2009: 200-264; Eshetu et al., 2009: 3-7). As to these writers, most secondary school students have poor performance in science and mathematics subjects. The wash back effect of the national examination and lack of emphases on continuous assessment and student-centered learning methods were mentioned as some of the main reasons of students' high rate of failure in these subjects (Eshetu et al., 2009: 3-4). Mostly, the teaching-learning process and classroom assessments focused on rote learning and memorization of facts (MoE, 2010:19). This was a big challenge to achieve innovation in the education sector, i.e., to develop higher order thinking and problem solving skills.

Therefore, assessment-led educational reform is vital to meet the long-term vision of the country. Many research evidences widely acknowledged the impact of assessment on students' learning, particularly in science education (Black & Wiliam, 1998: 7-21; Stiggins, 2002: 4-7; Irons, 2008: 20-29; Havnes & McDowell, 2008:3-7). In line with this, Irons (2008: 20-32) suggested that the quickest way to change students' learning and achievement is to change the assessment system. Thus, new modes of assessment which intertwined instruction, learning objectives, and real world tasks need to be prioritized in teachers' daily instruction to improve students' achievement and to equip them with higher order thinking and problem-solving abilities.

The Ethiopian education and training policy also envisages the use of continuous assessment at all levels of the education system to enhance students' learning standards (MoE, 1994: 18; MoE, 2005: 30-31). The policy assumption behind continuous assessment is that integrating assessment and instruction will contribute to the quality of learning to achieve the desired objectives.

However, it is difficult and needs time to change the existing assessment practice which has been dominated by standardized tests and summative assessment. This is because changing the mode of assessment needs major paradigm shifts in thinking about learning and teaching. Similarly, Twing, Boyle & Charles (2010: 2) emphasized the close connection between assessment and instruction to improve students' learning. They stated that it needs radical changes to become fully integrated and to become fully formative to improve students' learning quality.

According to Irons (2008:28), in order to make the change successful, students, teachers, and other stakeholders need to have clear understanding about the benefits and strategies of new assessment practices: For this reason, the Ethiopian education system has introduced various in-service training programme for science and mathematics teachers to communicate the change effectively. Some of them are Continuous Professional Development (CPD) and English Language Quality Improvement Programme (ELQIP) (MoE, 2005: 9-10; Eshetu et al., 2009: 4-6). In addition, teacher training institution and school curricula were revised at different levels to include formative assessment (continuous assessment) and student-centered teaching methods in order to improve students' learning. Consequently, teachers were endorsed to treat assessment as an integral part of their daily instruction at all levels both to maintain the quality of learning and to realize the desired competence (Eshetu et al., 2009: 4-6). Therefore, this research tries to assess the actual practice of integrating assessment with daily instructions in science subjects and identify possible factors that hinder such practice in the selected schools.

## **1.2 The research problem**

Ample research evidences indicated that assessment for learning, or formative assessment, or continuous assessment, is a powerful tool to improve students' learning more than summative assessment, or assessment of learning can do, particularly, in science subjects (Stiggins, 2005: 324-325; Stiggins, 2008: 2-3; Assessment Reform Group (ARG), 2002: 2; Harlen, 2005: 208; Harlen, 2003: 9; Black & wiliam, 1998b: 2). On the other hand, researchers such as Black & Wiliam (1998a: 17-20), ARG (1999: 3-5), Stiggins (2005: 324-326), and Heritage (2011:1), identified lack of assessment for learning practice among classroom teachers. From their comprehensive study, they reported that teachers' practices of classroom assessment largely focused on memorization of simple facts, giving of marks, and competition among students rather than on improving learning and instruction. Along with their study results, the main reasons for such poor practices were teachers' lack of knowledge and skills about formative assessment, pressure of standardized tests and summative assessments, and teachers' negative perceptions on their own and students' role in the teaching learning process. Accordingly, Black & Wiliam, (1998a: 17-20) and ARG, (1999: 3-12) strongly suggested the importance of conducting research on teachers' practice of assessment for learning to examine the existing

practice and to identify the problems and, as a result, to take possible interventions for effective implementation of assessment for learning.

However, in Ethiopia currently there is little or no complete research that has been done to examine teachers' practice of assessment for learning in second cycle secondary schools science education and to identify the challenging factors they face for effective implementation. Mostly, previous studies on assessment have mainly focused on summative assessment, continuous evaluation of students' learning, marking, and grading but not on the practice of integrating assessment into daily instruction to improve learning (Yigzaw, 2013:1489-1490). Moreover, many of the existing studies were conducted in higher institutions.

Therefore, this study is initiated to fill these gap in many ways: first, it was conducted in second cycle secondary schools, particularly with a focus on science education; second, it focused on how teachers currently integrate formative assessment strategies into their daily instruction to improve students' learning instead of using it as a tool to collect pieces of marks at the end of each lesson; third, it emphasized on examining assessment for learning strategies and on identifying teachers' awareness and practice; and fourth, it centered on identifying the major factors that hinder the practice of using formative assessment strategies as an active teaching method.

Moreover, policy documents and assessment guidelines that are distributed to schools did not provide effective continuous assessment orientation for learning strategies and did not show best practices or models that indicate how it can be integrated into daily instruction (cf. Amhara Regional Educational Bureau continuous assessment guideline, 2010). From my observation, mostly, school officials ordered teachers to use different continuous or formative assessment tools at the end of lessons or chapters to collect pieces of marks to decrease students' attrition rate rather than encouraging them to use it to adjust their instruction and to improve students' learning.

It was evident that the integration of assessment, teaching, and learning is very important to improve learning, but at the same time it is very difficult to put it in to practice. From my

observation and mini interview conducted in one of the second cycle secondary school, it seems that teachers struggle with this process. Thus, formative assessment or assessment for learning, or continuous assessment is most commonly understood for what it is not. Even though, there is strong evidence of the effectiveness of formative assessment on students' learning improvement, there is still significant gap in understanding the nature of this assessment among stakeholders (Organization for Economic Co-operation and Development (OECD), 2005:7). Even educational experts at different levels and school officials are only speaking about the term 'continuous assessment' without understanding it.

Under such circumstances, curriculum designers, textbook writers, teacher training institutions, school leaders, and teachers all followed their own way in implementing this assessment system. According to Mebea (2008: 3), in Ethiopia, school principals, supervisors, students, and teachers are still in doubt about the effectiveness of continuous assessment to improve students' learning. So, there is a big gap between the government policy intentions and assessment practices at classroom level. However, to improve the quality of education observing and improving the teaching-learning process at classroom level has a great importance. Without understanding what exactly happens in the classroom, all programs that teachers support and other efforts may not adequately prepare teachers to solve the challenges they face in the classroom.

Research is needed on the implementation of effective formative assessment strategies in the classroom because changes in the modes of assessment represent major paradigm shifts in all stakeholders thinking about learning and teaching. According to Libman (2010:63), curriculum change will not be effective to bring quality education without making corresponding changes in assessment.

Therefore, it is essential to study second cycle secondary school science teachers' assessment for learning practice and other factors related to their assessment practice in order to improve students' learning. This study sought to investigate science teachers' practice of assessment for learning and identify hindering factors in the implementation of this new and powerful assessment to improve learning by raising the following research questions:



## **Main research question**

- How do science teachers integrate assessment in their pedagogical practices and what are the major factors in their choice of assessment strategies?

## **Sub-research questions**

- To what extent do teachers integrate assessment in their instruction to improve students' science learning?
- What are the major factors that affect teachers' practice of assessment for learning in science subjects?
- How do second cycle secondary school science teachers' perceive assessment for learning?
- To what extent do professional (pre-service and in-service) trainings help science teachers to integrate assessment with their daily instruction?
- What type of support schools provide for the effective practices of assessment for learning in science subjects?
- Are there significant relationships between teachers' perception and school supports with teachers' overall practice of assessment for learning?
- How do the training and background characteristics (teaching experience, teaching subject, work load per week and class size) of teachers' affect their practice of assessment for learning?

### 1.3 Significance of the study

Even if there is a strong assertion on the assessment reform in Ethiopia, the way in which the new paradigm (i.e. assessment for learning) put into practice on regular bases to improve students' learning seems challenging. It needs great effort from policy makers, school administrators, researchers and practitioners (teachers and students) themselves. Thus, the current study has its own practical, theoretical, and policy wise contributions for different stakeholders at different levels.

Practical contribution:

- This study can help science teachers to identify the gaps in their classroom practice. As a result, it may be used as a guideline to improve their knowledge and skill to integrate formative assessment strategies into their daily instruction for the purpose of learning.
- The findings of this study are useful for school leaders to engage in continuous discussion with teachers regarding the practice of formative assessment to improve students' learning, to encourage teachers to collaborate and conduct action research, and to promote professional trainings centered on formative assessment.
- The outcomes of this study can also help policy makers and teacher training institutions to evaluate the relevance of pre-service and in-service assessment trainings in order to integrate formative assessment into daily instruction. As a result, the study can serve as a framework for developing professional trainings which focus on assessment for learning.

Theoretical contribution:

- The findings of this study can add a new knowledge to the exiting assessment theory and practice, particularly, in the context of Ethiopia. As mentioned above, most studies in Ethiopia mainly focused on the summative value of formative assessment and on grading, but not on the learning value of formative assessment or on the integration of assessment and instruction. Therefore, it is hoped that this study can add a new dimension to the theory and practice of integrating different formative assessment strategies into daily

instruction to improve students' learning. For instance, on the power of formative assessment on learning, effective strategies of formative assessment, teachers' perception, and other possible factors of integrating formative assessment into daily instructions are elements to be considered.

- The findings are also informative for policy makers, educational experts, school leaders, teachers, and other concerned bodies about the current practice of assessment for learning in the selected second cycle secondary schools and the importance it has to science teachers to improve students' learning. Professional trainings on formative assessment can be effective if the designers of the training program and other stakeholders know what teachers are already doing and what they are not doing well. This study, therefore, has an informative value in this regard.
- Finally, this study will serve as a bridge for further research in the area of assessment for learning or formative assessment.

## **1.4 Theoretical framework and literature review**

### **1.4.1 Overview of assessment for learning theories**

Like other approaches in the education system, assessment for learning (the new paradigm) has its own theoretical and philosophical backgrounds. Over the past centuries, different learning theorists carried out a debate to describe how individuals learn things. In line with the debate on how peoples learn, learning theorists also put their own views on learning assessment. Thus, this section provides a brief overview of behaviourism and constructivism perspectives on learning and assessment procedures, techniques, and principles respectively. A detail exploration of the theoretical framework of assessment for learning is discussed in chapter two.

Along with behaviourism, learning was viewed as a stimulus and response association (Shepard, 2000: 9-10). Due to this fact, behavioural leaning theorists focus on the observable behaviour of the learners to describe how learning is occurred rather than on the covert or mental processes.

Thus, learning is conceived as a process of changing the observable behaviour through selective reinforcement of learner's response to events that occur in the environment (Peel, 2005:20-21). Here, the external environment is the determinant factor in learning to change the behaviour of the learner toward the desired outcomes. As a result, the effectiveness of the teaching-learning process was judged based on the changes in the observable behaviour of the learner.

Similarly, in behaviourist assessment was viewed as tests and exams that teachers administered at the end of lesson to evaluate the changes in students' observable behaviour. And the change was interpreted in line with the correct and incorrect answers students' got in the exam. As a result, the role of the teachers is to teach learners how to perform better on a test or exam (Shepard, 1991:2) and students are working hard to listen and read accurately to memorize and repeat concepts and formulas to score high on exams. Thus, the instruction relies on the transmissions of knowledge which is largely passive, teacher-directed, and controlled (Richardson, 2003: 6; Hackmann, 2004: 697; Taber, 2011: 40-41) and on the memorization of concepts, principles, and formulas for the sake of exams (Taber, 2011: 42).

So, the purpose of assessment in the behaviourist paradigm is to evaluate learning outcomes, to select, rank, and to give grades accordingly, but not for internal improvement (Shepard, 1991: 2-14; Anderson, 1998: 6-8). For this reason, assessment techniques such as true-false, short answer, matching, multiple-choice, and restricted essay questions are preferred in this perspective (Anderson 1998: 8; Shepard, 2000:7; Fautley & Savage, 2008:18). Yet, such objective tests are inadequate to reveal the student's higher-order thinking, understanding, and problem solving skills (Shepard, 1991: 2; Anderson, 1998:6; Scholtz, 2007: 43) especially in science subjects. Generally, for behavioral learning theorists, "assessment and instruction are viewed as two separate entities" (Anderson, 1998: 9; Shepard, 2000:5; Adair-Hauck et al., 2006:363).

However, constructivism, which was the current learning perspective, emerges and directs our thinking toward the constructive alignment of assessment, learning, and instruction. According to constructivists' view, learning is an active mental process of constructing new ideas and fitting this new ideas into the existing knowledge or schema (Shepard, 2000: 6-7; Hackmann, 2004:

697; Underhill, 2006:168; Improving Quality of Science Teacher Training in European Cooperation (IQST), 2008: 2; Singh et al., 2012: 198). Here, learners learn to learn by constructing meaning and their own systems of meaning (Richardson, 2003: 1628-1630). Unlike behaviourist view, constructivists' belief that learning requires active engagement of learners and determined by what goes on in learners' heads. As a result, the role of the teachers and students in the instruction becomes changed and formative assessment emerges as an integral part of their pedagogic practice. For this new perspective of learning, assessment should be integrated in the teaching-learning process to improve students' learning (Graue, 1993: 291). Thus, students are viewed as active participants of classroom assessment and instructions as opposed to the old paradigm of assessment in which the students were passive receivers of information from teachers or text books (Anderson, 1998: 5; ARG 1999:2; Shepard, 2000:6; Chappuis & Stiggins, 2002:1-2; Singh et al., 2012:198).

This new paradigm of learning comprises cognitive and social constructivism learning theories (Heritage, 2010: 7). Cognitive constructivist perspective focuses on what is going on in the mind of the learner while social constructivism focuses on the social aspects of learning. In cognitive constructivism, learning is almost completely a function of the learner's interpretation of events in the environment (Swan, 2005:3; Heritage, 2010: 7-9). Similarly, cognitive constructivist emphasizes on the importance of providing ongoing meaningful feedback to learners and the integration of self-assessment to the learning process (Underhill, 2006:168). Therefore, for cognitive constructivism, good assessment practice gives emphasis to learning with understanding and application, not on memorization of facts (Swan, 2005:8).

Within the social constructivist perspective, which grows out of the work of Vygotsky, learning is viewed as an interaction between the individual and the social environment, where "the learner both shapes and shaped by the community of practice" (James, 2006:57). This perspective focuses on the nature of social interactions and how these interactions influence learning (Richardson, 2003: 1625; Bennett & Gitomer, 2008: 7).

Accordingly, this theory suggested that teacher needs to create conducive environment to stimulate learners to think and act in authentic tasks beyond their current level of competence

(Heritage, 2010: 7-9). Thus, tasks need to be collaborative and students must be involved in the generation of problems and solutions (Shepard, 2000:12). Teachers and students jointly solve problems and all develop skill and understanding (Swan, 2005: 2; Heritage, 2010: 8). Like cognitive constructivist perspective, a social perspective of learning is a powerful explanatory theoretical framework for teachers seeking to engage in assessment for learning practices to take into account the role of interaction and joint collective action in the teaching-learning process (Heritage, 2010: 8).

Generally, the supporters of constructivist perspective have done a lot of scientific research work to make classroom assessment practice effective. And they proposed that assessment should be aligned to instruction and its purpose is to inform and improve learning, rather than simply to assess whether the learners have mastered the learning objectives. As stated by constructivist, it is, therefore, necessary to provoke students' mental models through dialogue, higher-order questioning, peer and self-assessment, open-ended assignments, thinking, and concept-mapping to scaffold them to apply concepts and strategies in novel situations (ARG 1999: 7-8; Black et al., 2004:11). A detailed explanation of the theoretical framework is found in chapter two.

#### **1.4.2 Overview of assessment for learning practice**

Many research evidences confirmed the effectiveness of assessment for learning to improve students' learning, particularly in science education. On the word of Black & Wiliam (1998), "substantial learning gains are possible when teachers introduce assessment for learning into their classroom practice." (Wiliam, Lee, Harrison, & Black, 2004: 49). Moreover, formative assessment or assessment for learning is vital to increase students' internal motivation to learn and to develop positive self-esteem, confidence, and self-independence among students (Nicol and Macfarlane-Dick, 2006: 199-208; Irons, 2008: 27-50). Thus, assessment for learning is a crucial component of effective instruction to achieve higher order learning outcomes or standards (Dunn & Mulvenon, 2009:9) and to develop lifelong learners (Fautley & Savage, 2008:91). However, teachers need to understand it well and integrate it with their daily instruction for the purpose of learning.

Accordingly, current researchers identified basic elements or strategies of assessment for learning to integrate it with daily instruction to improve learning. These comprises: sharing of learning target and success criteria with students, involving students in peer and self- assessment, providing effective written and oral feedback that helps students to recognize next steps in learning, and using effective questioning and dialogue throughout the instruction (Black & Wiliam, 1998: 17; ARG, 2002: 7; Lee, 2006:44; Council for the Curriculum, Examinations and Assessment (CCEA), 2007: 2; Department for Children, Schools and Families (DCSF), 2010:6).

Harlen (2003: 20) also added the phases in formative assessment or assessment for learning to integrate it effectively into daily instruction to improve students' learning. These are: identify learning objectives and success criteria, collect learning evidences in line with the learning objectives, interpret the evidences against learning objectives or identify the learning gaps, decide the next steps or proposed strategies to fill the gaps, and take actions to fill the gaps or use evidences for future learning. According to her, students become the center of any activity in the cycle of assessment for learning to achieve the learning objectives identified.

Thus, implementing the basic strategies and follow the cyclic nature of formative assessment is important to improve learning standards in science subjects through formative assessment. Moreover, the assessment tasks designed need to be integrated with daily instruction, authentic or practical or integrated with learning objectives and real world tasks, and continuous (King, 2006:33. Kvale, 2007:66).

However, assessment for learning or formative assessment was not practiced or integrated into daily instruction and real world tasks effectively to achieve the underlying objectives, particularly in science education (Lambert & Lines, 2001:133-135; Black & Wiliam, 2001: 4-6; Ottevanger, Akker, & Feiter, 2007: 19; Kvale, 2007: 65; Black & Wiliam, 1998 as cited in Fautley & Savage, 2008:92; CERI, 2008:3; Swaffield, 2009:9; Fisseha, 2010: 103-105). According to these researchers, most teachers predominantly focus on tests, marks, grades, sorting and ranking of students, provision of judgmental feedback, and on norm referenced interpretation of results. Moreover, teachers work with an insufficient picture of students' learning needs and give less attention for students' previous performance. Therefore, teachers

give low attention for the learning value of formative assessment in their practice; rather they predominantly focus on the traditional perspective of assessment.

With this regard, research evidences pointed out the possible reasons for such failures of the value of formative assessment practices in the classroom. The main barriers in teachers' practice of assessment for learning are: teachers' perception of science learning and assessment (Harlen, 2003:11; Graue & Smith, 2004:25; Stobart, 2008: 168; Mansell, James, & ARG, 2009: 18-19; Carless, 2005 cited in Hill, 2011: 348); the pressure of summative tests or exams or standardized tests (CERI, 2005: 24; Dysthe, 2008: 15; Falchikov & Boud, 2008: 92-93; CERI, 2008: 3; Mansell, James, & ARG, 2009: 18-19; Carless, 2005 cited in Hill, 2011: 348); fear of resource and time pressure (Harlen, 2003:11; Graue & Smith, 2004:25); lack of professional development opportunities which focus on formative assessment during pre-service and in-service trainings (Harlen, 2003:11; Ottevanger, Akker, & Feiter, 2007: 19; Mansell, James, & ARG, 2009: 18-19); lack of intensive supports from school leaders and educational experts at different level (Harlen, 2003:11; Graue & Smith, 2004:25; Mansell, James, & ARG, 2009: 18-19; Carless, 2005 cited in Hill, 2011: 348); and students' perception and past experience on assessment (Falchikov & Boud, 2008:15). This is not exception in Ethiopia. Thus, further research works are vital to examine the current practice of assessment for learning and to identify the possible factors for such practices to achieve the intended objectives through formative assessment. Detailed descriptions of assessment for learning strategies, power of formative assessment on students' learning, qualities of formative assessment, and barriers in teachers' practice of assessment for learning were presented in chapter two.

## **1.5 Research methodology**

### **1.5.1 Research design**

The main purpose of this study was to examine science teachers' current practice of integrating assessment for learning with their daily instruction to improve learning and identify major factors of their assessment for learning practice. To address the research questions mentioned in section 1.2, mixed methods design (Creswell, 2009:19) was applied. From different mixed



method design concurrent triangulation was selected. It is one of the most widely used mixed method designs in social and human science researches (De Vos, Strydom, Fouche, & Delport, 2011: 442; Creswell, 2009: 248). According to Creswell (2009:13-14), a mixed methods approach is one in which the researcher tends to base knowledge claims on pragmatic grounds. As pragmatic knowledge claims using consequence-oriented, problem-centered, and pluralistic approaches to derive knowledge about the problem were the most important. "Mixed methods research eliminates different kinds of bias, explains the true nature of a phenomenon under investigation and improves various forms of validity or quality criteria" (De Vos et al., 2011: 436). "A mixed methods design is useful to capture the best of both quantitative and qualitative approaches" (Creswell, 2009: 24). Therefore, the researcher employed both quantitative and qualitative methods to cross-validate, confirm, and verify the findings within the study. A more detailed exploration of the research design is presented in chapter three.

### **1.5.2 Sampling**

In East Gojam Zone there are 18 second cycle secondary schools. The second cycle of secondary education covers grade 11 and 12 and prepares students for continuing their studies at the higher education level or selecting their own vocations (MoE, 1994:15). Therefore, the population of this study was science subject (Mathematics, Biology, Chemistry and Physics) teachers in these second cycle secondary schools.

Then, to select appropriate sample size, the schools were clustered in to two based on their year of establishment. Those schools 'older' than ten years were clustered in one group (cluster one) and those schools 'younger' than ten years were clustered in the second group (cluster two). Then, because of time and financial constraint, five schools were selected randomly from each cluster. Both grade levels (grade 11 and 12) were included in the study. For quantitative section, both science teachers in the selected schools were included to fill the questionnaire. Totally, there are 186 science teachers in the selected schools. For qualitative section, the participants were selected purposively from the nearby four schools by considering their teaching experience, educational background, and subject taught until data saturation. Those individuals who were

included in the interview were also included in the observation sessions (see the details in chapter three).

### **1.5.3 Data collection techniques**

To answer the research questions both quantitative and qualitative data were collected. The multiple data sources allow for triangulation of data to reduce bias and at the same time to develop a deeper understanding of the issues under study (De Vos et al., 2011: 436; Creemers, Kyriakides & Sammons, 2010: 116-120; Creswell, 2009: 240- 248). For quantitative section, survey or self-administered questionnaire was used whereas, for qualitative aspect, semi-structured interview and observation were implemented to address the research question mentioned in section 1.2 (see details in chapter three).

### **1.5.4 Reliability and validity**

De Vos et al., (2011: 177) stated that increasing the number of items in the questionnaire and interview or number of observations conducted will raise the reliability of the instrument. Accordingly, pilot study was conducted in this study to ensure the reliability of the instrument. And then Cronbach alpha and split-half reliability estimation methods were implemented to analyze the reliability of items in the pilot and final questionnaire. Moreover, Creswell (2007: 209) added that for qualitative approach “reliability can be enhanced if the researcher obtains detailed field notes by employing good quality tape for recording and by transcribing the tape”. This is also done accordingly.

The questions in the questionnaire and interview were reviewed by peers and selected experts in the area to establish the validity of the question. Content validity and face validity of the instrument can be established on the bases of researchers’ and other experts’ judgment (De Vos et al., 2011:173). Prolonged engagement on-site, persistent (consistent) observation, triangulation and peer debriefing are important strategies to prove validity in qualitative methods (Ridenour and Newman, 2008: 57-58). Likewise, Creswell (2007: 208-210) described important opinions to increase the validity of qualitative methods. These are: peer review, prolonged engagement and

persistent observation, use of multiple and different sources, methods, investigators and theories (i.e. triangulation), and involvement of respondents to view the credibility of findings and interpretations. For this reason, peers review, triangulation, and two phase observation were employed in this study. More detail is provided in chapter three.

### **1.5.5 Method of data analysis**

For quantitative part, descriptive statistics such as frequency, percentage, mean and standard deviation were employed to give general description about teachers' background, perception, practice, professional trainings given for teachers' and its' relevance, school supports provided, and identify possible factors. Moreover, multiple linear regression was used to analyze the relationship between teachers' perception and school supports with teachers' overall practice of assessment for learning. Besides, separate multivariate analyses of variances were employed to assess the effect of teachers' training and background characteristics on the overall practice of assessment for learning. If significant effects are observed, univariate tests are also used to assess the effect on each factor loadings. The difference is tested at  $\alpha = 0.5$ . For qualitative part, the semi-structured interview questions and observation were analyzed qualitatively using words. A detail exploration is found in chapter three.

### **1.5.6 Ethical considerations**

According to Cavan, (1997:810) ethics has been defined as 'a matter of principled sensitivity to the rights of others, and that 'while truth is good, respect human dignity is better' (as coted in Cohen, Manion, & Morrison, 2007: 58). Thus, research ethics are principles that guide the researcher how to treat participants during the research process. In this study, the researcher considered ethical issues during data collection, interpretation, and dissemination. For this reason, the researcher developed an informed consent form through incorporating the purpose of the study, participants' role in the study, how participants are selected, their right to ask questions, to participate or withdraw at any time, and their privacy in the study. Moreover, in the

interview and observation part the researcher used coding for names and aliases or pseudonyms for individuals and places to protect identities.

## 1.6 Definition of key concepts

- **Assessment:** It is the ongoing process of gathering, collecting, reflecting on, and using evidence about student progress in order to support and enhance their future learning.
- **Formative assessment or assessment for learning or continuous assessment:** It is an ongoing process in which teachers and students gather evidence throughout the instruction about students' learning to identify the gaps for future learning.
- **Summative assessment (assessment of learning):** It is the attempt to summarize or evaluate student learning at some point in time, say at the end of a lesson, chapter or course.
- **Feedback:** It is a comment or information that learners receive from a teacher, from other learners, or from themselves upon reflection, on the product of a learning task.
- **Learning targets:** It refers to the learning objectives that the learners achieve at the end of the subject
- **Success criteria:** It refers to the assessment criteria or standard that the learners need to realize after each assessment tasks. Success criteria are formulated based on the learning objectives or goals.
- **Science subjects:** Science in this study includes Mathematics, Physics, Chemistry, and Biology subjects in grade 11 and 12.

## 1.7 Outline of chapters

The first chapter includes introduction, research problem, research questions, significance of the study, research methodology including research design, sampling, data collection techniques, methodological norms, method of data analysis, ethical issues, definition of key concepts, and Outline of chapters.

Chapter two presents the theoretical framework of the study and review of related literature. This includes the underlying assessment procedures from the standpoint of behaviorists and constructivists learning theories. This chapter also includes concept of assessment for learning practices, research evidences on the effectiveness of formative assessment to improve students' learning and learning standards with special reference to science and mathematics education in secondary schools. Moreover, it includes the challenging factors of science teachers' practice of assessment for learning and features of effective assessment for learning.

Chapter three presents the research design and methods used to conduct the investigation. Besides, population, sample size, sampling technique, data collection methods, methods of data analysis, methodological norms (reliability and validity of instruments), pilot study, and ethical issues are described in detail.

Chapter four consists of the results of both the quantitative and qualitative data and discussion of results in relation to the basic research questions.

Chapter five provides the conclusions, contributions, recommendations, limitations, and summary of findings.

## **1.8 Summary**

In chapter one the background of the study, research problem, research questions, significance of the study, research methodology including research design, sampling, data collection techniques, methodological norms, method of data analysis, ethical issues, definition of key concepts, and outline of chapters were explained. The next chapter will present the theoretical framework that guide assessment for learning and review of related literature on the same issue.

## **CHAPTER TWO**

### **2. THEORETICAL FRAMEWORK AND REVIEW OF RELATED LITERATURE**

#### **2.1 Introduction**

In the past, scholars focused on the notion of overt behavior and stimulus response association for the acquisition of knowledge or learning to occur, but now the focus was shifted away from stimulus response learning towards an active process of mental construction and sense making (Shepard, 2000:4-9). This historical paradigm shift from behavioural revolution to cognitive revolution is also associated with the current innovation in the educational assessment (Heritage, 2010: 8; Cowie, Moreland & Otrell-Cass, 2013:11). Hereafter, a paradigm shift occurs in the educational assessment from assessment of learning or summative assessment to assessment for learning or formative assessment. In contrast to summative assessment, formative assessment or assessment for learning emphasizes on the learning process rather than the end product and primarily focus on improving students' learning. Assessment for learning engages students actively in their own learning through effective questioning, dialogue, descriptive feedback and peer and self-assessment. Unlike summative assessment, in formative assessment instruction and assessment is integrated and the assessment activities are aligned with the real life situations.

For this reason, the intention of this chapter is to present a review of the literature on the theoretical framework of assessment for learning or formative assessment strategies, procedures and principles. Like other approaches in the education system, assessment for learning has its own theoretical and philosophical backgrounds. Furthermore, this chapter explores a review of related literature on the effective strategies to put assessment for learning into practice, the empirical research evidences of how formative assessment improves students' learning, particularly in science education, qualities of new modes of assessment or formative assessment, and the inhibiting factors for effective implementation of formative assessment in secondary school science subjects.

## **2.2 Theories underlying assessment practices**

As literature indicates, over the past century, learning theorists have carried out a debate on how do individuals acquire, organize and deploy skills and knowledge. Particularly, how do people think and learn has been viewed as one of the most powerful questions of science education since the 20th century (Cizek, 1997: 6-7; Shepard, 2000: 4-9; Cowie, Moreland, & Otrell-Cass, 2013: 6-13). For this reason, there is a significant change in science education, specifically in assessment practices today. In line with the debate on how people learn; learning theorists put their own views on assessment (Swaffield, 2009: 4). By supporting this idea Cowie, Moreland, & Otrell-Cass (2013: 6) stated that “all assessment is underpinned by a view of learning.” Therefore, assessment may be practiced in many different ways depending on our views of learning. Thus, under here, we discuss the implication of behaviourism and constructivism perspectives on science learning, learners, instruction and assessment and the integration between them.

### **2.2.1 Behaviourism**

The origins of behavioural learning theories may be traced back to the late 18th and early 20th century with associative learning principle (Cizek, 1997: 6-7; Bennett & Gitomer, 2009: 59). Behaviorism is comprised of individual learning theories that have a shared common assumption within them (Peel, 2005:20-21). This traditional view of learning most popularly associated with the work of Pavlov, Watson, Skinner and Thorndike (James, 2006: 54). All of them emphasize on the observable behavior and the power of external environment to determine what is learned rather than the individual. For them, the learning environment is the determinant factor for learning to occur (James, 2006: 54). Behaviourists view that learning is brought through stimulus and response association (Shepard, 2000: 9-10) and the complex associations between response and stimulus are built up gradually through the teaching inputs (Underhill, 2006:166).

Furthermore, behavioural learning theory focuses on the observable behaviour rather than the covert or mental processes of the learners (Stobart, 2008: 161), because the main feature of behaviourists view of learning is that learning should be considered in terms of evidence of behaviours or actions. The divorcing of the mental process from the behaviour is a key feature of

learning in behaviourist view (Fautley & Savage, 2008:16). Similarly, James (2006: 54) pointed out the ignorance of the concept of mind and intelligence to explain learning in behavioural learning theory. Thus, learning is conceived as a process of changing observable behavior as result of selective reinforcement and punishment of an individual's response to events that occur in the environment (Peel, 2005:20-21; James, 2006: 54; Race, 2007:1). Accordingly, the effectiveness of the instruction was judged based on the changes in the observable or overt behavior of learners.

For the behaviourist, when the individual is born, mind is seen as an empty vessel or 'tabula rasa' to be filled by parents or teachers (Cizek, 1997: 6-7; Anderson, 1998: 8; Taber, 2011: 46). Therefore, for them, learners are considered as passive receiver, and knowledge is viewed as stable, objective and external (Taber, 2011: 40-41). Generally, according to behavioral learning theorists man is little more than an elaborative machine, whose behavior is largely determined by his/her environment (teachers) (Shepard, 2000: 6). Hence, it relies on a transmission of knowledge which is largely passive, teacher-directed and controlled (Richardson, 2003: 6; Hackmann, 2004: 697; Taber, 2011: 40-41). Within a behavioural learning view, complex tasks should be broken down in to smaller fragments to be mastered sequentially (Bennett, & Gitomer, 2009:59; Fautley & Savage, 2008:16), because learning means accumulation of specific facts and memorization of these facts sequentially.

From behavioural perspective, knowledge is a hierarchical set of discrete package which teachers can teach sequentially and which learners learn progressively more complex ideas (Bell & Cowie, 2002:18). Therefore, the role of the teacher is to state objectives in relation to students' age, transmit it and measure the objectivity and clarity of transmission using valid tests of content mastery (Anderson, 1998: 8; Adair-Hauck et al., 2006:363).

Then, the assessment method in the behavioural perspective most widely comprises traditional forms of assessment which are unidirectional, semester or year-based, and paper and pencil exams (Fautley & Savage, 2008: 16-18). Here assessment is viewed as a way of estimating how much objectives a student has acquired and make judgment about the change in the observable behaviours of the student (Dysthe, 2008:19).



Here, teachers teach to make the learners performance better on a test or exam (Shepard, 1991:2). From this perspective, students' achievement is associated with the accumulation of skills and discrete knowledge and the memorization of facts in a given subject. Therefore, for behaviourism, teaching focuses on rote learning and assessment on the memorization of concepts, principles and formulas for the sake of exams (Taber, 2011: 42). Accordingly, the role of the students becomes working hard to listen and read accurately to memorize and repeat specific scientific concepts and formulas in science subjects, and teachers' role is to train learners to respond correctly and rapidly to the instruction.

Within a behaviourist paradigm, assessment focuses on measuring outcomes, grading and marking (Shepard, 1991: 2; Anderson, 1998: 6). In this perspective, the purpose of assessment is to select, rank and control, but not for internal modification (Shepard, 1991: 14; Anderson, 1998:8). Likewise, Bell & Cowie (2002: 19) stated that in this view formative assessment is used only to check whether a student has achieved the learning objectives or not and to provide judgmental feedback accordingly (Underhill, 2006:166; Scholtz, 2007: 43).

As behaviourists students' progress was judged based on the correct and incorrect answers they got on the paper and pencil tests. Moreover, their "poor performance was remedied by extra practice in the incorrect items, sometimes by deconstructing them further and going back to even more basic skills" (James, 2006: 55), because as behaviourist repetition is seen as one of the greatest means to ensure mastery of objectives (Shepard, 2000:5; Taber, 2011: 41). Therefore, assessment techniques in behaviourist approaches comprise of objective tests such as true-false, short answer, matching, multiple-choice and restricted essay questions (Anderson 1998: 8; Shepard, 2000:7; Fautley & Savage, 2008:18). Yet, such objective tests are inadequate to reveal the student's higher-order thinking, understanding, and problem solving skills especially in science subjects (Shepard, 1991: 2; Anderson, 1998:6; Scholtz, 2007: 43). If teachers want students to do more than memorization of facts and concepts, they need to assess students in various ways using multiple tools.

Overall, as behavioural learning theorists, “assessment and instruction are viewed as two separate entities” (Anderson, 1998: 9; Shepard, 2000:5; Adair-Hauck et al., 2006:363). However, the question of the integration between instruction and assessment and comprehensive ways of assessing were answered by constructivists’ views of learning that build on the work of Piaget, Vygotsky, and other cognitive theorist (Graue, 1993: 285-288; Cizek, 1997: 6-7; Anderson, 1998: 7; Lambert & Lines, 2001: 127; Hackmann, 2004: 697; Roos, 2004: 5; James, 2006: 53; Scholtz, 2007: 44; Bennett & Gitomer, 2009:58).

### **2.2.2 Constructivism**

The term constructivism got its meaning at the first time by Jean Piaget in 1929, in his book “The construction of Reality in the Child” (Sjoberg, 2007: 4). Then constructivism has grown by the work of Lev Vygotsky who put more stress on the social and cultural aspects of learning (James, 2006: 56; Sjoberg, 2007: 6).

In contrast to behavioural learning theory, constructivism holds that “meaning is constructed in our minds as we interact with the physical, social and mental worlds” (Swan, 2005:1). For them, all learning is an active process and tied with the individual experience (existing idea) about the world (Shepard, 2000: 6; Stobart, 2008: 151-161; Cowie, Moreland & Otrell-Cass, 2013: 11). Since, knowledge is not passively received from the outside environment rather than actively constructed by the learner in the constructivist view of learning (Sjoberg, 2007: 4; Swan, 2005:2; Fautley & Savage, 2008:19). Hence, the teaching learning process should consider the existing experience of learners about the phenomena.

Currently, constructivism is a widely accepted learning theory in the teaching learning process (Anderson, 1998: 7; Shepard, 2000: 6-7; Shepard, 2005: 66; James, 2006: 55). These current views of learning broadly refer to the multiple fields of inquiry concerned with how students learn. According to constructivists’ views of learning, the instruction should emphasis on understanding, problem solving skills, conceptual development and the enhancement of students’ abilities to recognize and use meaningful patterns of information (Shepard, 2000: 7; Shepard; 2005:7). Constructivism learning theories have many implication for instruction because it

emphasis on student-centred, knowledge-centred, assessment-centred and community-centered learning environments (Swan, 2005:6-9).

Unlike the behaviourist view, learning, in the constructivist view requires the active involvement of students to construct and build up new knowledge for themselves from what they already know (Shepard, 2000: 6-7; Hackmann, 2004: 697; Underhill, 2006:168; Stobart, 2008: 161; Singh et al., 2012: 198; Cowie, Moeland, & Otriel-Cass, 2013: 11). Thus, students fit new ideas into their already existing knowledge and experience (Shepard, 2000: 10; Sjoberg, 2007: 4; Heritage, 2010: 7; Improving Quality of Science Teacher Training in European Cooperation (IQST), 2008: 2). Generally, in the constructivist view, learners are active meaning makers for themselves (Richardson, 2003: 1628-1630; Swan, 2005: 6).

Furthermore, the increasing focus on the application of science skills in the recent days is closely aligned with the emerging research on the theory of constructivism. Since, Piaget, who was the originator of constructivism, his theory was discovered by science educators. His theory had durable influence on science and mathematics education and turned it into real science education (Sjoberg, 2007: 5-6). Therefore, constructivism learning theory becomes the basis for science and mathematics education (IQST, 2008: 2). For the reason that, constructivism is a dominant and a powerful theoretical perspective in the field of science education, particularly on the assessment practice today (Sjoberg, 2007: 6; Bell & Cowie, 2002:114-115).

This new perspective has changed our thinking and understanding both for the instruction and assessment. In contrast to the traditional theory, here instruction focuses on conceptual understanding rather than content coverage and rote learning. Now, we understand that assessment is an integral part of the instruction (Graue, 1993: 291; Shepard, 2000: 8), a continuous process throughout the instruction (Heritage, 2010:8; Bell & Cowie, 2002: 6-23), and aligned with classroom objectives and real life problems (Graue, 1993: 292; Gulikers, Bastiaens & Kirschner, 2008:73-77; Havnes, 2008: 108). For them, meaningful assessment focuses on the learner's entire conceptual network, not just on discreet facts and principles (Shepard, 2000:8). Assessment processes, then, should enable students to demonstrate deep understanding of concepts and give continuous evidence of their understanding and thinking for teachers and for

themselves to close the gap between the existing understanding and the new understandings (Swan, 2005:8).

As a result, formative assessment (assessment for learning) appears as an integral part of teachers' pedagogic practice in constructivists approach (Graue, 1993: 281; Shepard, 2000: 6; Stobart, 2008: 161; Cowie, 2012:679-680; Cowie, Moeland, & Otrell-Cass, 2013: 11), and the role of teachers and students became changed (Heritage, 2010: 8). The change was occurred as a result of the beneficial power of assessment for learning on students learning. For constructivist, therefore, it is necessary to provoke students' prior knowledge or schema through effective dialogue, questioning, self and peer- assessment, open-ended assignments, thinking and concept-mapping to support them to apply concepts and strategies in real situations (ARG, 1999: 7-8; Black et al., 2004:11; James, 2006: 56; Hodgson & Pyle, 2010: 20).

In addition, providing constructive feedback and sharing clear learning goals and success criteria are important to deepen students understanding as to constructivist view (Cowie, 2012:679-680). Opposed to the old paradigm in which assessment is viewed as a way of evaluating students' accumulation of factual knowledge and grading, in the current perspective assessment was viewed as an activity designed primarily to improve students' learning by actively engage them in the assessment process (Anderson, 1998: 5; ARG, 1999:2; Shepard, 2000:6; Chappuis & Stiggins, 2002:1-2; Singh et al., 2012:198). That is why, currently, constructivism has received great "attention for its implication in relation to teaching and assessment" (James, 2006:55).

Constructivism has two main approaches: cognitive constructivist and social constructivist (James, 2006: 53; Fautley & Savage, 2008:16; Heritage, 2010: 7). The two approaches of constructivism have some differences because of their different research agenda on the nature of learning. Piaget, who is the father of cognitive constructivism, was more interested in the internal development of mental structure within the individual, while the father of social constructivism (Vygotsky) was more interested on the social and cultural nature of learning (Swan, 2005:2-5; Sjoberg., 2007:6).

### **2.2.2.1 Cognitive constructivism**

Cognitive constructivist perspective focuses on, what is going on in the mind of the learner rather than thinking of learning as a simple stimulus-response connection as the behaviourist view. For this view learning is both an active, developmental and reflective process of mental construction linked to interactions with the environment (Swan, 2005:3; Heritage, 2010: 7-9). As Piaget, we sense our world by assimilating new information into our prior knowledge (Schema) and by accommodating our prior knowledge into the new information (Swan, 2005:2-3). Opposed to behaviourist, cognitive scientists focus on the mind of the individual (James, 2006: 55; Bennett & Gitomer, 2009: 58). For them, students must go beyond memorization of facts to know what, how and why to be prepared to inter into in the global competitive market.

In this theory the learner is an active agent in the learning process (Steadman, & Svinicki, 1998:13; Heritage, 2010: 7-9). Moreover, prior knowledge, deductive reasoning, inductive reasoning, self-monitoring and self-regulation are the vital components of the teaching learning process (James, 2006: 53). Here, learning is almost completely a function of the learner's interpretation of events, since knowledge in the cognitive constructivist represents more than the accumulation of discrete facts or information. Explicitly, knowing is being able to integrate knowledge, skills, and also procedures in ways that are helpful for interpreting real-life situations and solving problems in it.

The central point of assessment in this perspective is that assessment is a way of knowing what students know and how, when, and whether they use what they know to solve real problems (Dysthe, 2008: 19). For them, assessment is not just a tool the teacher uses at the end of a science lesson or year. Rather, assessment is a way of gathering and interpreting evidences minute-by-minute throughout the instruction for the purpose of improving learning and to obtain a comprehensive picture of a student's progress. Therefore, unlike the traditional paper and pencil assessment methods which provide information only on a certain aspects of learning, teachers should utilize a wide variety of assessment strategies for these purposes, especially for science subjects. Because, as cognitive constructivist, higher-order science thinking, procedural knowledge to solve science related questions, and the application of knowledge into real-life

contexts can be assessed and improved using such alternative assessment methods (Dysthe, 2008: 19-20).

Moreover, Cognitive constructivist emphasizes the importance of providing ongoing constructive feedback to learners and the integration of self-assessment to the learning process (Swan, 2005:8; Underhill, 2006:168). Since, as to them, self-assessment encourages learners to continuously construct and reconstruct their knowledge in response to feedback. Assessment and feedback are important in students learning (Swan, 2005:3; Heritage, 2010: 4-6). For cognitive constructivists, good assessment practice give emphasis to learning with understanding and application, not on memorization of facts (Cizek, 1997: 6-7).

Generally, teachers' assessment practice in science subjects needs to inspire students to use existing knowledge to build up new information in different ways, such as to generate their own scientific concepts, to design an experiment, to analyze their data, to solve problems, to report what they found, and to apply their knowledge, skills, and thinking processes into real-life situations. When considering the issue of knowledge application in the outside school environment, we are referring the social constructivist learning perspective.

#### **2.2.2.2 Social constructivism**

After the mid-1990s, social constructivism views of learning got considerable attention in education and particularly in science education (Bell & Cowie, 2002: 19 and 115). Science educators develop a social constructivism view of science learning which focuses on the whole parts of the classroom, because, cognitive constructivism gives primacy to the individual mental structures and thought processes at the expense of social and cultural aspects of learning (Bell & Cowie, 2002:114-115; Stobart, 2008: 161; Cowie, Moeland & Otrell-Cass, 2013: 6-17).

Within the social constructivist paradigm, learning is occurred through interaction between the student, the teacher and the tasks in the social environment, where “the learner both shapes and shaped by the community of practice” (James, 2006:57). For them, knowledge is constructed and reconstructed both individually and socially. Therefore, the nature of both individual and social

aspects influence the process of meaning making or problem solving (Stobart, 2008: 151; Dysthe, 2008:20; Bennett & Gitomer, 2009: 58; Cowie, 2012:680). Since, in contrast to the behavioural and cognitive perspective, social constructivism recognizes the social aspects of learning: conversation, interaction, and the application of knowledge to the real-life situations (Richardson, 2003: 1625; Havnes, 2008: 108).

This perspective grows out of the work of Vygotsky (Swan, 2005: 4), and suggested that teacher needs to create conducive environment to stimulate learners to think and act in authentic tasks beyond their current level of competence (Heritage, 2010: 7-9). As Vygotsky, high-order cognitive thinking and skills develop from the social environment in which the individual live and actively practiced (Swan, 2005: 4). Hence, students should actively participate in the social interaction to build up their understanding. Here, learning occurs through ongoing participation of the individual in the cultural practices (James, 2006: 56-57). Besides, from social perspective, learning, action and thinking cannot be separated (Willis, 2011: 401). That is why; learning is seen as an integral and inseparable aspect of social and individual practice. Thus, tasks need to be collaborative, interactive and dynamic and students must be involved in the generation of problems and solutions (Shepard, 2000:12).

In these perspectives, language has also important role in the learning process to facilitate student's cognitive development (Swan, 2005: 4; James, 2006: 56-57; Fautley & Savage, 2008: 19-21) since, teachers and students work jointly to develop their skills and understanding to solve real-life problems (Swan, 2005: 2; Heritage, 2010: 8; Cowie, Moeland & Otrell-Cass, 2013: 6-13). Mostly, this perspective view learning as an apprenticeship process, in which students work on authentic problems with the help of teachers and peers (Havness, 2008: 102-109). So, teachers need to create conducive classroom environment in which students can be motivated to participate, think and act in real-life tasks (like apprentice) beyond their current level of competence (James, 2006: 57) or provide scaffolding of learning in the students' 'zone of proximal development' (Swan, 2005:4; Heritage, 2010: 8). Zone of proximal development is most often describe in terms of the difference between an individual independent performance and his or her performance in collaboration with others (teachers and peers) (Fautley & Savage, 2008: 19-20).

Like cognitive constructivist perspective, formative assessment, has also been examined from the standpoint of social constructivist. As many literatures indicated, social constructivist learning perspective gives a powerful theoretical framework for assessment for learning (James, 2006: 56-58; Stobart, 2008: 151-161; Orr, 2008: 133-134; Gioka, 2008: 146; Cowie, Moeland & Otrell-Cass, 2013: 6-17). For this perspective, assessment for learning is seen as an active social process, particularly accomplished by the quality of teacher-student interaction in the learning context (Dysthe, 2008:19; Heritage, 2010: 8; Willis, 2011:402). Thus, they inform us that formative assessment is a collaborative process between teachers and students rather than only depends on an individual mind (Bell & Cowie, 2002:116). Within the socio-cultural theory of learning, formative assessment is a multidirectional, contextualised and social activity which used to identify the gaps in students' understanding to examine the next step in learning (Bell & Cowie, 2002:18-19; Heritage, 2010:8).

Assessment is an interactive, dynamic, and collaborative process between the teacher and students, as well as between and among students. This means that the teacher should not only be interested in what students can do on their own (i.e. summative assessment), but what students can achieve in collaboration with the teacher within the 'zone of proximal development'. One function of assessment would be then, to help to identify this 'zone' for each particular student and to support progress within it. (Gioka, 2008: 146)

Thus, students and teachers must adopt a substantial change in their classroom practice to establish a common understanding during formative assessment process to improve learning. Moreover, according to socio cultural perspective, such culturally situated patterns of participation in formative assessment allows students to develop a sense of ownership (Willis, 2011: 402) and plays a great role in students' identity formation (Cowie, Moeland & Otrell-Cass, 2013: 17). Here, the language that the teachers and peers used in assessment practice is important to help students to view themselves as able learners (Dysthe, 2008:20-21). Since, formative assessment provides different opportunities for learners to reflect their ideas, to evaluate their understandings and others (peers) and to understand valuable practices in the classroom community (Black & Wiliam, 2006: 13-15; Willis, 2011:402).



As a result sharing of learning goals and assessment criteria, providing constructive oral and written feedback, providing time for self- and peer-assessment and developing effective dialogue is crucial in formative assessment to implement it effectively (Havnes, 2008: 109; Gioka, 2008: 146; Willis, 2011: 402).

Furthermore, within the social constructivist view, assessment should be used to integrate the classroom learning activities to the outside real-life contexts (work places) (Shepard, 2000: 5; Dysthe, 2008:20). By supporting this idea Havnes (2008: 102), stated that “assessment has a vital role to play in bridging the boundaries between education and work and the wider society.” Generally, constructivist learning perspectives put greater effort on our classroom assessment to shift it from decontextualized to authentic; from using one single measure to using multiple assessment methods to build a student’s learning profile; from assessing only low level of competence and understanding to assessing high level of cognitive thinking and skills also; from assessing a few to assessing many dimensions of learning; from isolated assessment to integrating assessment within the learning and teaching practices; and from teacher directed assessment to increasing student responsibility in the assessment process. Then paradigm shift comes into the assessment practice (i.e. from assessment of learning to assessment for learning) (Black & Wiliam, 1998: 7; Anderson, 1998:9; ARG, 1999: 2; Bell & Cowie, 2002: 12; Harlen, 2003: 13-28; Torrance, 2007: 281-282). Therefore, in the following section we will discuss more about the concept of assessment for learning.

### **2.3 Putting assessment for learning into practice**

Many research evidences showed that formative assessment is more powerful to improve learning standards, particularly in science education. Hence, formative assessment should be integrated with instruction, authentic and continuous to improve learning. However, there are many factors that inhibit teachers to practice such formative assessment in their daily instruction. This section presents the concept of assessment for learning, research evidences on the role of formative assessment on learning, good qualities of formative assessment and factors that hinder the practice of assessment for learning.

### **2.3.1 The concept of assessment for learning**

According to ARG (2002:2) assessment for learning has been defined as “the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there.” It is a crucial element of an effective instruction (Dunn & Mulvenon, 2009:9). Ample research evidences indicated that assessment for learning is powerful to improve students’ science performance, to increase learning independence, to inspire students’ motivation to learn and risk-taking behavior, to develop self-confidence and positive self-esteem, to enhance positive relationship and reflection, to create conducive classroom atmosphere, and to facilitate two way communication (ARG, 2002: 3; Harlen & Crick, 2003: 170-175; Black & Wiliam, 1998a and 1998b cited in Black et al., 2004: 10; CCEA,2007:3-5; Moss & Brookhart, 2009:28).

However, effective implementation of assessment for learning requires the application of principles and specific elements to achieve the desired results. For this reason, Assessment Reform Groups identified ten guiding principles on their leaflet to make the practice of assessment for learning effective. For them, assessment for learning should be a part of effective planning, focuses on how students’ learn, be central to the classroom practice, be a key professional skill, be sensitive and constructive, foster motivation for learning, promote understanding of goals and criteria, help learners to know how to improve, develop the capacity for self-and peer assessment and recognize all educational achievement (ARG, 2002: 2). Its core principles are to promote the use of assessment to support learning through the active involvement of the learners in the learning process to take responsibility for their own learning (Mansell, James, & ARG, 2009: 9-10; CERI, 2008: 10; Irons, 2008: 79).

Moreover, other research evidences have shown the decisive element of assessment for learning to put it into practice effectively in science subjects. These comprise: sharing of learning target and success criteria with students, providing constructive written and oral feedback that helps students to recognize next steps in learning, involving students in peer and self- assessment, and using effective questioning and dialogue strategies (Black & Wiliam, 1998: 17; ARG, 2002: 7; Lee, 2006:44; CCEA, 2007:6-7; Hodgson & Pyle, 2010: 20; Department for Children, Schools

and Families (DCSF), 2010:6). This means that students are given the chance to clearly understand the learning targets and their own progress toward those learning targets, to assess their own work, and that of their peers, to collaborate with their teachers and peers in developing criteria and norms for their work, and to ask questions and to be a part of the discourse community. Generally, assessment for learning is not a revolutionary concept but a product of constructivist scientific research work.

As stated above, many scientific studies confirmed the effectiveness of assessment for learning to improve learning. Under here we discuss those effective elements of assessment for learning in science fields. These are: sharing of learning targets and criteria for success, formative feedback, peer and self-assessment, and questioning and dialogue.

### **2.3.1.1 Learning targets and success criteria**

The establishment of students learning goals is the first step in the whole assessment for learning process (CCEA, 2007:8; Moss & Brookhart, 2009: 24). The goal of assessment for learning is to guide students toward the development of their own learning to learn skills. Hence, to develop the required abilities and skills learning targets and the criteria for success must be clear to students (Shepard, 2000:11; Lee, 2006: 62). Students can only achieve learning objectives if they understand those objectives, take responsibility, and assess their own progress in line with those objectives (Black & Wiliam, 1998: 23-24).

Moreover, students who have a clear understanding about learning targets and the criteria by which their work will be assessed are better able to take responsibility for their own learning (CCEA, 2007: 8; Fautley & Savage, 2008: 47; Moss & Brookhart, 2009: 24-28) and they are more likely to be challenge seekers rather than challenge avoiders (Lee, 2006: 46). Clarity in assessment procedures, processes and criteria is important to enhance individual and institutional achievement (Torrance, 2007: 282). Hence, to implement assessment for learning productively learning targets and success criteria should be clear (Stiggins, 2005: 328) and these have a significant impact on what teachers teach and students learn (Keeley, 2008:12). Transparency in

learning intention and success criteria enables learners to assess their own progress during a lesson (Lee, 2006:62).

In addition, sharing of learning intentions and success criteria has positive impact on students' motivation to learn (CCEA, 2007: 8; Moss & Brookhart, 2009: 28). Research has shown that students who regularly share learning targets and success criteria are: "more focused for longer periods of time; more motivated; more involved in their learning; and better able to take responsibility for their own learning" (CCEA, 2007:8). Thus, assessment and instruction must take place with a clear target and criteria in mind to improve learning.

Chappuis & Stiggins (2002:5) also discussed its implication on active involvement of students in the assessment process. Students engage in the assessment for learning process when they use assessment information to set goals, make learning decisions related to their own improvement, develop an understanding of what quality work looks like, evaluate themselves in relation to the stated criteria and communicate their status and progress toward the established learning goals. Moreover, student who has an awareness of learning targets and success criteria is better able to set goals, develop a variety of learning strategies, develop meta-cognitive skills (Keeley, 2008:11-13) and control and evaluate his or her own learning process (Torrance, 2007: 283; Bloxham & Boyd, 2007: 59-62).

Likewise, sharing success criteria guides students' learning (Keeley, 2008:11-13), help them to see their progress against the learning objective (Lee, 2006: 46), encourages an independent approach to learning (CCEA, 2007:12) and provides a framework for assessment tasks and for the interpretation of evidence (Bloxham & Boyd, 2007: 59-62). Also, sharing success criteria foster a positive classroom environment; encourages students to be active participate in learning; build students' self-esteem and confidence and strengthen teacher-student relationship (CCEA, 2007: 13).

For these reason, Nicol & Macfarlane-Dick and other researchers mentioned some effective strategies for teachers to elucidate learning targets and standards for students to increase their own learning ownership. These include: providing carefully defined assessment criteria and

objectives; discussing and reflecting assessment criteria and standards in class (CCEA, 2007:10-11); involving students in peer and self-assessment in relation to defined criteria and standards; and conduct collaborative workshops (Nicol & Macfarlane-Dick, 2006: 205) and show sample or model of good work and further explanation (CCEA, 2007: 14; Wiliam, 2008: 37; Dysthe, Engelsen, Madsen & Wittek, 2008: 129; Moss & Brookhart, 2009: 27).

Realistically, most teachers and supervisors believed and practiced that writing learning objectives and success criteria on the board or telling it for students is enough, but it is not sufficient to understanding it to achieve the desired outcomes (Fautley & Savage, 2008:46-47; Moss & Brookhart, 2009: 24-25; Havnes, Smith, Dysthe, & Ludvigsen, 2012: 22). Therefore, teachers should share learning targets and standards using appropriate assessment tasks, actively involve students in the assessment process starting from planning (Havnes et al., 2012: 22), and follow-up their activities (Fautley & Savage, 2008:46-47).

Thus, learning targets and success criteria should be: planned in advance to reduce workload, shared with learners using appropriate language (CCEA, 2007:10-11), and fully understood and used by students (Lee, 2006: 46) because, learning targets and success criteria provide the standard against which evidence is elicited, performance is compared, and feedback is generated to close the gap between current learning and desired outcomes.

### **2.3.1.2 Formative feedback**

Formative feedback is defined by Irons (2008:7) as “any information, process or activity which affords or accelerates students learning based on comments relating to either formative assessment or summative assessment.” The information supplied by feedback may come from themselves, peers, books and teachers (Gamlem & Smith, 2013:150). However, successful provision and use of this feedback is dependent on the nature of the classroom climate in which the learning is taking place (Heritage, 2011, 19). Therefore, feedback requires more interactive and collaborative learning environment to be effective (Mentkowski, 2006:55-56).

Feedback to the learner is an essential part of formative assessment (Black et al., 2003: 42, Lee 2006: 56; Cheung, Cody, Hirota, Rubin, Slavin, Coffey, et al., 2006, 207; Irons, 2008:23 Sadler, 2010: 535; Havnes et al., 2012:21). Yet, to be formative feedback, it must help students to identify the gap in their learning and indicate next steps to fill the gaps (Lee 2006: 56; Heritage, 2010: 4). Researchers have shown that formative feedback is central to the development of effective learning (Moss & Brookhart, 2009: 56; Sadler, 2010: 536). Therefore, the provision of quality feedback is important to bring improvement in students learning (Irons, 2008:8; Sadler, 2010: 536). Feedback can be written or oral, or it can be given in the form of demonstration in the case of pre-school (Brookhart, 2008:55; Moss & Brookhart, 2009: 44).

There is considerable research evidence to show that effective feedback leads to learning gains. In short, Cross (1996) stated the importance of feedback as, “learning without feedback is like learning archery in a darkened room” (cited in Peterson, 2008: 7). In addition, there is a large body of complementary research studies demonstrating the effects of feedback on learning (Lee, 2006; Brookhart, 2008; Irons, 2008; Moss & Brookhart, 2009). Similarly, Hattie & Timperley mentioned the powerful effect of feedback on students learning than their prior knowledge, socioeconomic background, and reduced class size (cited in Heritage, 2010: 5).

Effective feedback is descriptive, criterion-referenced, constructive (not judgmental), incremental, positive, clear, specific, supportive in tone and focuses on the work and the process (Manitoba Education, Citizenship and Youth (MECY), 2006: 33; Norton, 2007:97; Brookhart, 2008:56). If such things are put into practice in the classroom, then students viewed “mistakes” as opportunities for learning, and they did not afraid of asking for help as needed (Moss & Brookhart, 2009: 56).

Moreover, many writers agreed that effective feedback should initiate thinking, give them insights about their work (Moss & Brookhart, 2009: 45), allows learners to develop self-reflection and empower learners to become self-regulated learners (Lee, 2006: 56; Fautley & Savage, 2008: 43; Moss & Brookhart, 2009: 45; Hodgson & Pyle, 2010: 16-18; Sadler, 2010: 536; Orsmond & Merry, 2011: 125).

Besides, Nicol & Macfarlane-Dick (2006:204) identified seven principles of good feedback practice from literature that might help learner to be self-regulated learners. Good feedback practice: helps to clarify what good performance is; facilitates the development of reflection and self-assessment in learning; delivers high-quality information to students about their learning; encourages teacher and peer dialogue around learning; encourages positive motivational beliefs and self-esteem; provides opportunities to close the gap between current and desired performance; and provides information to teachers that can be used to help shape the teaching.

Thus, feedback should be connected to the learning targets and success criteria, given on time, found where success has occurred, identified where and how improvement can take place, allowed time for improvement, accessed to the students in terms of meaning, and integrated in the formative assessment process (Weaver, 2006: 379; CCEA, 2007:15; Moss & Brookhart, 2009, 55). Students' learning improves and their understanding deepens when feedbacks are given timely, ongoing, embedded in the learning process and targeted feedback on their work (Black et al., 2003: 42-49; MECY, 2008: 9; Irons, 2008: 8-9; Peterson, 2008: 3; Fisseha, 2010: 103-105).

“Feedback to any pupil should be about the particular qualities of his or her work, with advice on what he or she can do to improve, and should avoid comparisons with other pupils” (Black & Wiliam, 2001: 6). The feedback should be: understandable and valued by students, “close the gap” on their understanding, of appropriate quality, timely, and provide appropriate classroom atmosphere for dialogue (Irons, 2008: 43-45).

Moreover, feedback which is provided through dialogue enhances students understanding rather than simple transmission (Nicol & Macfarlane-Dick, 2006: 208; Fautley & Savage, 2008:42-45). Particularly Dialogue between teachers and students is essential for effective feedback (Bell & Cowie, 2002: 16-17). Thus, teachers need to create a positive discourse community in their classroom to advance students' learning.

Even though many influential researchers have recently challenging the view, feedback is still generally conceptualised as a transmission of simple facts based on stimulus-response learning theories (Bell & Cowie, 2002: 16-17; Nicol & Macfarlane-Dick, 2006: 200). As a result, learning

improvements became absent and teachers' workload increases year by year as student numbers and class sizes increases (Nicol & Macfarlane-Dick, 2006: 200). However, within a constructivist framework, feedback is viewed as a dialogue between teachers and students rather than as merely receiving feedback from a teachers (Bell & Cowie, 2002: 16-17; McDowell, 2008: 241; Sadler, 2010: 535) because, the emphasis has shifted from memorization of facts and quantitative judgments to divergent and complex assessment tasks that are typically assessed using qualitative and quantitative judgments (Sadler, 2010: 535) and students are "not passive recipients of the teacher's judgments of their work" (Harlen, 2003: 19).

Quality feedbacks are highly specific, directly reveal what actually resulted, clear to the performer, and available or offered in terms of specific targets and standards (Bloxham & Boyd, 2007: 104-112; Peterson, 2008: 7; Teaching Development Unit (TDU), 2009:8). In other words, feedback should provide information about how the performer did in light of what he or she attempted, and further, about how this performance could be improved (Shepard, 2000: 11; Shepard, 2005: 68; Havnes et al., 2012:21). According to the Assessment Reform Group principle, learners need information and guidance in order to plan the next steps in their learning. As a result, teachers should pinpoint the learner's strengths and advice on how to develop them; be clear and constructive about any weaknesses and how they might be addressed; provide opportunities for learners to improve upon their work (ARG, 2002:2).

The language used in responding to feedback also affects the way in which a student receives feedback. Traditionally, teachers' feedback consists of right and wrong answers and judgmental statements and which focus on the self, such as "Good report.", "You need to study hard." "Your handwriting is very nice." to register their approval or disapproval of student performance (Shepard, 2000:11; Stobart, 2008, 148). But, such evaluative feedbacks have no room for improving students learning (Harlen, 2003: 29), can cause anger (Weaver, 2006: 380), and can actually have negative effects on students' desire to learn and self-esteem (Chappuis & Stiggins, 2002: 41; CCEA, 2007: 16; Stobart, 2008, 148; TDU, 2009: 7).

Likewise, if teacher feedback focuses on praise, reward, criticism and punishment, it has low impact on students learning (TDU, 2009:7; Stobart, 2008, 164-168; Yeh, 2013: 174). Moreover,



general comments that lacked guidance and focused on the negative side and unrelated to assessment criteria are unhelpful to improve learning (Bruno & Santos, 2010:112). Stobart (2008, 160-168) considered formative feedback which focuses on the self as a 'good murder'. By contrast, feedback has the potential to have a significant effect on students learning, when the information in the feedback is about a task, descriptive, constructive, give guidance on how to do it more effectively, and students are willing to dedicate effort (Gamlema & Smith, 2013: 152), and is clearly related to the learning goals (Black & Wiliam, 2001: 8; Jones, 2005: 3; Swaffield, 2011: 435; Havnes et al., 2012:21). Therefore, comments should be concrete, contextualized and related to the student's work. According to Nicol & Macfarlane-Dick (2006: 203) and Bloxham & Boyd (2007:20-23), to provide effective feedback whether it is oral or written, learners must know the learning objectives and success criteria for the task, the extent to which they have achieved the learning objectives and success criteria, and how to close the gap between what they have done and what they could do.

Feedback should play a role in enabling students to improve their performance on future assessment tasks (Moss & Brookhart, 2009: 45; Havnes et al., 2012:21-22). "The priority in giving feedback is to challenge students to tease out their assumptions and to help them be critical about the quality of arguments" (Black et al., 2004: 17). Providing effective feedback also reinforce students' capacity to self-regulated learning (Swaffield, 2011: 435), develop intrinsic motivation and self-esteem, inspire confidence and hope (Gioka, 2008: 147), fill gaps, and encourage dialogue among teachers and peers (Nicola & Macfarlane-Dick, 2006: 201-215; Irons, 2008: 2- 8; Heritage, 2010:4-6).

Furthermore, effective feedback enhances students' autonomy and motivation, helps to develop self-monitoring and meta-cognitive skills in the learner (Gioka, 2008: 146; Bruno and Santos , 2010:112), encourage deep learning (Higgins, Hartley & Skelton, 2002: 54), fosters resiliency and persistence, and provides students with specific strategies for next steps in their learning (Moss & Brookhart, 2009: 45-47). How the student interprets and deals with feedback is critical to the success of formative assessment (Poulos & Mahony, 2008: 144; Higgins, Hartley & Skelton, 2002: 54). The most useful feedback focuses on the qualities of student work or the processes or strategies used to do the work rather than the self (Brookhart, 2008:55). Therefore,

providing effective written and oral feedback is a key aspect in assessment to enable students to learn from assessment.

Research studies have also shown that marks and grades have little or no value to improve students' learning (Black & Wiliam, 2001: 8; Cheung et al., 2006, 216; Stobart, 2008, 168; Fautley & Savage, 2008:42-45; MECY, 2008: 9; Heritage, 2010:13) because, such feedbacks do not provide direction for next steps, do not encourage students to set and revise learning goals and do not provide specific guidance they need to learn (Black & Wiliam, 2001: 8; Kvale, 2007:65; Wiliam, 2008, 37; Stobart, 2008, 168). In addition, focusing on grades and marks can create a competitive culture in the classroom (CCEA, 2007: 15). When students are given test scores or grades, they focus on their performance compared to other rather than on their learning, and it can demotivate them to do further (Lee 2006: 56). Yet, formative feedbacks focus on improvement, de-emphasis competition and improve motivation and learning ambition (Lambert & Lines, 2001:137). So, feedback is useful when it is given as comment rather than in the form of grade or mark which focused on self.

Feedback is also important to inform teachers about modifications they need to make to promote more effective learning (Harlen, 2003: 19; Cheung et al., 2006: 216; Bloxham & Boyd, 2007:20-23). Thus, formative feedback is vital not only for students but also for teachers to adjust their instruction and to take actions that will help students to improve learning. However, according to Moss & Brookhart (2009: 46), many teachers perceived that giving marks, grades and detailed corrections as effective feedback. Generally, without effective assessment for learning strategies, which involves high-quality feedback to learners being embedded within daily instruction, the personalized learning agenda becomes a dream that will never be realized.

### **2.3.1.3 Peer and self-assessment**

Assessment is a process of collaborative communication in which information about learning flows between teacher and student; it is not something that teachers do to students (Lambert and Lines, 2001: 138; Wragg, 2001: 68). Rather, students need to be actively involved in the assessment process to achieve learning goals (Lambert & Lines, 2001: 141). This two-way

exchange of information between teachers and students is the heart of assessment for learning. To improve learning standards through assessment, students must be involved in peer and self-assessment (MECY, 2008: 10; Lee, 2006: 61), because collaboration between students to students and teachers in the assessment process is vital to share responsibilities in the learning process (Heritage, 2010: 14; Heritage, 2011:18-19).

Peer assessment is the involvement of students in assessing the work of other students or their peers; while Boud (1991) defines self-assessment as “the involvement of students in identifying standards and/or criteria to apply to their work and making judgments about the extent to which they have met these criteria and standards” (cited in Liu & Carless, 2006:281). Peer- and self-assessment are important forms of assessment for learning that engage students in talking about their learning (Lee, 2006: 61). Peer assessment often goes hand-in-hand with self-assessment (Black et al., 2004: 14) and some writers express it as a prerequisite for self-assessment to occur (Liu & Carless, 2006:281). Peer assessment helps students to learn the skills of self-assessment and also provides a rich resource of ideas that students can use in their own learning (Lee, 2006: 61). Developing peer- and self-assessment skills are very important in order to make learners more autonomous and to help them to identify their own learning needs and next steps in their learning (Hodgson & Pyle, 2010: 4; CCEA, 2007:26).

Furthermore, researchers pointed out the effectiveness of peer and self–assessment for academic staffs to save instructional time and reduce workload (Liu & Carless, 2006:281; CCEA, 2007: 26; Irons, 2008: 32 and 80) and when there is resource constraints (Liu & Carless, 2006:281; CCEA, 2007: 26). Peer- assessment benefits both students who give the feedback and for those who receives the feedback to deepen their understanding about their own learning (Heritage, 2010: 14). Researchers revealed that peer-assessment is important to relate students with multiple strategies and procedures to reach solutions; develop detachment of judgment; and increase motivation to persist on a task (Nicol & Macfarlane-Dick, 2006: 208; Bloxham & Boyd, 2007:23-24), and it is valuable for students because most of the time they accept critiques of their work from peers rather than teachers (Black et al., 2004: 14). Moreover, peer-assessment is central to facilitate dialogue among teachers and students, which is powerful for effective learning (Irons, 2008: 79).

Similarly, self- assessment helps learners to become self- reliant, increase their motivation and self-esteem (Lambert and Lines, 2001: 142), develop higher level of commitment, and to be an effective lifelong learner (Falchikov, 2005: 120-121; Lee, 2006: 61). The notion of becoming a lifelong learner is to develop students' capacity in assessing their own learning (Boud & Falchikov, 2006: 402). Likewise, enhancing self- assessment in learning is one effective way to develop self-regulation skills and reflection on learning progress (Lee, 2006: 61; Nicol & Macfarlane-Dick, 2006: 205; Moss & Brookhart, 2009: 82), to cultivate independent learning, to develop meta cognitive skills (Stobart, 2008, 149; Heritage, 2011: 19), and to advance social skills (MECY, 2008: 13). Self-assessment establishes a positive and collaborative working environment through breaking the pattern of passive learning (Lambert & Lines, 2001: 143).

The process of self-assessment based on an explicit set of criteria increases students' responsibility over their own learning and makes the relationship between teacher and student more collaborative (Lee, 2006: 61). This further develops students' self-regulation and self-direction of their own learning which are critical skills for lifelong learning in the 21st century knowledge based economy (Panadero, Alonso-Tapia, & Reche, 2013: 1). Therefore, "self-assessment is a very powerful teaching tool and crucial to the assessment for learning process" (Jones, 2005: 19).

Moreover, encouraging self-assessment helps students to avoid 'helplessness' by reducing their attribution of their success and failure towards external and uncontrollable events and encourage students to do their work confidently (Lambert & Lines, 2001: 143). As Ross et al. (2002: 45) cited in Moss & Brookhart (2009: 82) indicated that, "Students with greater confidence in their ability to accomplish the target task are more likely to visualize success than failure." Therefore, classroom assessments should build students' confidence in themselves as learners and to take responsibility for their own learning so as to lay a foundation for lifelong learning.

Various research evidences have shown that when students take an active role in monitoring and regulating their own learning, the rate of their learning is dramatically increased (Black & Wiliam, 1998a: 26; Nicol & Macfarlane-Dick, 2006: 205; Irons, 2008: 79) and perhaps doubled.

It is widely recognized that when learners are fully engaged in the learning process, learning increases (Jones, 2005: 19). As well, large learning gains can occur when students serve as learning resources for one another, as evidenced by CCEA (2007: 27) and Wiliam (2008: 37). To make peer assessment valuable, students must work as a group or team and each student must be responsible in some way and students need to be trained the skills of collaboration in peer assessment (Black et al., 2004: 15; Lee, 2006: 62; CCEA, 2007: 28). However, peer assessment is difficult or even unfair to put in to practice within a competitive working environment between students (Wragg, 2001: 71; Kvale, 2007: 66-67).

For peer and self-assessment to be productive, learning targets, success criteria and assessment tools should be clear for students (Lambert & Lines, 2001: 143; Lee, 2006: 61; Liu & Carless, 2006: 280), good self- and peer assessment models should be showed for them (CCEA, 2007: 28), time should be given to practice and refine their self- and peer assessment skills, they should not be required to assign marks for their own and/or peers work, timely and descriptive feedback should be given on their work (Nicol & Macfarlane-Dick, 2006: 206; MECY, 2008, 13; Moss & Brookhart, 2009: 80), and students should be trained and guided in self and peer assessment skills (Black & Wiliam, 2001: 6; Bloxham & Boyd, 2007: 63-64).

However, most of the time, the term assessment is perceived as referring to marking, grading, measuring or ranking (Liu & Carless, 2006: 280) and as a result self and peer assessment is regarded mainly as students giving marks or grades to their own work or to each other (Liu & Carless, 2006: 280; Moss & Brookhart, 2009: 80) and teachers use these skills as a “time-saver for grading tests” (Moss & Brookhart, 2009: 80). Beside, Tan (2008: 235) discussed, on the difficultness of changing teachers and students perception toward self-assessment. But, more significantly, self- and peer-assessment can be a very productive way to involve students’ actively in formative activities (Bell & Cowie, 2002: 19), and can help students to learn from their assessment and feedback practice and help them to fill the gaps in their learning (Bloxham & Boyd, 2007:23-24). Thus, these skills are processes that students do as part of their learning activities (Black et al., 2003: 49; Black et al., 2004: 14). On the other hand, peer and self-assessment helps teachers to understand the difficulties, progress and ideas of their students quickly and deeply (Lee, 2006: 61). Generally, to implement assessment for learning in their

classroom teachers should encourage peer and self-assessment as a result they will improve students' learning and achievement in science subjects.

#### **2.3.1.4 Questioning and dialogue**

Questioning is a key element to implement assessment for learning and science learning in particular (Lee, 2006:49; Hodgson & Pyle, 2010: 13). Equally, others identify it as an integral part of assessment for learning practice (Jone, 2005:11-15; CCEA, 2007: 21). But, questions should be effective and strategic to explore and encourage students to think (Lee, 2006: 49-50). According to Lee, to stimulate students' understanding and thinking, particularly, in science subjects rich and effective questions should be practice as a part of assessment for learning.

Researchers have shown that effective questioning and dialogue in the classroom is important to improve students understanding (Jone, 2005: 11-13; CCEA, 2007: 21; Irons, 2008: 63), to assess prior knowledge and to get evidence about the level of their understanding (Wragg, 2001: 33), to communicate learning targets (Moss & Brookhart, 2009: 29), to engage students in the lesson through thinking (Lee, 2006: 49-50) and to develop a learning culture of open discussion (Irons, 2008: 63).

Questions can be effective when forwarded from both teacher and student (CCEA, 2007:22; MECY, 2008: 7) because, the direct way of asking questions by teachers is not effective to improve students' learning (Black & Wiliam, 2001:7). The Northern Ireland Curriculum, positioned student-led question as a key process to achieve learning gains, because, it promotes their engagement; a sense of self-reliance; and develops ability to reflect, to work with difficulties, and to evaluate their own learning (CCEA, 2007: 25).

Alongside, studies identify effective strategies of questioning to be productive. These are: encouraging students questions, planning questioning, asking effective questions, asking questions better, giving enough time to think, and dealing with answers productively (Black et al., 2004: 11-13; CCEA, 2007: 21-25; Moss & Brookhart, 2009: 96). Moreover, questions should

be challenging, cause deep thinking, provoke discussion, explore the full range of learning targets, and buildup from previous learning (Lee, 2006: 50).

Besides, the questioning strategy should involve dialogue between teacher and student, not merely the teacher telling the student what she/he did correctly and incorrectly. Building dialogue among teachers and students is important for both to increase students learning and for teachers to respond and direct students thinking (Bloxham & Boyd, 2007:113). Hence, “the dialogue between pupils and a teacher should be thoughtful, reflective, focused to evoke and explore understanding, and conducted so that all pupils have an opportunity to think and to express their ideas” (Black & Wiliam, 2001:7).

Questioning is a skill that needs careful planning. For example, Moss & Brookhart (2009: 102-103) specified that strategic questions are not asked ‘on the fly’ rather than planned in relation to the learning targets. Teachers need to plan questions rather than hope that a question pops into their head at an appropriate moment (Jone, 2005: 11-13). Questions that are carefully planned encourage classroom discussions, actively engage students in the learning how to learn skills and increase teachers listening ability (Moss & Brookhart, 2009: 104). Besides, effectively planned questions can speedy students to think and provide teachers with information to adjust instruction (Wiliam, 2008: 37).

In addition, strategic questions are essential for students to increase their self-efficacy, to set goals, to develop self-regulation skills, and to attribute their success and failure towards the learning strategies they used and the amount of effort they put into the learning targets (Moss & Brookhart, 2009: 29 and 102). Therefore, questions do more than auditing learning during formative assessment process. However, “many teachers do not plan and conduct classroom dialogue in a way that might help students to learn” (Black et al., 2004:11).

Furthermore, researches have shown that giving sufficient thinking time for students after asking a question is an important part of effective questioning (Black & Wiliam, 2001: 7-8; Black et al., 2003:32-33; Black et al., 2004: 11-13). Wait time (thinking time) is essential to engage every student actively in answering higher order questions (Black et al., 2004: 13; Lee, 2006: 52) and

helps students to attribute their success and failure on their effort rather than on their ability (Black et al., 2004:13). Similarly, Rowe (1974) explained the significance of waiting time. She found that waiting time during questioning increases the explanation of answers, students' confidence in their own ability to respond in meaningful way and the number of students who knows the answer (cited in Black et al., 2003: 13; Fautley & Savage, 2008:38). As well, wait time help teachers to ask better question that reveal where the learners are in their learning (Stobart, 2008: 148).

On the other hand, research has shown that many teachers give less than two second of wait time between posing the question and asking for the answer (Black et al., 2004:11). This is not exception in Ethiopia. From my experience many teachers are looking for a particular expected answer within a short time and if no answer is coming, they ask other question or answer the question themselves. Most of us lack flexibility to deal with the unexpected ones.

Yet, short wait time is useful only for few students and questions that are answered quickly (memorization of facts) (Black & Wiliam, 2001:8; Black et al., 2003:32; Black et al., 2004:12). These low levels of questions and the involvement of few students in answering such questions can keep the lesson going but is actually out of touch with the understanding of most of the class students. As a result, the question-answer dialogue becomes superficial; one in which all plan and thoughtful involvements suffers (Black & Wiliam, 2001:8). Consequently “coverage without understanding is pointless and may even harmful, as pupils will learn that they cannot successful at learning” (Lee. 2006:52). Wragg (2001:32) emphasized the importance of open ended higher order questions to deepen students understanding. Accordingly, teachers need to use effective questioning techniques that keep all students engaged and gauge the understanding of the whole class instead of just selected students (Wiliam, 2008:37).

Additionally, creating supportive climate in the classroom, giving value and explore both right and wrong answers and no hand up strategies are important in helping students to think and answer the question (Lee, 2006: 51-52; CCEA, 2007: 25). Thus, science teachers should pay attention not only for correct answers, but also to the students thinking because wrong answers might indicate unexpected thinking rather than lack of understanding.



### **2.3.2 Research evidence on how formative assessment improves learning**

The purpose of this section is to show the empirical research evidences on the effect of formative assessment on students learning, particularly, on secondary school science education. Therefore, this section presents the effect of formative assessment on students' achievement, motivation to learn, self-esteem, self-efficacy, and self-regulation.

#### **2.3.2.1 Formative assessment in science education**

Research evidences suggest science education as a crucial element for personal, economical, and social development in one's nation (Goodrum, Hackling, & Rennie, 2001: 23; Kola, 2013: 225). Science education is imperative to produce good citizens who can actively engage in poverty reduction, environmental protection and health improvement (MoE, 2002: 31; MoE, 2010: 11; FDRE, 2010: 1), because it plays a fundamental role in developing scientific literacy (Goodrum, Hackling, & Rennie, 2001: 6-11). Similarly, the Ministry of Education in Ethiopia promoted the role of quality science education to develop critical, innovative and creative thinkers who can actively participate in policy development (MoE, 2005: 31).

Science education is a foundation for science and technology which strongly shapes our world nowadays (MoE, 2010: 9; Kola, 2013: 225-226), because "Science and technology plays a prominent role in our lives today, and will into the future" (Cowie, Moreland, & Otrell-Cass, 2013: 4). Currently, many countries including Ethiopia critically emphasis on the provision of quality science education at all levels of learning (Kola, 2013: 225). The government of Ethiopia realizes science education as the major instrument to become a middle income country within 2025 (FDRE, 2010: 1) because, quality science education for all is crucial to achieve the Millennium Development Goals adopted by the world's leaders in 2000 (MoE, 2005: 31). Therefore, science education should be taught to develop students who are scientifically literate citizens rather than to prepare them only for academic careers in science (Hofstein, Mamlok, & Rosenberg, 2006: 139). Particularly, secondary school science education is a basis to produce citizens who are qualified in science and technology (Adula & Kassahun, 2010: 32) so that

students need to develop the required knowledge, skills and attitude that enable them to be successful in their profession.

However, students' achievement is low in science and mathematics subjects throughout the world (King, 2006:3; Velayutham, Aldridge, & Fraser, 2011: 2160). Due to these facts, science and mathematics curriculum, instruction, and assessment have undergone a change from an authoritarian and teacher-centred approach to a collaborative, student-centred and inquiry-based approach (Farenga, Joyce, & Ness, 2006:41). Moreover, the domination of constructivist learning theories in education empowered this shift (King, 2006:3; Velayutham, Aldridge, & Fraser, 2011: 2160). Constructivist learning theories emphasized on the use of student-centered and formative assessment methods in the classroom to actively engage students in the learning process (Hasan & Ehsan, 2013:577) because, student-centered assessment and teaching methods have a positive influence on students' science achievement (Harlen, 2003: 20; Black et al., 2004: 10; Odom, Marszalek, Stoddard, & Wrobel, 2011:2361-2370).

Accordingly, policy makers in Ethiopia designed “high quality, dynamic, practical, interactive and attractive science and mathematics curricula at teachers training institutes as well as in primary and secondary schools” (FDRE, 2010: 5), and focus on the practices of better assessment techniques to enhance students' science achievement (MoE, 2005:48). New modes of assessment is vital to improve the quality of science education (Goodrum, Hackling, & Rennie, 2001:20; Bell & Cowie, 2002:20-21; Harlen, 2003: 7- 20) to improve students' learning and to get full picture of their progress in science education (Allen, 2007:4) because, “science involves much more precisely delineated concepts than many school subjects, so assessment ought to be more exact” (Wragg, 2001: 63). Therefore, science curriculum, instruction, and classroom assessment should be organized based on these core science concepts (Vitale, Romance, & Dolan, 2006: 3). Ample research evidences have revealed the special gain of formative assessment especially for science education to improve students' understanding and application of scientific concepts and achievement (Bell & Cowie, 2002:80; Harlen, 2003: 7; Lauvas, 2008: 163; Keeley, 2008:19; Velayutham, Aldridge, & Fraser, 2011: 2159; Hasan & Ehsan, 2013:577). In line with this, Black & Harrison (2004) explained:

Science provides the means by which learners can interact with the world around them and develop ideas about the phenomena they experience. So, when they attempt activities such as germinating seeds, their studies in science equip them with ways to observe and question what is happening. Through experiments they can begin to work out and predict what might happen if conditions for the growing plant change. To be able to learn science in this way, student needs help in developing process skills to investigate and communication skills to question and discuss findings. Formative assessment fits well into this learning scenario, since its purpose is for teachers to sift the rich data that arise in classroom discussion and activity, so that professional judgments can be made about the next steps in learning. (p. 3)

Similarly, Harlen (2003: 7) emphasized the impossibility of practicing inquiry-based science learning in the classroom without formative assessment. Formative assessment is important “to empower students by nurturing the belief that they can succeed in science learning and to cultivate the adaptive learning strategies required to help to bring about that success” (Velayutham, Aldridge, & Fraser, 2011: 2159). It helps them to understand complex science concepts (Bell & Cowie, 2002: 19-21; Harlen, 2003: 7-25). Then the understanding of such complex science concepts helps students to know and explain the principles that underlie science phenomena using coherent evidence (Liu, Lee, & Linn, 2011: 164).

However, assessment, which emphasis on the recall of isolated facts, encourages shallow learning and memorization of scientific facts (Liu, Lee, & Linn, 2011: 164) rather than deepening students’ understanding of scientific concepts and skills which are crucial in today’s’ education (Harlen, 2003: 10). Moreover, Henriques, Colburn, & Ritz (2006: 15-16) pointed out the limitation of these traditional forms of assessment such as multiple choice, true-false, short-answers, and matching items to assess such understandings. Multiple choice items do not provide students with the opportunity to explain, reason, or justify their choices and promote complex arguments or standards-based coherent ideas and understanding (Liu, Lee & Linn, 2011: 164).

Science teaching standards also give less emphasis on testing the acquisition of factual information and asking for recitation of acquired knowledge (Farenga, Joyce, & Ness, 2006:42) because, such factual scientific question do not allow students to fully understand the principle, the cause and effect relationship between variables and the application of scientific concepts in

the real situations. (Wragg, 2001: 63; Odom et al., 2011:2352). Assessment which focuses on recall of information is likely to reinforce surface or rote learning (Bell & Cowie, 2002: 20-21) and this affects students' future learning (Odom et al., 2011:2352). It may also motivate teachers to emphasize on superficial memorization of science facts rather than promoting deep scientific understanding (Liu, Lee, & Linn, 2011: 165).

Yet, the teaching standards in science education highly emphasized on students' coherent understanding and use of scientific concepts, knowledge, ideas, and inquiries in real life situations (Liu, Lee, & Linn, 2011: 164-166). Now, students need to develop "the capacity to readily acquire new knowledge, to solve new problem, and to employ creativity and critical thinking in the design of new approaches to existing problems" (President's Committee of Advisors on Science and Technology in United States reports, 1997 quoted in Harlen, 2003: 8). Therefore, how to assess such complex concepts in science education is critical issue (Cowie, Moreland, & Otrell-Cass, 2013: 4).

Hence, the teacher should provide opportunities for students to actively engage in scientific inquiries, discussion and debates, questioning, feedback and peer and self-assessment to improve their learning (Harlen, 2003: 19-20; Chin, 2006: 1315; Farenga, Joyce, & Ness, 2006:42). However, due to lack of appropriate assessment methods and teachers confidence in using them, students' science achievement was hardly assessed using laboratory works, group and individual projects, reading scientific articles, constructive feedback, peer and self-assessments and dialogues. As a result, students received final grades based mainly on their abilities on paper-and-pencil tests (Hofstein, Mamlok, & Rosenberg, 2006: 140).

Therefore, science education reform is important to integrate assessment with curriculum and instruction (Hofstein, Mamlok, & Rosenberg, 2006: 140) and students need to develop a variety of learning experiences to deepen their understanding of scientific concepts (Vitale, Romance, & Dolan, 2006: 3). As Farenga, Joyce, & Ness (2006:4) argued, "the purpose of a particular methodological approach to teaching is dependent on what a particular society values at any given time." Therefore, assessments that support students' learning should be prioritized

(Harlen, 2003: 19-22). Assessment should focus on conceptual understanding which really have value in science education rather than on concrete information (Allen, 2007: 5).

### **2.3.2.2 Formative assessment and students' achievement**

Many researchers discussed that 'formative assessment' and 'assessment for learning' are identical (Black et al., 2004: 10; Fautley & Savage, 2008:38) because formative assessment (assessment for learning) is closely connected to effective learning (Lee, 2006: 1-2). Therefore, the researcher uses the terms interchangeably in this study. There are many definitions of formative assessment. Heritage (2007: 141) described it as "a systematic process to continuously gather evidence about learning." According to Moss & Brookhart (2009:6), formative assessment is "an active and intentional learning process that partners the teacher and the students to continuously and systematically gather evidence of learning with the express goal of improving student achievement."

Similarly, Harlen (2003:7) conceptualized formative assessment as the process of "gathering and use of information about students' ongoing learning by both teachers and students to modify teaching and learning activities." "Formative assessment is not a test, or instrument", but it is a practice that empowers teachers and students to give their best to enable learning (Heritage, 2011:19). Most of the definition reflects views of learning in which students actively participate to construct their own knowledge rather than being passive receivers of knowledge. In recent times, educational practices are inclining towards constructivist learning theories which focus on student-centred models to support independent, self-regulated and self-motivated learning (Hobbs, Brown, & Gordon, 2009: 55).

Nowadays, the power of formative assessment in improving students' learning and raising standards, especially, in science subjects is unquestionable (Bruno & Santos, 2010:111). Variety of assessment books, research works and journals suggest a theoretical framework on the importance of formative assessment to improve science learning (Harlen, 2003: 7-20; Black et al., 2004: 10; Fautley & Savage, 2008:38; Keeley, 2008: 19-28; Harlen & Gardner, 2010: 16-40). Research evidences gathered and analyzed from 250 studies by Black & Wiliam around the

world across different subjects and grade levels over the past two decades demonstrated that formative assessment is the most significant tool to realize profound and substantial gains in all student achievement, especially for low achiever students (cited in Lambert & Lines, 2001:147; Harlen, 2003: 7; Black et al., 2004: 10; Moss & Brookhart, 2009: 9; Stiggins, 2008:8-9; Stiggins, 2009: 21; Greenstein, 2010: 7; Harlen & Gardner, 2010:21). Hence, formative assessment plays an important role to reduce the range of achievement gaps (Moss & Brookhart, 2009: 10) and to achieve the goal of understanding and thinking that are valued in education for the twenty-first century (Harlen & Gardner, 2010:21).

Following the comprehensive research works of Black & Wiliam in “Inside the Black Box” and “Assessment and Classroom learning” in 1998, Wiliam, Lee, Harrison, & Black (2004: 58-64) examined the achievement of secondary school students in mathematics and science subjects who were exposed and not exposed to formative assessment practices. And, the research team got the effects of formative assessment on learning outcomes and found a mean effect size of 0.32 when exposed to the intervention. Moreover, recent research conducted by Kingston & Nash’s (2011) shown significant effects of formative assessment on students achievement (cited in Herman, 2013: 6).

Moss & Brookhart (2009:5) associated formative assessment with the power of wind which generates energy. For them, formative assessment energizes or pushes students to understand and use learning targets, set their own learning goals, select effective learning strategies, and assess their own learning progress. Formative assessment actively engages and encourages students in the process of constructing, modifying and deepening their knowledge (Keeley, 2008: 16-26; Harlen & Gardner, 2010:19-20).

Many researchers indicated that the primary purpose of formative assessment is to promote students’ learning (Black et al., 2004: 10; Moss & Brookhart, 2009: 6). Therefore, to be formative, assessment should improve students understanding of the subject matter (Birenbaum, Kimron, Shilton, & Shahaf-Barzilay, 2009:13; Harlen & Gardner, 2010: 16), and shapes student’s learning towards the learning standards (Wiliam, 2006: 285). Thus the major goal of formative assessment is to identify the gap and understand what the student actually knows

(Stiggins, 2005:328; Heritage, 2007: 141-142; Birenbaum et al., 2009:131) and provide opportunity for teachers to advise their students to fill the gap between their actual understanding and the desired one (Black et al., 2004: 10; Bruno & Santos, 2010:111; Greenstein, 2010: 32-33).

Formative assessment encourages students to use their existing ideas and to build them to understand, explain, predict, hypothesis, evaluate and justify ideas, evidences and phenomena (Keeley, 2008: 16-17). However, memorizing of facts or a fixed set of procedures through paper and pencil tests does not support students to develop these abilities to apply their learning in the new context (Harlen & Gardner, 2010: 20). Additionally, formative assessment enhances the quality and quantity of teacher-student interactions through providing opportunity to examine and reflect scientific ideas (Keeley, 2008:16-17; Moss & Brookhart, 2009: 22) because, both teachers and students collect evidences together to advance students' learning and achievement (Heritage, 2007: 141; Greenstein, 2010: 7; Heritage, 2011:19). Formative assessment also increase teachers' professional quality, build teachers capacity and improve school standards (Moss & Brookhart, 2009: 1; Fautley & Savage, 2008:38) which are important to meet the demands of today's education.

Formative Assessment is cyclic in nature and it consists of five phases: defining objectives and success criteria; collection of evidences about students' skills, knowledge and attitudes; interpretation of the collected evidences (identify the gap); implementing interventions to close the gaps, and assessing the effectiveness of the interventions in closing the gaps (Harlen, 2003:20; Birenbaum, et al., 2009:131). Moreover, Keeley (2008:26-28) suggested the following points to build up the link between assessment and instruction through formative assessment practice and as a result to improve students science learning. These are: explore students' ideas in depth and analyze their thinking; encourage students to use their prior knowledge and experience to predict and construct their explanation during scientific inquiry; create a classroom culture of ideas, not answers; develop a discourse community (science talk); encourage students to take risks and listen other ideas carefully; use a variety of formative assessment techniques and grouping configuration; and encourage continuous reflection.

Moreover, to be effective in promoting thinking and learning, formative assessment should incorporate metacognition and reflection strategies which are a key component of assessment that promotes learning (Keeley, 2008: 18-26). Yet, such formative assessment is rarely implemented in classroom cultures (Black et al., 2003: 6; Moss & Brookhart, 2009:1).

Generally, formative assessment improves students understanding in science subjects. Therefore teachers need to implement formative assessment in science subjects through effectively practicing the components and phases of formative assessment.

### **2.3.2.3 Formative assessment and students' motivation to learn**

Motivation is something 'that energizes, directs, and sustains behaviour towards a goal' (Moss & Brookhart, 2009: 15). Motivation is crucial affective element in science learning (Velayutham, Aldridge, & Fraser, 2011: 2161). Ample research evidence indicated the impact of motivation on students science learning (Black et al., 2004:18; Harlen, 2006: 61-67; Moss & Brookhart, 2009: 15; Cauley & Mcmillan, 2010: 4-5; Harlen, 2010:40-42; Velayutham, Aldridge, & Fraser, 2011: 2161). Motivation determines students' engagement, achievement, learning strategies, and critical thinking in science learning (Velayutham, Aldridge, & Fraser, 2011: 2161- 2163).

Researchers consider it as an engine that drives students' science learning in the twenty-first century to adapt the ever changing conditions in their lives (Harlen & Crick, 2003: 173; Harlen, 2006: 61-67; Moss & Brookhart, 2009: 15). According to them motivation is also a crucial element to develop learners competences for effective lifelong learning. Students are willing to devote more time and energy to any activity (Black et al., 2004: 18; Nicol & Macfarlane-dick, 2006: 209; Irons 2008: 32-40), become more successful in their learning (King, 2006:33) and utilize their knowledge, skill and understanding in their lives beyond schooling (Fautley & Savage, 2008:88) when they are highly motivated.

However, students' motivation for learning is highly influenced by a variety of external and internal psychosocial factors (Harlen & Crick, 2003: 173). Therefore, it is important to be aware of what aspects of teaching and learning practice act to promote or impede it (Harlen, 2006: 76-



77). Numerous researchers indicated assessment as one of the key factor that affect students' motivation to learn (Harlen & Crick, 2003: 173; Harlen, 2006: 76-77; Fautley & Savage, 2008:88) and consequently their outlooks toward themselves (Black et al., 2004: 18). Conversely students' motivation for learning also affects teachers' assessment practice in the classroom (Fautley & Savage, 2008:88).

Views of motivation were highly influenced by learning theories. Harlen (2006:67) positioned the views of behavioural and constructivist learning theories on students' motivation for learning. According to Harlen, behaviourists focus on the importance of reinforcement and punishment to motivate students to do any given task (Harlen, 2006: 67). Moreover, MECY (2006: 6) pointed out the attention of behaviourists on assessment results and grading to motivate students to work hard and to learn. As stated by Stiggins (2005:324), this traditional view of motivation emphasis on the use of final exams, unexpected quizzes and failing report cards to motivate their students. For them, maximizing students' anxiety leads to high scores and good grades (Stiggins, 2005:324).

However, research evidences have shown that assessment that focuses on grades and marks motivate only few students who are high achieving but demotivate those students who do not do assessment activities successfully to invest time and effort in a given task (Black et al., 2004: 18; Nicol & Macfarlane-dick, 2006: 209; Moss & Brookhart, 2009: 16). As a result, such assessment systems affects their self-esteem and leads to a sense of helplessness (Black, et al., 2004: 18; Nicol & Macfarlane-dick, 2006: 209), and school dropout rate increases among less successful students (Harlen & Crick, 2003, 170; Harlen, 2006: 76-77). Moreover, such assessment systems, which focus on marks and grades, have a negative impact on lifelong learning preparation (Harlen & Crick, 2003: 170; Moss & Brookhart, 2009: 15).

Feedback, given in the form of grades or marks, allows most students to focus on their ability rather than on the importance of effort to be successful (Black et al., 2004: 18). This implies that students who believe, their successes or failure are due primarily to their effort will have stronger motivation and staying power to complete a challenging work (MECY, 2008, 9). Generally, according to Harlen (2006: 77), classroom environment that focuses on one-way learning,

competition among students, teaching the test and judgmental feedbacks in terms of scores and grades can have a negative impact on student motivation to learn.

On the other hand, constructivist learning theorists' emphasized on the active involvement of students in the learning process to motivate them to learn (Harlen, 2006: 67; MECY, 2006: 6). For them, when students have ownership in their learning particularly in the assessment process they are more motivated to invest time and energy to learn (Irons, 2008: 35). Harlen (2006: 67) also, mentioned the focus of constructivist theory on the importance of stimulating intrinsic interest of students to improve learning. As constructivist perspective assessment enhances student motivation when it emphasizes on success rather than failure; provides constructive feedback rather than judgmental; develops students sense of responsibility for their own learning; builds students confidence and self-esteem; is relevant, and appeals to students' imaginations; and provides learning targets and success criteria (MECY, 2006: 6). Thus, such assessments motivate students to take more responsibility for their own learning (Cauley & Mcmillan, 2010: 2), to make assessment an integral part of their learning experience, and to embed or integrate it in activities that stimulate students' abilities to create and apply a wide range of knowledge (Libman, 2010:63).

Motivation has two forms, extrinsic and intrinsic, which are associated with different learning strategies (Harlen & Crick, 2003: 170-175; Fautley & Savage, 2008:89; Moss & Brookhart, 2009: 15). From their review, extrinsic motivation in learning activities is associated with external drivers such as rewards, praise whereas intrinsic motivation is the student's internal desire to learn. Most researchers associated extrinsic motivation in learning activities to 'shallow' learning or memorization of facts (Libman, 2010:63) and intrinsic motivation is connected to 'deep' understanding (Bloxham & Boyd, 2007:27-29) and lifelong learning (Harlen & Crick, 2003: 170-175; Irons, 2008: 37 ). In fact, researchers have also included the negative effect of extrinsic motivation on intrinsic students' motivation (Harlen & Crick, 2003:175; Moss & Brookhart, 2009: 16). On the other hand, Fautley & Savage (2008:91) highlight the importance of combining the two kinds of motivation in classroom activities. Consistently, Harlen (2006: 63) indicated that intrinsic motivation enhances students' desire to continued learning than learning motivated extrinsically.

Motivation for learning is a combination of different elements which have a power to promote students' effort required for learning (Harlen & Crick, 2003: 170-175). These are: students' interest, goal-orientation, self-attribution, self-esteem (Black et al., 2004: 18; Nicol & Macfarlane-Dick, 2006: 210-212), self-efficacy, and self-regulation (Harlen & Crick, 2003: 170-175; Harlen, 2006: 64; Moss & Brookhart, 2009: 16). Therefore, students' sense of self, beliefs, habits of thinking, attribution of success and failure and higher order thinking affects their intrinsic motivation to learn (Harlen & Crick, 2003: 174; King, 2006:33) and consequently affects their academic achievement (Zimmerman & Schunk, 2001 cited in King, 2006:33). There is also a good deal of evidence that assessment has a fundamental role in promoting or inhibiting these key elements of motivation and hence affects what students' value in their education (Harlen, 2006: 64; CCEA, 2007: 2; Irons, 2008:36; Stiggins, 2008:8-9). Therefore, assessment strategies that build students' academic self-efficacy and motivation hold great promise as educational interventions.

Formative assessment is the crucial element to increase students' motivation to learn (Nicol & Macfarlane-Dick, 2006: 209; Irons, 2008:32; Moss & Brookhart, 2009: 16). In formative assessment, particularly, feedback is a powerful technique to support or inhibit student motivation and achievement (Irons, 2008:37; Cauley & Mcmillan, 2010: 5). Feedback helps students to address their internal desire through encouraging them to learn and 'to close the gap' on their understanding of subjects (Irons, 2008:36). However, students can be demotivated when the feedback is unfair, unclear, not understandable, not related with their work, not given on time, overly critical and not constructive (i.e. grades and marks are used) (Irons, 2008:36). Students' motivation and future learning is greatly affected by what we communicate to them about their learning, how we do it, and when we do it (MECY, 2006: 9).

Also, research evidence has shown the significant effects of students' participation in the assessment and learning process on their motivation to learn (Chappuis & Stiggins, 2002: 40; CCEA, 2007: 2; Stiggins, 2008:8-9). However, if students' needs are ignored due to incompetent assessment tasks, it will harm student confidence, motivation, self-efficacy and learning, as well as student and teacher efficacy (Stiggins, 2008:4). Students' emotional reaction toward assessment results affects their thinking, feelings and motivation to learn (Stiggins, 2006: 16).

Furthermore, teachers have a significant role to increase or decrease students motivation to learn (Cauley & Mcmillan, 2010: 4). Fautley & Savage, (2008: 91) articulated, “the teacher holds the keys to motivation”. Therefore, teachers should provide clear learning targets to their students; offer feedback about progress toward meeting learning targets; attribute student success and mastery to moderate effort; encourage student self-assessment and help students’ to set attainable goals for improvement to increase student motivation and learning through their formative assessment practice (Cauley & Mcmillan, 2010: 4-5). Each of the technique can enhance student motivation as well as achievement because through assessment, teachers can destroy or enhance students’ motivation to learn quickly and permanently than other tools (Stiggins, 2001: 36) as cited in (Harlen, 2006: 62).

Besides these, Nicol & Macfarlane-Dick (2006: 212) placed other important strategies for teachers to enhance high levels of motivation among students to improve their learning. These include: providing marks on written work only after students have responded to feedback comments; allocating time for students to re-write selected pieces of work; pretesting with feedback; and resubmissions. Therefore, the above review of related literature indicated that science teachers should create a positive classroom environment by implementing effective formative assessment strategies to enhance students’ motivation to learn to achieve the desired objectives.

#### **2.3.2.4 Formative assessment and students’ self-esteem, self- efficacy and self-regulation**

“The goal of assessment for learning is not to eliminate failure, but rather to keep failure from becoming chronic and thus inevitable in the mind of the learner” (Stiggins, 2007:25). Research evidences showed that formative assessment is vital to generate students’ confidence, high self-esteem, positive self-efficacy, self-regulation and motivation to learn and ensure all students an experience of success and make all of them competent learners (Nicol & Macfarlane-Dick, 2006: 199-208; Irons, 2008: 27-50; Moss & Brookhart, 2009: 5; Harlen, 2010: 40-42). When students become competent, they develop deep understanding, assess their own progress and others and able to direct their own learning (Irons, 2008: 27; MECY, 2008: 10) which are crucial to meet the challenges of 21th century. But, judgmental assessment types lead students towards anxiety

and threatening feelings about themselves as learners (Harlen, 2010: 40-42). Generally, building students' confidence, self-esteem and self-efficacy in their own skills and knowledge and their ability to manage their own learning is important to achieve the desired goals (CERI, 2005: 48). Therefore, under here, we can see the effect of assessment on students' self-esteem, self-efficacy and self-regulation

#### **2.3.2.4.1 Self-esteem**

According to Harlen (2006: 66), "self-esteem refers to how people value themselves both as people and as learners. It shows the confidence that the person feels in being able to learn." (Harlen, 2006: 66). Similarly, Fautley and Savage (2008: 86-87) defined, "a person's sense of self-worth." In other words, Self-esteem refers to "how one values oneself as a person and as a learner" (ARG, 2002: 3). However, there is a good deal of evidence on the power of assessment practices to hinder or cultivate students' sense of positive self-image which mutually affects their learning (Black & Wiliam, 2001; ARG, 2002: 3-10; Harlen, 2006: 64-66; Nicol & Macfarlane-Dick, 2006: 199-208).

Assessment practices have both positive and negative effects on students' sense of self (Black & Wiliam, 2001:6; Stiggins, 2009: 419). Assessment shapes how we see ourselves as learners and as people (Stobart, 2008: 145). And Stiggins (2009: 419) emphasized the impact of upper elementary assessment experiences on students' sense of self-worth in their future life because learners are the ultimate users of assessment information in order to improve their learning. The negative effects are related to assessment which focuses on grading, marking and ranking (Bell & Cowie, 2002: 15; Irons, 2008: 50). Students who, "encounter difficulties and poor results are led to believe that they lack ability, and this belief leads them to attribute their difficulties to a defect in themselves about which they cannot do a great deal" (Black and Wiliam, 2001:6). Therefore, such assessment activities induce low achievers to attribute their poor results to their abilities rather than efforts to strive in the next learning (Bell & Cowie, 2002: 15; Harlen, 2006: 66).

Consequently, these students develop a habit of low self-esteem. Mostly, judgmental feedbacks and repeated tests undermine their self-esteem and confidence (Black & Wiliam, 2001:6; Bell & Cowie, 2002: 15; ARG, 2002: 4 and 2003: 8) and do not allow them to use their time and efforts to challenge new tasks (Harlen, 2006: 66; Irons, 2008: 85) and negatively affects self-regulation, their motivation to learn and academic achievement (Nicol & Macfarlane-Dick, 2006: 209). CERI, (2005: 49 and 271) also added the negative effects of competition and social comparison activities in schools on students' academic self-esteem.

Whereas, the positive effects are related to assessment practices that actively involve students in the assessment processes. Black & Wiliam (2001:6) associated effective formative assessment to a powerful weapon to realize significant results in all students, especially, in low achievers. Because, assessment for learning emphasis on the learning process and on students' effort rather than on their abilities and dispositions (Stobart, 2008: 145). As a result, such assessment activities develop high self-esteem among all students (Black & Wiliam, 2001:6) and then they develop a habit of success or confidence on their ability to do any activity independently and effectively (Harlen, 2006: 66; Stiggins, 2006:15; Stiggins, 2007:25).

Reciprocally, students' levels of self-esteem have both positive and negative effects on their future learning. Students with low self-esteem become demoralized to learn, perceiving learning as threatening to the self, view all feedback as a judgment of their ability and met different problems in their learning (Brooks, 2002: 44; Harlen, 2006: 66; Weaver, 2006: 381-382). Whereas, students with positive self-esteem tend to be strongly motivated to learn, interpret feedbacks constructively, and use their efforts to challenge difficult and new tasks (Brooks, 2002: 44; Harlen, 2006: 66; Weaver, 2006: 381-382). Thus, they believe on themselves as able learners which intrinsically motivate them to take the risks of trying in future learning and as a result it leads them to be confident and develop a strong sense of academic self-efficacy (Stiggins, 2009: 419). To conclude, those students who have high self-esteem and confidence develop positive self-efficacy and become self-regulated learners.

#### **2.3.2.4.2 Self-efficacy**

Self-efficacy refers to “student’s perception/beliefs that she/he can learn particular content or skills and will be successful in doing so” (Moss & Brookhart, 2009: 82). It is associated with self-esteem (Fautley & Savage, 2008: 86-87) except self-efficacy “is more directed at specific tasks or subjects” (Harlen, 2006:66). Students’ self-efficacy will decrease if they fail at a particular task unless they have a very strong sense of self-esteem (Fautley & Savage, 2008: 86-87). Anderson & Bourke (2000: 35) explained, self-efficacy “It refers to how capable the learner feels of succeeding in a particular task or type of task” (cited in Harlen, 2006: 66). Moreover, Ormrod (2006) defined self-efficacy, as cited in Fautley & Savage (2008: 86):

An impression that one is capable of performing in a certain manner or attaining certain goals. It is a belief that one has the capabilities to execute the courses of actions required to manage prospective situations. Unlike efficacy, which is the power to produce an effect (in essence, competence), self-efficacy is the belief (whether or not accurate) that one has the power to produce that effect.

As many researchers, students learn more when they have greater feelings of efficacy about their academic abilities (Fautley & Savage, 2008: 86-88; Stiggins, 2009: 419). Hence, “Students with greater confidence in their ability to accomplish the target task are more likely to visualize success than failure” (Ross et al., 2002: 45 quoted in Moss & Brookhart, 2009: 82). Especially, the feedback system is vital (Nicol & Macfarlane-Dick, (2006: 208). Feedback enhances students’ confidence in their academic abilities when it is given constructively and openly manner. Thus, students become confident to ask and answer question, discuss their work with teachers and peers, and find out new strategies to solve problems they faced (Irons, 2008: 50). Similarly, Zimmerman & Schunk (2001) cited in King (2006:33), reported that there is high correlation between students’ self-efficacy and self-regulated learning and students’ motivational belief and their academic achievement.

Moreover, Albert Bandura (1994: 71) cited in Stiggins (2009: 420) stated that:

Strong sense of efficacy enhances human accomplishment and personal well-being in many ways. People with high assurance in their capabilities approach difficult tasks as challenges to be mastered, rather than as threats to be avoided. Such an efficacious

outlook fosters intrinsic interest and deep engrossment in activities. They set themselves challenging goals and maintain strong commitment to them. They heighten and sustain their efforts in the face of failure. They quickly recover their sense of efficacy after failures or setbacks. They attribute failure to insufficient effort or deficient knowledge and skills, which are acquirable. They approach threatening situations with assurance that they can exercise control over them. Such an efficacious outlook produces personal accomplishments, reduces stress and lowers vulnerability.

Hence, assessment practices have a key role to build a strong sense of self-efficacy (ARG, 2002:6; Harlen, 2006: 66). Moss & Brookhart (2009: 82) also specified the importance of developing the skill of self-assessment among students to develop self-efficacy. Thus, teachers should implement assessment practices that build a strong sense of academic self-efficacy among students to improve their learning (Stiggins, 2009: 420) through sharing learning targets and success criteria or standards, providing constructive feedback, using effective dialogue, and involve students in peer and self-assessment activities (ARG, 2002:6; Gioka, 2008: 146; Stiggins, 2009: 420). Since, as cognitive and social constructivist views of learning involving students directly and actively in the assessment process and in their own learning is crucial to develop self-efficacy (Chappuis & Stiggins, 2002: 40).

#### **2.3.2.4.3 Self-regulation**

Self-regulation, according to Harlen (2006: 67), “refers to learners' consciously controlling their attention and actions so that they are able to solve problems or carry out tasks successfully.” Similarly, Assessment Reform Group defined it “the capacity to evaluate one’s own work and to make choices about what to do next” (ARG, 2002: 3). As Moss & Brookhart (2009: 82), self-regulation is broader than self-efficacy. Pintrich & Zusho (2002: 64) defined, “Self-regulated learning is an active constructive process whereby learners set goals for their learning and monitor, regulate, and control their cognition, motivation, and behaviour, guided and constrained by their goals and the contextual features of the environment” (cited in Nicol & Macfarlane-Dick, 2006: 201).

Most self-regulated learners take responsibilities of their own learning and control it, actively engage in producing, collecting and interpreting assessment evidences (Nicol & Macfarlane-



Dick, 2006: 199); select, use and monitor strategies to close the gap in next learning; and assess their own progress (Harlen, 2006: 67; Moss & Brookhart, 2009: 82). Moreover, Nicol & Macfarlane-Dick, (2006: 204) added that self-regulated learners are more persistent, resourceful, confident, independent and higher achievers.

Hence, promoting self-regulation is important to help students to become effective learners (Nicol & Macfarlane-Dick, 2006: 204) and to improve their achievement (Harlen, 2006: 67; Moss & Brookhart, 2009: 82). Researchers put formative assessment followed by effective feedback as a powerful means to develop self-regulated learners (Nicol & Macfarlane-Dick, 2006: 199-211). And they proposed seven principle of good feedback practice for teachers to develop students' self-regulation. These are: clearly share the learning targets, standards and successes criteria; actively involve students in self-and peer assessment; provide constructive and timely feedback; inspire a discourse community; cultivate positive motivational beliefs and self-esteem by emphasizing on assessment for learning; provide opportunities for students to produce a piece of work and to modify it after external feedback to fill the gap in their learning; and use feedback information as an input to shape their teaching (p., 204- 211). Therefore, students must be provided the opportunity to develop positive self-esteem, self-efficacy and self-regulation skills which play an important role in preparing lifelong learners and promoting positive learning in future (Nicol & Macfarlane-Dick, 2006:212; Fautley & Savage, 2008:88).

### **2.3.3 Qualities of new modes of assessment**

The purpose of this section is to discuss the characteristics of new modes of assessment or formative assessment which makes it different from summative assessment. Formative assessment is different from other traditional forms of assessment, because formative assessment is integrated with the instruction and the real-life situations (authentic), and the assessment is conducted continuously throughout the instruction.

### **2.3.3.1 Integrated with instruction**

Traditionally, as behavioural learning theories assessment and instruction were viewed as separate activities in which classroom instruction is followed by assessment (Fautley & Savage, 2008:5-14) and only paper and pencil tests or exams are emphasized (Crisp, 2012:33-34).

However, currently assessment is viewed as an integral part of the daily instruction (Shepard, 2000: 5; Adair-Hauck et al., 2006: 360, Fautley & Savage, 2008:5-91) and an integration of different assessment methods to see the complete picture of students and to improve learning (Crisp, 2012:33-34). Biggs (1999) named it “constructive alignment” as cited in (Bloxham & Boyd, 2007:27). Single assessment method should not be conducted to achieve the main objectives of learning (Bass & Glaser, 2004: 8). Assessment system which is integrated with instruction improves learning than assessment of learning or high-stake summative tests (Twing, Boyle & Charles, 2010: 2-3).

Mostly, summative assessments or high-stakes exams are separated from real classroom learning (Shepard 2000:9) because these assessments are not aligned with daily classroom instruction, (Harris, McNeill, Lizotte, Marx, & Krajcik, 2006:69), therefore, they are less likely to be sensitive to improve student learning. Thus, in order to improve daily lesson and students’ learning, assessment practices should be connected to classroom instruction in seamless fashion (Adair-Hauck et al., 2006: 360).

Assessment and teaching should be compatible and act to support one another (Bloxham & Boyd, 2007:27) since good teaching is aligned with good assessment (Libman, 2010:63). Similarly, for Cowie (2012:687-688) curriculum, instruction, assessment, learning and students’ achievement are strongly linked and both of them are equally powerful to influence one another. Ample research evidences showed the significance of integrating assessment with teaching and learning to improve learning standards (Bass & Glaser, 2004: 7; Mentkowski, 2006:59; Boud & Falchikov, 2007:191 Bloxham & Boyd, 2007:5; Stobart, 2008:144; Fautley & Savage, 2008:91; Boud, 2009:37).

Such constructive alignment between assessment and daily instruction is vital to provide an accurate picture of learning (Harris et al., 2006:69; King, 2006:33), to assess higher order thinking (Higgins, Hartley & Skelton, 2002: 53), to reinforces consistent expectations for teaching and learning (Bass & Glaser, 2004: 8) and to prepare learners for lifelong learning (Fautley & Savage, 2008:91). Moreover, it increase students' self-efficacy and intrinsic motivation, and allows them to take a greater degree of responsibility for their own learning (Fautley & Savage, 2008: 93-94) and deeper students' understanding (Keeley, 2008:15-16; Wyatt-Smith & Gunn, 2009: 91). Equally, it “encourages students to take risks, think creatively, engage in reflection, and discover their own strengths” (King, 2006:33). Correspondingly, students have opportunities to actively involved in self-and peer- assessment; set learning standards; develop different strategies to involve in real-life problems, issues and situations; use feedbacks to fill the gaps in their learning; and engage in activities that have value in future (Crisp, 2012: 41).

As well, integrated assessment helps teachers to “extract and interpret information about what learners know and understand through careful observation, interactive dialogue, and varied expressions (e.g., writing and drawing) of scientific reasoning”, as a result “they make instructional changes to maximize the benefits of learning events” (King, 2006:32). Furthermore, Keeley revealed that:

Integrating assessment and instruction help teachers to learn more about what students need in order to be successful learners of science. Teachers can identify students existing ideas, charting, and monitoring learning paths that will eventually lead students to discover, understand, and use the important ideas and processes of science. (Keeley, 2008:15-16).

Generally, integrating assessment with daily instruction and using multidimensional assessment methods provides useful and comprehensive insights about students' progress (Fautley & Savage, 2008: 93-94), improves the quality of teaching and raises students' achievement by allows teachers to incorporate numerous formative assessment techniques (Keeley, 2008:15-16; Crisp, 2012: 33-41) and normally such assessments are practical and usable for students, teachers and researchers to adjust learning in the future. Hence, the effectiveness of assessment depends on its alignment with instruction and learning activities (Harris, et al., 2006:69; Irons, 2008: 13;

Suuramm, Koch & Arden, 2010: 400; Libman, 2010:63) and especially on formative assessment practices and the feedbacks it follows (Biggs, 1999 cited in Higgins, Hartley & Skelton, 2002: 53) since, formative assessment or assessment for learning is an integral part of the daily teaching learning process (Fautley & Savage, 2008: 93-94; Twing, Boyle, & Charles, 2010: 3-5). Therefore, clear and consistent alignment between assessments and daily instruction is important to enhance classroom learning.

### **2.3.3.2 Authentic**

Authentic assessment is a form of the new mode of assessment (Kubitskey, Fishman, Margerum-Leys, Fogleman, Brunvand, & Marx, 2006: 108; Kvale, 2007:66; Fisher & Frey, 2007: 114; Gulikers, Bastiaens, & Kirschner, 2008: 73). Authentic assessment refers to the alignment of assessment tasks to the classroom objectives and to the tasks to be mastered outside the school (Kvale, 2007:66). Equally, Gulikers, Bastiaens, & Kirschner, (2008: 76) defined authentic assessment as “an assessment requiring students to demonstrate the same (kind of) competencies, or combinations of knowledge, skills, and attitudes, that they need to apply in a criterion situation derived from professional practice” which reflects a real world application. Authentic assessment refers to “the extent to which the assessment of educational courses matches with the intended outcomes” (Murphy, 2006: 44).

Authentic assessment plays a great role to integrate curriculum and assessment (Buxton, 2006:179; Murphy, 2008:43). It is “more practical, realistic and challenging” (Torrance, 1995 cited in Bloxham & Boyd, 2007: 29). Authenticity depends on “the degree of correspondence between the assessment task and learning objectives” (Bachman & Palmer, 1996:23) as mentioned in (Murphy, 2008:43) and the “resemblance between the assessment and professional practice” (Gulikers, Bastiaens, & Kirschner, 2008: 83).

Traditionally, there is a large gap between instruction, assessment tasks and professional practice in the world of work (Gulikers, Bastiaens, & Kirschner, 2008: 73). But now, the culture of education was changed to competency-based education that focuses on the development of competent students and future employees (Segers et al., 2003) as cited in Gulikers, Bastiaens, &

Kirschner, 2008: 73). Therefore, authentic assessments are important to marry the curriculum and assessment tasks and as a result to achieve the stated educational goals (Murphy, 2006: 44-45).

Unlike traditional paper and pencil tests, authentic assessments engage students as active constructors of meaning (Murphy, 2008:43) and comprise a variety of innovative assessment methods to solve real-life problems (Falchikov, 2005: 70-71). Theoretically, authentic assessment was emanating from cognitive learning theories (Falchikov, 2005: 70). Thus, they are in line with the current views of learning (Murphy, 2008:43) which focuses on the 'constructive alignment between instruction, learning and assessment' (Biggs, 1996) as quoted in (Gulikers, Bastiaens, & Kirschner, 2008: 73).

Researchers pointed out the importance of authentic assessment to maximize students' learning (Murphy, 2006: 44-45; Murphy, 2008:43; Swaffield, 2011, 434) and particularly it helps low achieving students to express their confidence and competency in a range of skills ( Buxton, 2006: 178). Moreover, it is vital for lifelong learning (Falchikov & Boud, 2008:91; Gulikers, Bastiaens, & Kirschner, 2008: 73) and apprenticeship training (Kvale, 2008:201) because it has a progressive focus, and is true representations of real-life problems (Falchikov & Boud, 2008:91). Likewise, authentic assessment has a positive influence on students' motivation to learn and on their professional competency development (Gulikers, Bastiaens, & Kirschner, 2008: 73).

Additionally, authentic assessment improves students' engagement in their learning, self-regulation and achievement gains (Swaffield, 2011, 447). Since, students are interested on assessment activities which have value beyond completing the task and meaningful by themselves (Bloxham & Boyd, 2007: 27-29). Also, on the word of Swaffield (2011, 434-447) authentic assessment improves teachers' motivation and professional practice, the relationship among students and teachers and the culture of classroom environment.

Authentic assessment does not encourage rote learning and passive test-taking (Murphy, 2006: 44-45). Instead, it focuses on students' problem solving and analytical skills; ability to integrate and coordinate what they learn; creativity; ability to work collaboratively; and written and oral

expression skills. It values the learning process rather than the finished product (Birenbaum, 1996 cited in Gulikers, Bastiaens, & Kirschner, 2008: 75-76). This implies that authentic assessment requires students to demonstrate their knowledge and skills in meaningful context (Swaffield, 2011, 434).

Therefore, it should be designed by considering these different kinds of competencies and the philosophy of assessment for learning (Murphy, 2006: 44-45) to resemble with the current goals of education which emphasise on the development of competent future employees (Gulikers, Bastiaens, & Kirschner, 2008: 75-77). As a result, it helps students to apply or demonstrate their knowledge and skills in real-life problems (Gulikers, Bastiaens, & Kirschner, 2008: 73; Swaffield, 2011:434). According to Gulikers, Bastiaens, & Kirschner, (2008: 74-83) authentic assessment tasks have high degree of predictive, constructive and consequential validity than traditional tests.

Furthermore, Gulikers, Bastiaens, & Kirschner, (2008: 77- 81), pointed out five theoretical dimensions of authentic assessment. These are: *the nature of task*: tasks in the authentic assessment should require students to integrate knowledge, skills and attitude and resemble to real-life tasks (p., 79). Authenticity depends on “the nature and context of assessment tasks that the students are required to perform” (Swaffield, 2011:434); *the physical context*: authentic assessment needs available resources in the school (i.e. equipment, safe and clear classrooms, time etc...); *the social context*: assessment tasks should resemble to the social process in reality (p., 80); *the assessment form*: teachers should implement a wide variety of assessment methods rather than single tests to address real-life problems (Kubitskey et al., 2006: 108; Buxton, 2008: 178); and *the criteria*: it refers to the development of assessment criteria which are explicit and transparent to the learners beforehand and built from real-life criteria for the development of professional competencies (P., 81). Generally, according to them, the authenticity of any authentic assessment activity depends on these five theoretical frameworks. Therefore, assessment activities in science subjects should be authentic since students need to learn how to apply their knowledge and skills in the real-life context in the ever changing world.

### **2.3.3.3 Continuous**

Formative assessment is a continuous process which is used to gather evidence about students learning minute-by-minute to identify the next step in learning (Bell & Cowie, 2002: 6-23; Sebba, 2006: 194-195; Stiggins, 2006:15; King, 2006: 31; Stiggins & DuFour, 2009:641). Likewise, formative assessment is an ongoing process and focuses on improving students learning rather than simply testing and grading (Henriques, Colburn, & Ritz, 2006: 15; Greenstein, 2010: 29).

If assessment is given in continuous manner, it allows teachers to use a variety of assessment methods to assess a more holistic picture of student performance (Ottevanger, Akker, & Feiter, 2007: 20) and to give guidance for needy students (Shepard, 2000: 8) on a regular basis and thus it provide input for future learning process (Martínez, Stecher, & Borko, 2009: 81). Moreover, it gives evidence to the learner not only their current achievement but also improvements in their capabilities (Stiggins, 2008: 3).

Students who are engaged in this ongoing metacognitive experience are also able to monitor and scaffold their learning (Shepard, 2000: 8), make corrections, and develop a habit of mind for continually reviewing and challenging what they know (MECY, 2008, 10). In general, continuous assessment gives students a greater control over their achievements and progress in learning (Hofstein, Mamlok, & Rosenberg, 2006: 145; MECY, 2008, 10) and “a way of disciplining students and forcing them study from day one” (Lauvas, 2008: 163).

Continuous assessment “reduces high rates of repetition and dropout and as a result it will also save resources being wasted for more productive uses and it maintains both the quality of learning as well as the desired efficiency coefficient” (MoE, 2005:30-31) because, if tasks are designed well, continuous assessment is important to capture student attention and effort, generate appropriate learning activity and provide constructive feedback for students. (Lauvas, 2008: 164). Therefore, we must use continuous classroom assessment to help students to take responsibility for their own academic success and to improve learning (Stiggins, 2004:25).

Others also noted the importance of continuous assessments to provide valid, reliable and greater volume of evidence about students' knowledge and abilities (Gradner, 2005: 202; Hofstein, Mamlok, & Rosenberg, 2006: 141-146; Wiliam, 2008:276). Assessments that employed a variety of methods continuously will yield valuable information about students' learning (Harris et al., 2006: 70). Assessment becomes powerful to improve learning when it is frequent, cyclical, dynamic, and sensitive to individual difference and occurs constantly throughout the instruction (Gradner, 2005: 202; King, 2006: 31; Wiliam, 2008:276; MECY, 2008: 10). Equally, Black & Wiliam (1998) provided the powerful and positive effects of ongoing assessment combined with appropriate feedback on students' science achievement as cited in (National Research Council (NRC), 2001: 226).

As well, assessment which reveals the next step in learning is "a powerful booster of confidence and motivation' for students as well as for teachers whereas assessment that focuses on judgmental feedbacks exposed students to be confused, frustrated and hopelessness" (Stiggins, 2008: 3). That is why teachers are recommended to use continuous and a variety of assessment methods in science and mathematics subjects: such as classroom dialogue, questioning, seatwork and homework assignments, formal tests, less formal quizzes, projects, portfolios, and so on to improve students' overall performance (NRC, 2001: 226).

For this reason, most countries in Sub-Saharan Africa have introduced continuous assessment as a new ways of assessment to improve students' learning (Ottevanger, Akker, & Feiter, 2007: 11) because introducing continuous assessment in addition to final examination is important to improve learning (Lauvas, 2008: 163). Furthermore, teaching standards in science education gives high emphasis on continuous assessment of students' understanding (National Research Council (NRC) 1996: 52) as cited in (Farenga, Joyce, & Ness, 2006:42).

Thus, assessment must be comprehensive, continuous and integrated into the learning cycle to provide evidences and to know what comes next in the learning rather than given at the end of teaching (Shepard, 2000: 8; NRC, 2001: 256; Bell & Cowie, 2002: 6-7; CERI, 2005: 198; Sebba, 2005: 194-195; Stiggins, 2006:15; Stiggins, 2008: 3- 4; Alkharusi, 2008: 248;Stiggins & DuFour, 2009:641; Tucker, Fermelis, and Palmer, 2009: 174; Martínez, Stecher, & Borko, 2009:



81; Greenstein, 2010: 153) because, as Heritage (2007:140-145) assessment becomes formative if it used various methods to gather evidence about students' learning continuously throughout the instruction.

However, using summative tests or paper and pencil tests continuously is not formative assessment (Bell & Cowie, 2002: 8). In such repetitive summative assessment students may achieve high marks but their learning is not deepened to apply it in real situation (Race, 2007:35-37). Similarly, continuous summative assessment encourages cheating, memorization and regurgitation of facts and rote learning and greatly weakened and lost students' motivation to do more and their legal rights in learning (Lauvas, 2008: 164). Moreover, such assessment system is important only for few students who have good memories of facts, those who work well under pressure and those who like leaving things to the last minute (William, 2008: 276).

To be formative, continuous assessment must show the current progress of students and what comes next in the learning or it should provide evidence for students and teachers to fill the gaps in learning (Stiggins, 2008: 4). According to Bell & Cowie (2002: 8), assessment is "considered as formative assessment if some action to improve learning during the learning was involved." Therefore, what makes any particular assessment formative is not the specific assessment tool employed continuously but how the information gathered from the tool is used to improve learning and to adjust instructional strategies toward the learning goals.

However, in most Sub-Saharan Africa countries, "assessment at the school-level is very much summative in nature, and is hardly used for instructional purposes or to provide feedback to the learners to fill the gaps in the future" (Ottevanger, Akker, & Feiter, 2007: 20). Practically, continuous assessment squeezes formative assessment out of the instruction by adding more intensive summative assessment tests and quizzes (Lauvas, 2008: 164).

Mostly, assessment was used as tool to judge and rank students and schools rather than considered as a source of evidence to improve students' learning (Heritage, 2007: 140). In many cases, teachers and students do not use assessment evidence as an input for future learning (Henriques, Colburn, & Ritz, 2006: 15). However, students should be given continuous

assessment feedback rather than judgmental feedback to see their improvement over time and to understand what comes next in their learning and to develop positive belief about themselves (NRC, 2001: 256-259; Stiggins, 2007:25; Stiggins, 2008: 4; Alkharusi, 2008: 248). To be effective, assessment must be continuous, integrated, and dynamic and yield accurate information for next steps in learning. Generally, teachers should use ongoing, frequent and varied assessment methods and pertinent feedback collaborates with their students to move learning forward. Therefore, teachers can observe and understand students existing belief, knowledge, and thinking overtime and can easily identify gaps.

### **2.3.4 Inhibiting factors for effective practice of assessment for learning**

The objective of this section is to explain the major inhibiting factors to implement formative assessment in secondary school science subjects. The section includes the general overview and factors which affect the implementation of assessment for learning such as teachers' perception of science, learning and assessment, their professional learning, the burden of summative tests, and the school support system and the availability of resources.

#### **2.3.4.1 Overview of the inhibiting factors**

As indicated in the previous section formative assessment is one of the most effective strategies to improve students' learning, to insure equity and to develop students' learning to learn skills. However, it is not straightforward to achieve the vision of formative assessment. "There is no 'quick fix' that can be added to existing practice with promise of rapid reward" (Black, et al., 2003:15). Similarly, "formative assessment is not some kind of passing bandwagon that we need to be seen to jump on to be proved 'with it', or up-to-date" (Lambert and Lines, 2001: 130). Mostly, policy makers viewed assessment changes as simple practice (Hawe, 2002, cited in Orr, 2008: 140).

However, they should implement different strategies to practice formative assessment effectively because policy shift cannot be easily implement as it is stated (Orr, 2008: 140). Especially, it is difficult to implement formative assessment when the curriculum is over-detailed and rigid, and

assessment is driven by high-stakes tests (Stobart, 2008: 168). In such cases, formative assessment becomes a series of marking that used to improve grades rather than a way to improve students' deep understanding. Particularly, in secondary school science subjects formative assessment was not practiced effectively to achieve its objective (CERI, 2008: 3). Therefore, the assessment practices fail to serve the underlying values to be formative assessment (Swaffield, 2009: 9).

Ample research evidences conducted in UK and many other countries including Africa showed the inadequacies of teachers' assessment for learning practice in science subjects (Lambert & Lines, 2001:133-135; Black & Wiliam, 2001: 4-6; Black et al., 2003: 10-11; Ottevanger, Akker, & Feiter, 2007: 19; Kvale, 2007: 65; Black & Wiliam, 1998 as cited in Fautley & Savage, 2008:92; Fisseha, 2010: 103-105). According to them, the shortcomings of teachers formative assessment practice includes: focuses on tests which encourage rote and superficial learning ; over-emphasized on marks and grades rather than the quality of students' learning; focus on sorting and ranking of students; focus on judgmental feedback which lower the self-esteem of low achiever students at the expense of giving useful advices for improvement; de-moralize students by focusing on competition among students rather than on personal improvement; give feedback which serves social and managerial purposes rather than helping students to learn more effectively; work with an insufficient picture of students' learning needs; priorities given to record marks rather than using them as an input for future learning; and give less attention for students previous performance.

Besides, studies suggest the possible reasons for such classroom assessment practices. Accordingly, the main inhibiting factors of teachers' assessment for learning practices are: teachers' view of science learning and assessment (Harlen, 2003:11; Graue & Smith, 2004:25; Stobart, 2008: 168; Mansell, James, & ARG, 2009: 18-19; Carless, 2005 cited in Hill, 2011: 348); the pressure of external high stake tests (CERI, 2005: 24; Dysthe, 2008: 15; Falchikov & Boud, 2008: 92-93; CERI, 2008: 3; Mansell, James, & ARG, 2009: 18-19; Carless, 2005 cited in Hill, 2011: 348); fear of resource and time pressure (Harlen, 2003:11; Graue & Smith, 2004:25; CERI, 2005: 24; CERI, 2008: 3; Mansell, James, & ARG, 2009: 18-19; Carless, 2005 cited in Hill, 2011: 348); lack of professional development opportunities in pre-service and in-service

trainings (Harlen, 2003:11; Ottevanger, Akker, & Feiter, 2007: 19; Mansell, James, & ARG, 2009: 18-19; Carless, 2005 cited in Hill, 2011: 348); teachers self-efficacy on their ability (Carless, 2005 cited in Hill, 2011: 348); lack of intensive support (Harlen, 2003:11; Graue & Smith, 2004:25; Mansell, James, & ARG, 2009: 18-19; Carless, 2005 cited in Hill, 2011: 348); lack of consistency between assessments and evaluations at the policy, school and classroom levels (CERI, 2005: 24; Ottevanger, Akker, & Feiter, 2007: 19; CERI, 2008: 3); barriers within students (related to students past experience on assessment); lack of alignment between curriculum, assessment and future learning (Falchikov & Boud, 2008: 92-96); and lack of alignment between learning theories and assessment practices (Dysthe, 2008: 15). Thus, in the following section we discuss about the key constraints of secondary school teachers' formative assessment practices in science education.

#### **2.3.4.2 Teachers professional learning**

Nowadays, considerable research evidences reflected teachers' quality as the crucial agent of quality education than any other factors in education (Yung, 2006: 1; Moss & Brookhart, 2009: 9), So professional learning to increase the quality of teachers is vital in educational reform (Harlen, 2010:127). Accordingly, professional learning is one of the key elements to ensure meaningful and sustainable change in teachers' assessment practice (Moss & Brookhart, 2009: 13; Harlen, 2010: 100; Gardner, 2010: 8). That is why, Assessment Reform Group in their pamphlet declared, teachers' professional development as one of the fundamental principles in effective implementation of assessment for learning (ARG, 2002:2).

So, school administrators and officials should plan sustainable and ongoing professional development programmes around the key aspect of teaching and learning in pre-and in-service trainings (Harlen, 2003: 40; Stiggins, 2009: 421) to improve the quality of teachers (William, 2006 : 287) and to achieve the desired purpose of formative assessment (William et al., 2004:49). Particularly, evidences showed that professional learning which focused on formative assessment has a greatest positive impact on teachers' teaching practice and students' achievement (William; 2006: 287; Moss & Brookhart, 2009: 9-13; Stiggins, 2009: 421; Harlen, 2010:127; Hill, 2011: 348) because, formative assessment is a powerful means to achieve quality education than

traditional testing (William et al., 2004:49) and hence it depends largely on teacher's professional knowledge, skill and experience (Bell & Cowie, 2001: 548; Harlen, 2010: 101). Professional learning is vital to identify and use a variety of assessment methods, gather evidence based on the stated objectives, interpret the evidence in line with the learning standards and to take action, to fill the gaps in future learning.

However, higher officials in many countries do not give credit for teachers' professional learning that would enable them to acquire the necessary skills to implement formative assessment in science subjects effectively (Yung, 2006: 3). Teachers and new school administrators "come into the profession with little or no pre-service preparation in assessment, let alone assessment for learning" (Stiggins, 2009: 421). Study by Moss & Brookhart, (2009: 9) also indicated that most teachers implement formative assessment activities blindly and students too because they are ordered to adopt a program or technique prescribed and presented by outside experts in a one-shot workshop. For these reason, teachers frequently use a series of tests which are simple to score and assess students' memorization of facts rather than scientific concepts (Kubitskey et al., 2006: 108; Moss & Brookhart, 2009: 13-15). Yet, to really know how and what students learn from their science classroom activities, teachers must use a variety of assessment tools (Kubitskey et al., 2006: 108).

According to Moss & Brookhart (2009: 13-15), most teachers misunderstood formative assessment as: a series of tests administer to audit students' learning; a program or method that they must adopt and add to what they already do; and a way of improving a program rather than a way to improve their day-to-day instructions. As stated by them, such misconception can dilute the effectiveness of formative assessment and block its consistent use in improving students' learning. It leads teachers into dilemmas in assessment for learning practices (Stobart, 2008: 144). As a result, it becomes challenging to develop self-regulated learners who are crucial for today's society demand.

But, evidence from USA by Khattri et al., (1998: 144) noted that "new approaches to student assessment alone are not sufficient to improve teaching and learning. Rather, the principles and ideas underpinning assessment reform must be clearly defined and understood at all levels"

quoted in (Yung, 2006: 3). To be effective assessment change demands much more than simply adding a set of new technique to the existing practice (Harlen, 2010: 102) because formative assessment is not a series of tests and a prepackaged program or set of techniques that teachers adopt and enact rather it is an integral part of the teaching learning process and that uses comprehensive method of assessment to gather evidences about students daily learning by engaging both teachers and students targeted to improve learning (Moss & Brookhart, 2009: 13-15).

Hence, to effectively implement formative assessment in secondary school science subjects teachers need to get training on how to: plan formative assessment in their daily instruction; share learning targets in understandable ways; gather evidences using different assessment strategies (observation, listening, questioning...); plan strategic questions throughout the lesson; develop self-and peer assessment among students; interpret assessment evidences in relation to learning goals or criteria; provide descriptive, timely, clear and task specific feedback; and show students where they are in relation to the goal and what they should do next to close the gap (Jones, 2005: 3; Moss & Brookhart, 2009:19; Stiggins, 2009: 421; Baird, 2011:344). Moreover, studies by Harlen, (2003: 39), indicated that to improve students' performance in science subjects, teachers need to understand the basic elements in the formative assessment cycle and implement them effectively. Especially, in science subjects concepts are interconnected and complex, then teachers need to learn how to integrate new modes of assessment in their daily instruction (Kubitskey et al., 2006: 108; Harlen, & Gardner, 2010:23). Since, teachers' deep understanding is important for new practice to be successful, unless "it is all too easy for new practices to fade away" (Harlen, 2010: 102).

Evidences showed that teachers who are engaged in appropriate professional learning can implement assessment for learning successfully (Mansell, James, & ARG, 2009: 19); actively involve their students in the assessment process (Chappuis & Stiggins, 2002: 43); share their experiences and create a community which feel comfortable (Kubitskey et al., 2006: 114-115); took responsibility for what happened in their classrooms; and did not blame external circumstances to accept and implement the change (Swaffield, 2009: 9).

Hence, higher officials and school leaders should provide opportunities for secondary school science teachers to upgrade their professional knowledge and skills in formative assessment through discussion with their peers, observing best practices, visiting other schools, assessing their own practice with action research; participating in conferences and workshops, reviewing other related research works (Harlen, 2003: 38; Akerson, McDuffie, & Morrison, 2006: 161; Stiggins, 2009: 421; Bennett & Gitomer, 2009: 55; Harlen, 2010: 101-116) and giving professional courses founded on formative assessment in pre-service and in-service trainings (ARG, 2003: 1; Harlen, 2003: 39; MoE, 2005: 13; ARG, 2005 cited in Montgomery, 2010: 64). Similarly, other researchers mentioned ongoing regular conversations with teachers about formative assessment as the most powerful means to develop professional competence (Moss & Brookhart, 2009: 13; Harlen & Gardner, 2010:26).

Therefore, there is a need to make profound change in the traditional pre-service and in-service training in order to achieve the objective of teachers' professional learning in formative assessment in science teaching (Bennett & Gitomer, 2009: 55). Moreover, Cumming & Wyatt-Smith, (2009: 1), indicated the requirement of high attention on assessment courses in teacher-preparation programs. Based on these facts, teachers' pre-service training curriculum was revised and various types of in-service trainings were introduced in Ethiopia to improve the quality of teachers in implementing continuous assessment effectively (MoE, 2005: 11-14).

To bring effective change in assessment practices both pre-service and in-service trainings should be given continuously over a period of time rather than a day workshops (McWaters & Good, 2006: 53); actively and cooperatively involves teachers in the development of new procedures or practices or changes (Harlen, 2010: 101-105); and focus on deepening teachers' knowledge of the topic and different strategies that they used in their lesson (Bell & Cowie, 2001: 548; Akerson, McDuffie, & Morrison, 2006: 160; Wiliam, 2006: 287). Therefore, school leaders should take an active role to build teachers understanding of formative assessment through providing different opportunities (Moss & Brookhart, 2009: 13-15; Harlen, 2010: 101). Hence, secondary school teachers must get the required training on how to implement assessment for learning in science subjects.

Moreover, school leaders, practitioners, students, policy makers and parents also need training to understand the fundamental principles and features of assessment for learning to undergo the change effectively in secondary school science subjects (Bennett & Gitomer, 2009: 55; Swaffield, 2009: 14). Since, successful implementation of assessment change needs capacity development among all stakeholders in the education system (Nusche, Radinger, Santiago, & Shewbridge, 2013:15). As a result, everybody becomes committed to undergo the change smoothly by strengthening, developing and supporting secondary school science teachers for further understanding of assessment for learning and implementation.

Therefore, teachers' professional knowledge and skill in assessment activities should be a key question for policy makers and researchers (Shepard, 2000: 10-13; Nusche et al., 2013:15) to achieve the desired goals in the field of science education. Finally, and perhaps most importantly, we all need to look for ways to bridge the gap between theory and actual practices in science classrooms by engaging in collaborative studies between classroom teachers and researchers.

#### **2.3.4.3 Teachers' perception towards assessment for learning**

Research evidences showed that teachers' perception and belief about how students learn and knowledge are constructed strongly affects their classroom practice, particularly on the implementation of formative assessment in science education (Harlen, 2003: 34-36; Black et al., 2003:13-14; Firestone, Monfils, Schorr, Hicks, & Martinez, 2004: 67; Shepard, 2000:5; Yung, 2006: 3; Cheung et al., 2006: 210; Weaver, 2006: 380; Moss & Brookhart, 2009:13-18; Gulikers, Biemans, Wesselink, & Wel, 2013:122) because, such teachers' perceptions and beliefs are developed from their personal experiences and assumptions about the nature of subject, teaching and learning process (Dysthe, 2008, 19- 22). Moss & Brookhart, (2009:13-18) also indicated the effect of teachers perception on their instructional decisions on student achievement and motivation to learn and on the whole learning process.

Educators who strongly advocated behavioural learning theories viewed science learning as passive and believed that knowledge exists separated from the student, and it is built up from



pieces of simple, discrete and irrefutable information or facts that students have to memorize, which can then be built into more complex units (Shepard, 2000:4-5; Black & Wiliam, 2001: 9; Harlen, 2003: 34). The assumption is that the teacher is the source of knowledge and knows best whereas students are empty to be filled by the teacher (Harlen, 2003: 34). In the behavioural learning perspective, students' views, prior knowledge and experiences have no value, and so there is no need to gather evidence about their learning. However, students are not passive receivers of information (Black et al., 2003:14). Formative assessment is effectively implemented and learning gains is substantial when students actively involved in identifying the learning gaps and take action to fill this gap (Black et al., 2003:14).

Thus, those teachers who are more attached to this perspective might have developed negative perceptions or misconceptions about formative assessment or assessment for learning (Shepard, 2000:5). In such circumstances formative assessment is hardly implemented (Black & Wiliam, 2001: 9). Besides, Stiggins, (2009: 421) identified common wrong beliefs that prevent teachers from using assessment for learning. These wrong beliefs include: viewing standardized tests as the main determinant of school effectiveness; assuming report cards and grades as the vital purpose of classroom assessment; and seeing assessment as teachers-centered. That is why; teachers' current assessment conception is detached from the current views of assessment, which focuses on assessment for learning (Gulikers et al., 2013:122). Hence, teacher should realize assessment as an integral part of instruction, student-centered and a powerful means to improve science education (Cowie, 2012:684).

The other main issue that inhibits formative assessment practice in science education is “the beliefs that teachers hold about the potential to learn of all of their pupils” (Black & Wiliam, 2001: 9). When teachers develop beliefs that intelligence or ability of each student is innate or inherited or fixed, then it affects not merely their classroom assessment practices but also students' beliefs about themselves (Black & Wiliam, 2001: 9; CERI, 2005: 48). But, to implement formative assessment effectively and as a result to improve science learning in secondary schools, teachers need to perceive or develop a belief that everyone can learn and the importance of effort rather than ability to achieve the desired goals (Stiggins, 2005:326). Therefore, teachers need to believe on all students' capability and the importance of effort to

achieve a certain level of academic objectives, and then they need to use different instructional method and advise all students toward the attainment of learning standards.

Generally, teachers' views and beliefs about the subject matter, teaching, learning, assessment, and students and their relationships influence instructional change (Firestone et al., 2004: 67; Cheung et al., 2006: 208) and especially their formative assessment practice (Harlen, 2003: 36; Cowie, 2012:684). Moreover, a study conducted in Ethiopia on teachers' and students' perception of continuous assessment showed the impact of teachers' perception and belief on their assessment practice (Yigzaw, 2013:1490).

Therefore, "deep changes both in teachers' perception of their own role in relation to their students and in their classroom practice is important to implement classroom assessment changes effectively" Black et al., (2003:13). Similarly, Harlen, (2003: 34-36) and CERI, (2005:48), revealed the requirement of a change in teachers' perception and beliefs towards science education, students, how their students learn and assessment to implement formative assessment in science subjects successfully.

#### **2.3.4.4 Pressure from summative assessment**

Nowadays great emphasis is given for assessment especially in science education (Cumming & Wyatt-Smith, 2009: 2) because of its strong effect on how teachers teach, assess and how students learn (Dysthe, 2008: 15; Nusche et al., 2013:14). Even though, both formative (assessment for learning) and summative (assessment of learning) are valuable and have different purposes in formal education, it is not always easy to get the balance between the two purposes right in real classroom practices (Harlen, 2003: 20-31). Traditionally classroom assessments focus on summative assessment which is valuable to make judgment about students' achievement and to report it to others. Therefore, there is always a tension among formative and summative assessment in the classroom assessment practice when teachers and students used the same assessment to serve both functions (Falchikov & Boud, 2008: 90).

A great deal of research evidence in secondary schools revealed the negative effects of traditional forms of summative (high-stakes) assessment on the teaching and learning process and especially on formative assessment in science education (Shepard, 2000:9; Goodrum, Hackling, & Rennie, 2001: 22; Harlen, 2003: 10; Wiliam et al., 2004:49; CERI, 2005: 6; Stobart, 2008: 159; Hill, 2011: 34; Baird, 2011: 344-345). This happens because; in many countries secondary school assessment has long been dominated by traditional forms of summative assessment with end-of-topic, end-of-semester and end-of-year school tests. So, such modes of assessment have a great backwash effect on science education.

The negative effects of summative assessments on secondary school science education were summarized by different authors as follows: learning focuses on isolated facts and skills that are easy to test; teaching mainly focuses on the test content; students focus their learning on what they think they will be tested on; conventional teacher-centered methods becomes dominant; teachers devote considerable time to prepare students how to pass and get good examination results; assessment depends only on paper and pencil multiple choice items rather than using a wider range of assessment tasks; practical laboratory works and understanding of higher order basic scientific concepts are omitted (Goodrum, Hackling, & Rennie, 2001: 22-23; Firestone, Schorr & Monfils, 2004: 10-11; Stiggins, 2006:6; Walport, Goodfellow, McLoughlin, Post, Sjøvoll, Taylor, & Waboso, 2010: 43-44); assessment activities are not aligned with real problems; curriculum contents and students' learning experience were narrow down to test topics (Firestone, Schorr, & Monfils, 2004: 10; Harlen, 2006: 79; King, 2006:33); students motivation to learn (Harlen, & Gardner, 2010:23-24), confidence (Harlen & Crick, 2003, 170-171; Falchikov & Boud, 2008: 92-93), emotions and self-esteem becomes damaged (Broadfoot, 2008: 213); and are useless for measuring the affective and psychomotor domains of learning and higher order thinking (Harlen, 2003: 10; Race, 2007:35).

Moreover, summative tests provoke anxiety and science phobia among students (Harlen, 2006: 79; Falchikov & Boud, 2008: 92-93) and students become reluctant to participate in any change and fail to recognize formative feedback as a helpful signal and guide (Black et al., 2003:14). Also, for Shepard (2000:9) and others who are named in her study, high stake external tests leads teachers to “de-skilling and de-professionalization” or as Shepard “to the denigration of

teaching.” As a result, teachers become demoralized and lose their freedom to use their knowledge and skills (Goodrum, Hackling, & Rennie, 2001: 22-23; Firestone, Schorr, & Monfils, 2004: 10).

Reasons which aggravate the tension of summative assessment today include: national or local requirements for certification and accountability (Black et al., 2003:14; Cumming & Wyatt-Smith, 2009: 1-2); great attentions from professional bodies and government agencies to summative than to formative assessment (Black & Wiliam, 2001: 5; Wiliam et al., 2004:49; Harris et al., 2006:84; Walport et al., 2010: 43); great interest of employers on students grade to recruit (Falchikov & Boud, 2008:90- 92); socialization of teachers and students with the traditional assessment practices for a long period of time (Dysthe, 2008: 15); “a policy context that encouraged rushed curriculum coverage, teaching to the test and a tick box culture” (Swaffield, 2009: 9); evaluation of teachers efficiency largely based on students results from standardised tests (Nusche et al., 2013:14); and “dominant theories of the past continue to operate as the default framework affecting and driving current practices and perspectives” (Shepard, 2000:4) which was a combination of behaviorist learning theories, hereditarian theories of intelligence, curriculum theories and scientific measurement theories (Shepard, 2000:5; Dysthe, 2008, 19- 22). Generally, that is why, currently in many countries summative assessment becomes the dominant classroom assessment practices.

Despite the research evidence, hence, it becomes challenging and difficult to bring new modes of assessment into practice to promote learning (Dysthe, 2008: 15; Cumming & Wyatt-Smith, 2009: 1-2; Cowie, 2012:683). Mostly, teachers tend to use tests that encourage low cognitive level activities such as recall of isolated items of knowledge, which is unreflective and not aligned to the learning objectives; focus exclusively on what is tested to bring high scores (Shepard, 2000:4-9; Harlen, 2003: 10; Harlen, 2006: 79), and they “teach to the test” (Harlen & Gardner, 2010:23-24) rather than “teaching for understanding” (Schauble & Glaser 1996; Stiggins 1999; Wiggins 1998 quoted in King, 2006:33) and “giving less attention to students’ wider developmental and educational needs” (Nusche et al., 2013:14). All the mentioned challenges will occur particularly when high stakes accountability pressures are added (Wiliam et al.,

2004:49; King, 2006:33; Harlen & Gardner, 2010:22-24) and are used to evaluate teachers and schools effectiveness (ARG, 2003:1; Harlen, 2006: 79).

Specifically, in Ethiopian the secondary school educational system is highly directed by the national examinations given at the end of each year. Hence, what goes on in the classroom is largely dictated by what happens in the public examination halls, because, as stated in the Ethiopian Education and Training policy, admission to higher institutions is only based on students result on this paper and pencils multiple choice public examinations (MoE, 1994: 18-19). Moreover, many of the reference materials for secondary school science subjects which are available in the current market are mainly prepared based on the items of tenth and twelve grade public examinations. Unknowingly, such test oriented reference materials negatively affect the study habits, intentions and motivation of students in secondary schools and the teaching learning process in science subjects in general.

As stated by Shepard (2000:9); Harlen, (2006: 79), however, such powerful pressure on summative assessment will rise students result in the public examinations, but students lack the skills, competences and understanding of basic scientific concepts in science which are crucial in the 21st century. Harlen & Crick, (2003, 170), attached the increment of students test score over time with greater familiarity of teachers and students with the test items rather than because of their learning was improved. Furthermore, Hodson (1992:120) cited in (Goodrum, Hackling, & Rennie, 2001: 23) strongly criticized the traditional forms of assessment which measures isolated skills in the practical aspects of science as: “philosophically unsound (because it is not science), educationally worthless (because it trivializes learning), and pedagogically dangerous (because it encourages bad teaching).” Therefore, policy officials, governments, school administrative and other practitioners should recognize and minimize the negative effects of high-stake summative tests on science education and consider the importance of new modes of assessment to assess students overall progress and to improve their learning (Shepard, 2000:9; CERI, 2005: 6; Harlen, 2006:79-80; Nusche, et al., 2013:14). Moreover, Broadfoot (2008: 224) considered the effect of summative assessment as a disease and stated it in short as follows:

If we have not yet found a cure for this disease, we are nevertheless beginning to make progress. But, as with all diseases, the first step is to recognize that there is a problem. As such recognition develops, so the search for a cure will become more urgent and widespread.

Since, “effective, fit-for-purpose assessment is central to the outcomes and reputation of high quality education” (Walport et al., 2010: 43). Generally, to prepare self-regulated learners, creative thinkers, lifelong learners and informed citizens, who are pivotal for economic, social and political development of a country, we all must give high attention to ensure formative assessment practice in science education. Therefore, to strengthen the practice of formative assessment in science subjects, policy makers, school administrators, teachers and educational researchers should work together, until it strongly roots and bring the new learning culture with it.

#### **2.3.4.5 Support from school principals and supervisors**

School principals and supervisors are the prominent figures in secondary school to lead the teaching learning process effectively. Thus the effective implementation of any educational changes in secondary school depends on the effectiveness of principals and department heads. The support of school principals is very important to embed a wide variety of changes including assessment changes in the teaching learning process (Firestone et al., 2004: 66; Harlen & Gardner, (2010:23). Specifically, Harlen (2003: 39), emphasized the importance of school principal and supervisors support for teachers to implement inquiry based formative assessment in science learning. Providing substantial support is very important for the attainment of formative assessment in science education (Black & Wiliam, 2001: 6; Ottevanger, Akker, & Feiter, 2007: 20). Similarly, an exploratory study on the barriers and enablers of formative assessment practice conducted by Hill (2011: 347-360) in New Zealand secondary schools, showed the relevance of school principals’ and senior staffs support for teachers to implement assessment for learning in mathematics and science education effectively. Moreover, a study conducted in the eight countries by CERI (2005: 5-7), revealed the importance of policy makers’ and school leaders’ support to solve practical barriers in the use of formative assessment in secondary schools.

Hence, school principals and supervisors should be committed, responsible, and professional and have positive attitude toward the assessment change to support teachers to harmonize it in to the school culture. Ample research evidences mentioned different strategies for school principals and administrators to support teachers' to implement formative assessment effectively in science education (Black et al., 2004: 20-21; CERI, 2005: 5-7; Ottevanger, Akker, & Feiter, 2007: 20; Mansell, James, & ARG, 2009: 19; Moss & Brookhart, 2009: 18-20; Swaffield, 2009: 9; Hill, 2011: 350). According to them, the strategies include: using formal and informal discussions, using well-chosen starter statements that encourage shared inquiry, avoiding supervisory and judgmental perspectives, focusing on teacher's professional learning needs and provide opportunities for continuous professional in-service training, promoting collaborative work among teachers, encouraging mutual observation and the sharing of ideas and experiences within schools and with other schools, giving opportunity for teachers to trying out and evaluating new ideas, giving exemplary materials, giving time for teachers to meet on regular basis and providing opportunity to report their work on staff meetings. To realize all of these, school principals and supervisors should get appropriate and continuous professional training that enables them to give the necessary support to teachers for the effective implementation of formative assessment in secondary school science subjects.

#### **2.3.4.6 Resource availability**

The organization and the availability of appropriate instructional material, sufficient time, well-equipped classrooms and laboratories, reduced class size, and qualified teachers are important to all secondary schools to achieve high level of learning outcomes in science by inspiring more activity based, inquiry-oriented science in a safe environment because according to Walport et al. (2010: 43), "it is impossible to achieve visible learning outcomes, if time and other resources are limited and that the consequence is teaching to the test". Moreover, others reported the greatest impact of resource availability on the teaching learning process, particularly on science teaching (Lasonen, Kempainen, & Raheem, 2005: 40-41; Bines & Woods, 2007: 10; MoE, 2010: 19-27).

In addition, a wide variety of instructional materials, equipment and technological tools are required to implement student centered learning strategies (Firestone et al., 2004: 66). Specifically, research evidences revealed the effects of resources availability, layout and quality on formative assessment practice in science education (Bell & Cowie, 2002: 22; Wiliam, 2006: 287) and impact of shortage of time to share their experiences, to discuss on the barriers and enablers of their formative assessment practice, and to observe other model classrooms (Harlen, 2003: 38).

On the other hand, Iron (2008: 27-32), recommended that developing and using peer and self - assessment as a best strategy to reduce teachers workload and time pressure. For him, reducing the amount of summative assessment and replacing it with formative assessment will reduce teachers' workload. Moreover, peer-assessment (CERI, 2005: 5) and small group feedback discussions after students received written comments (Nicol & Macfarlane-Dick, 2006: 208-210) are important when teachers faced large classes.

As mentioned by Fautley & Savage (2008:84-87), the physical and the affective classroom atmosphere has a great impact on the teaching learning process. As to them, the physical classroom environment includes the accessibility, quality and the arrangement of seating and desks in the room; the availability of enough open spaces in the seating arrangement; the availability of bag storage; the visual environment through displays of work, posters and other resources; the quality of lighting in the room; the air quality and temperature; the provision and location of other resources, including books, items of other equipment, ICT, etc.; and outside environments, whereas the affective environment indicates the emotional (psychological) climate of the classroom which includes students' mental disposition, their level of interest and their motivation to learn. All of these have a great impact on students' science learning. Therefore, creating comfortable classroom environment is the first step to implement formative assessment effectively in science education.

For this reason, the Ministry of Education in Ethiopia promoted strategies to establish school clusters and resource centers at a certain area; as a result teachers in the clustered schools are strongly supported and use reference materials, models and equipment jointly in order to enhance



their competences and to achieve higher order learning outcomes in science education (MoE, 2002: 51-52). To sum up, the availability of resources is important to implement formative assessment or assessment for learning in secondary school science subjects to achieve the desired outcomes and to develop self-regulated learners who are crucial in the knowledge economy.

## **2.4 Conclusion**

This chapter has attempted to digest the theoretical framework of assessment for learning and key areas in assessment research as base for the enhancement of practice to improve students' learning in science subjects. On the basis of research evidences, formative assessment or assessment for learning is being advocated as a powerful means to increase students learning. Moreover, from the evidence formative assessment is crucial to increase students' motivation to learn, positive self-esteem, self-efficacy and self-regulation, which are all important qualities to help students to become active participants in the knowledge based society in which science subjects play an important role. However, formative assessment which supports students' learning is still not common practice in science classroom, particularly, in Ethiopia. Mostly, assessment practices in secondary school science classroom do not encourage students' motivation to learn, to develop positive self-esteem, self-efficacy and to be self-regulated learners. Moreover, comprehensive studies were not done in Ethiopia on formative assessment as science inquiry method to improve students' learning and to identify inhibiting factors of teachers' assessment for learning practice.

Research on formative assessment from cognitive constructivist and social constructivist perspective directs our attention towards the integration of assessment with instruction and active role of students within the assessment process to identify gaps. Therefore, cooperative works between policy makers, researchers, school administrators, teachers and students should be done to incorporate the new perspectives in the teaching learning process to improve students' learning in science subjects.

## **2.5 Summary**

Chapter two clarified the theoretical framework of assessment for learning, the component of assessment for learning, empirical research evidences on the effect of formative assessment on science learning, and characteristics of formative assessment. Moreover, the chapter comprised the inhibiting factors of the practice of formative assessment in secondary school science subjects. The next chapter will explain the research methodology. This includes: the research design, the research method, sampling, sampling technique, data gathering tools, pilot testing, methodological norms, and data analysis methods.

## **CHAPTER THREE**

### **3. RESEARCH METHODOLOGY**

#### **3.1 Introduction**

Chapter two discussed the theoretical framework of assessment for learning and related literature review on: the elements of assessment for learning, research evidences on how formative assessment improves students' learning, particularly in science education, qualities of formative assessment and possible factors that hinder teachers' practice of assessment for learning. This chapter describes the research methodology employed to investigate second cycle secondary school science teachers' practice of assessment for learning and inhibiting factors of their practice. Beside this, it explains the research design and the data collection techniques and procedures such as population, sampling technique, sample size, methods of data collection /instruments/, pilot testing, methodological norms, ethical issues, and method of data analysis.

#### **3.2 Research design**

According to De Vos, Strydom, Fouche, & Delport (2005: 132), research design is a plan or blueprint of how you intend conducting the research. These writers further note that research design provides the overall structure of the data collection procedures that the researcher follows to select samples, collect data and analyses data. The research design is a strategy aimed at gathering evidences to answer the research questions. Hence, in order to address the research questions mentioned in section 1.2, mixed research design was applied. Mixed-method approach has got a wonderful popularity in the social, behavioral and related sciences in recent years (De Vos et al., 2011: 434). Moreover, it is a currently accepted research design in education (Ridenour and Newman, 2008: 16). The main reason of applying mixed research design in this study was to triangulate the results obtained from quantitative and qualitative data, to validate

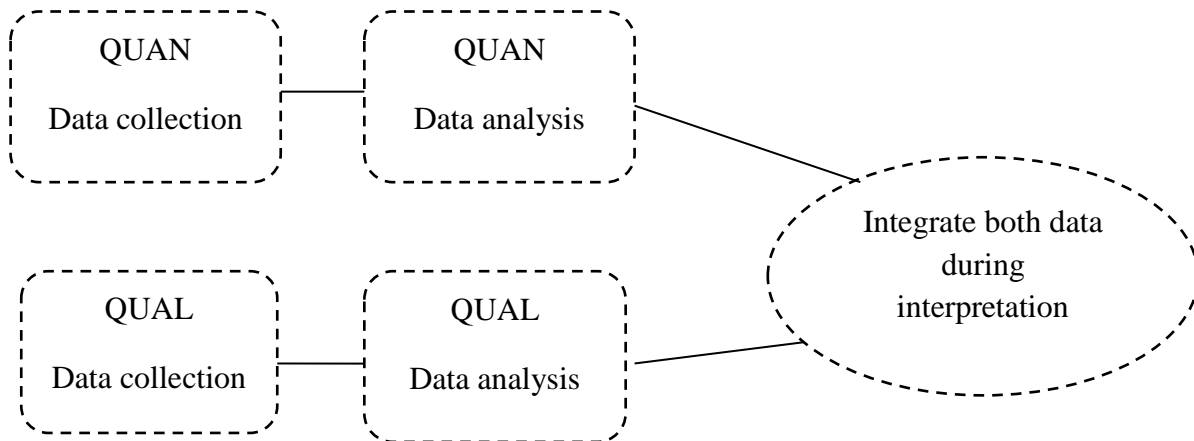
and to explore the different aspects of the phenomenon to get more detailed information about the issue.

According to Creswell (2009:13-14), a mixed methods approach is the one in which the researcher tends to base knowledge claims on pragmatic grounds. As pragmatic knowledge claims, solving the problem is the most important issue. For pragmatic, both quantitative and qualitative approaches are essential to derive knowledge or evidences about the research problem. For this reason, it is useful to capture the best from both quantitative and qualitative approaches to investigate the problem. As stated by De Vos et al. (2011: 436), “mixed methods research eliminates different kinds of bias, explains the true nature of a phenomenon under investigation and improves various forms of validity or quality criteria”.

The use of mixed method design helps the researcher to understand the problem clearly and allows us to answer research questions that are difficult to answer using quantitative or qualitative methods alone (Lopez-Fernandez & Molina-Azorin, 2011: 1460). Mixed method design employs problem-centred strategies of inquiry to clearly understand the research problem. Similarly, researchers who engaged in mixed method research argued that the use of both quantitative and qualitative methods provides a more complete evidence of a research problems than does the use of either quantitative or qualitative approaches alone (Fraenkel & Wallen, 2009: 557; De Vos et al., 2011: 436; Creswell, 2012: 535-540).

As a result, among different mixed methods design equal weight concurrent mixed design was selected for this study, because, it is one of the most widely used design from other mixed methods designs in social and human science researches (De Vos et al., 2011: 442; Creswell, 2009: 248). In this design, the researcher collects both quantitative and qualitative data simultaneously during the study and then integrates the evidence gathered from qualitative and quantitative methods in the interpretation of the overall results (Lopez-Fernandez & Molina-Azorin, 2011: 1461). The investigator believes that concurrent mixed methods design is better to understand the issue under investigation by converging both qualitative and quantitative data. It helps the researcher to offset the weakness of one method by other methods for more complete understanding of the problem (Creswell, 2012: 540). Moreover, it is useful to triangulate the data

for the validation of the findings of the issue under investigation. Therefore, the researcher collects both forms of data simultaneously, analyze each data separately, and finally integrate both forms of data in the interpretation part to see whether the findings from the two methods resemble or contradict each other. Graphically, the design was presented as follows.



*Figure 1: Equal weight concurrent mixed design*

### 3.2.1 Quantitative approach

Quantitative approach, particularly descriptive survey design is important to gather information about individuals' perception, opinion, attitude, belief and knowledge about the current issues in education (Fraenkel & Wallen, 2009: 390; Lodico, Spaulding, & Voegtle, 2010:26; Creswell, 2012: 376). It is a widely used design to determine the existing individual practice in education (Creswell, 2012: 376-377). Survey design is typically important to gather data at a particular time from a population to describe the nature of the existing condition and to identify key variables or factors (Cohen, Manion, & Morrison, 2007:170). Moreover, it is useful to include larger sample size and to generalize findings for the population (De Vos et al., 2011: 156-157; Lodico, Spaulding, & Voegtle, 2006: 156-160; Creswell, 2009: 174-176).

Thus, in this study, survey design will be used to examine science teachers' perception of assessment in relation to students' science learning, practice of integrating assessment to

improve learning, their professional learning, school support system and to identify possible factors that hinder their assessment for learning practice. Therefore, questionnaire or survey is the most appropriate method; and fits for the investigator's purpose.

### **3.2.2 Qualitative approach**

Qualitative research is more important to explore the issue or the problem in its natural settings. In qualitative approach, the researcher collects evidences at the site where the participants experience the phenomenon of the problem under investigation (Willig, 2008: 15-17). Thus, qualitative approach gives a complex and detail understanding of the issue or the setting by directly involving individuals to share their experiences in their work places.

The specific type of design employed in the qualitative approach of this study was phenomenological strategies, because the researcher thought that phenomenological design is useful to describe the meaning of the lived experience of individuals about the phenomenon (De Vos et al., 2005:270). As Creswell (2007: 58), it is essential to describe the question of ““what” they experienced and “how” they experienced it.” Moreover, phenomenological approach is more applicable particularly for psychological, sociological and educational researches to deeply understand the existing practice (Creswell, 2007: 60).

Thus, in this research the investigator employed phenomenological strategy to describe science teachers' lived classroom assessment experiences or practices of integrating assessment with instruction and to identify the lived barriers that hinder their assessment for learning practice. Particularly, in this strategy high attention was given to assess the actual practice of assessment for learning strategies (such as the sharing of learning targets and the criteria for success, provision of descriptive feedback, questioning strategies, involvement of students in their own learning or in peer and self-assessment, teacher-student and student-student interaction), the kind of support schools provided for teachers, the assessment training teachers received and its relevance to implement assessment for learning, the availability and organization of laboratory equipment, and the physical environment of the classroom (i.e. number of students, availability

of chairs, tables, light and instructional materials). Therefore, in this case in-depth semi-structured interview and observation are used to get the detail information of the problem.

### **3.3 Sampling**

The population for this study was second cycle secondary school science teachers in East Gojjam Zone, Amhara Regional State, Ethiopia. According to Creswell (2012: 142), population who shared the same characteristics refers to a group of people in the universe while sample refers to a sub-section of individuals who are selected from the target population or sampling frame to study about it. Sampling frame or target population is defined as “the actual list of sampling units from which the sample is selected” (Creswell, 2012: 381).

Totally, there are eighteen second cycle secondary schools in East Gojjam Zone, Amhara Regional State, Ethiopia. Thus, to select representative schools in the Zone for quantitative approach, random sampling was employed. Accordingly, the schools were first clustered into two based on their year of establishment. Those schools ‘older’ than ten years were clustered in one group (Cluster one) and those schools ‘younger’ than ten years were clustered in the second group (Cluster two). Then, five schools from each cluster were selected using simple random sampling, particularly lottery method. Next, all 186 science teachers in the selected schools were included in the study, and 156 of them agreed to participate in the study. In cluster one there were 127 science teachers and in cluster two there were 59 science teachers as shown in Table 3.1.

**Table 3.1 Number of science teachers in the selected second cycle secondary schools**

Schools	Population	Respondents	Percentage
<b>Cluster one</b>			
School 1	32	25	87.13%
School 2	25	22	80%
School 3	26	25	96.16%
School 4	14	11	71.43%
School 5	30	27	90%
<b>Cluster two</b>			
School 6	8	3	37.5%
School 7	15	12	80%
School 8	16	16	100%
School 9	8	5	62.5%
School 10	12	10	83.33%
<b>Total Sample</b>	<b>186</b>	<b>156</b>	<b>83.87%</b>

*Source: Respective schools*

For qualitative section, purposive sampling was employed and eight participants were selected purposively from the nearby four schools to the researcher based on their teaching experience, educational background and subject they teach to get detailed information. Purposive sampling is more applicable in qualitative research to choose individuals who have deep knowledge or experience about the issue under investigation (Cohen, Manion, & Morrison, 2007:115). Other researchers also explained the importance of using purposive sampling in qualitative research to select participants or sites that will best help the researcher to understand the problem and to answer the research question (Creswell, 2009: 21; Creswell: 2012: 206). Thus, eight science teachers participated in the interview and observation sessions to get detailed and live information about their actual practice of integrating assessment within science instruction and explore possible factors of not applying such practices in the selected schools.



### **3.4 Data collection techniques**

To answer the research questions, both quantitative and qualitative data were collected. The triangulation of methodological approaches and multiple data sources allow the researcher to reduce bias in data collection and analysis and at the same time to get deeper understanding about the issues under study (De Vos et al., 2011: 436; Creemers et al., 2010: 116-120; Creswell, 2009: 240- 248). Thus, for quantitative aspect closed-ended questionnaires were designed whereas for qualitative part in-depth semi-structured interview and unstructured observation were used to address the research question mentioned in section 1.2.

In this study the researcher employed a broad survey (questionnaire) in order to examine teachers' practice of assessment for learning, perception on assessment for learning, school support provided to implement assessment for learning in science subjects, professional training teachers gained and its relevance to implement formative assessment and to identify possible factors that hinder the practice of assessment for learning and to generalize results to a population. On the other hand, detailed semi-structured interviews and unstructured observation were employed to collect detailed views and live experience of participants about the same issue to clearly realize the research problem under investigation.

#### **3.4.1 Questionnaire**

According to Creswell (2012: 382), questionnaire “is a form used in survey design that participants in a study complete and return to the researcher.” It is the most commonly used instrument in survey research design to collect data about the opinions, perceptions and knowledge of people on a particular issue (De Vos et al., 2005: 166). Moreover, these authors argue that questionnaire is useful to gather the necessary and broad information from a larger population. In this study self-administered questionnaire was employed to gather the relevant data from second cycle secondary school science teachers. The draft questionnaire consisted of 164 closed-ended items and 3 restricted open ended questions. The questionnaire has six sections. These are: Background information (7 items), teachers' practice of assessment for learning (72 items), teachers' perception of assessment for learning (47 items), professional

learning they received and its relevance (12 items), school supports provided for the practice of assessment for learning(13 items) and possible factors of their assessment for learning practice (16 items). All of the items of the questionnaire were prepared based on the theoretical framework and related literature review.

A 5-point Likert scale, with the following options 1= Never, 2= Rarely, 3= Sometimes, 4=Regularly, 5= Often, was used to measure the practice sub-scale. The perception and school support sub-scales were also measured by a 5-point Likert scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree). A dichotomous “Yes or No” option was also used to assess the professional learning teachers received and its relevance to implement assessment for learning. Moreover, a 5-point Likert scale, i.e. 1=Very Unlikely, 2=Slightly Unlikely, 3= Neither, 4=Slightly Likely, 5=Very Likely, was used to assess the possible factors of not applying assessment for learning in science instruction.

### **3.4.2 Observation**

The observation was conducted through unstructured participant observation, because as to De Vos et al., (2011: 329-331), real feelings, expressions, experiences, words, and impressions in a specific situation can be only understood through participant observation. Moreover, these authors agreed that participant observation is the best data collection method for phenomenological approach to gain an in-depth understanding about live or natural experience of respondents. Thus, it provides opportunity for the researcher to get real evidences from real-life situation (Jones & Somekh, 2005:139-140). Generally, according to Cohen, Manion, & Morrison (2007: 305), observation is important for the researcher to gather data related to:

- The physical environment and its organization that is the availability of light, tables, chair, blackboard, neatness of rooms, and their arrangements.
- The human setting (the organization of individuals) that is the setting arrangement of students in the classroom.
- The interactional setting that is student to teacher, student to students’ dialogues or interaction.

- The programme setting (the availability of resources and their organization).

Therefore, in this study, lesson observation was employed to gather detailed and live data related to the practice of assessment for learning and possible factors that hinder such practices. While observed, attention was given primarily on the interaction between teacher with student and students to students, formative assessment strategies mostly used and their purpose, the physical environment of the classroom, seating arrangement of students, number of students in one class, the availability of instructional materials and laboratory rooms and equipment and their organizations. The observation was conducted two times for purposively selected teachers. However, for three teachers the observation was made only at one time because of their inconvenience. Therefore, there were 13 classroom observations for eight teachers. Moreover, two laboratory rooms were observed, one from cluster one and the other from cluster two. The observation of each class period was recorded in the form of field note.

### **3.4.3 Semi-structured interview**

Semi-structured interview allows the researcher to explore various issues in-depth with respondents (De Vos et al., 2011: 342-359; Ridenour & Newman, 2008:20-21; Creswell, 2009: 205-207; Creswell, 2007: 35-39). Thus, semi-structured interview was used in order to gain a detailed picture of a participant's perception and knowledge about a particular issue. Interview provides direct information about participants' knowledge and thinking on that particular phenomenon (Cohen, Manion & Morrison, 2007: 269). In this study, face-to-face semi-structured interviews were conducted with eight teachers in the school in order to address the research questions. In both cases the same protocol was followed.

The interview questions include teachers background information as an introduction, purpose of classroom assessments, teachers' perception on nature and purpose of formative assessment, formative assessment methods teachers mostly used, teachers' experience in sharing learning objectives and assessment criteria and implementing self-and peer assessment, type of feedback teachers provide, questioning techniques teachers used, learning contents mostly assessed, application of the collected assessment evidences, the type of assessment training they took and

its relevance to implement formative assessment, supports provided from school principal, supervisor, and department head to implement formative assessment, and possible factors that hinder teachers' practice of assessment for learning.

Moreover, interviewees were asked to suggest possible solutions to improve the implementation of formative assessment in science subjects to improve students' learning standards. All of the questions were prepared based on the literature and they are directly related to the research questions (See Appendix E). Generally, the aims of the semi-structured interview and the unstructured observation were to get more detailed and live information about the problem under investigation.

## **3.5 Methodological norms**

### **3.5.1 Validity**

According to Babbie (2007:146), as cited in (De Vos et al., 2011: 172), "validity refers to the extent to which an empirical measure adequately reflects the real meanings of the concept under consideration". Hence, validity refers to the trustworthy, meaningfulness and appropriateness of the instrument to measure what it intends to measure. There are four types of validity. These are: face validity, content validity, concurrent validity and construct validity. These forms of validity refer only the instrument validity in the quantitative and qualitative data gathering methods because, according to Ridenour and Newman (2008: 40), if the instrument is valid and reliable, then the research design is internally and externally valid. For them, measurement validity and design validity are dependent.

For the quantitative section, to establish the face and content validity of the questionnaire appropriate sampling of items for both sections were included and reviewed by peers and selected experts in the area. Content validity and face validity of the instrument can be established on the bases of researchers' and experts' judgment (De Vos et al., 2011:173) and by including items that comprehensively cover the domain (Cohen, Manion, & Morrison, 2007: 137). All of the items in the questionnaire were prepared based on different theories included in

the literature to increase their construct validity. Moreover, Exploratory Factor Analysis was employed to assess the relationship among items and factors in the practice, perception and possible factors sub-scales. The concurrent validity of the questionnaire items can be established through triangulation of different data sources.

For qualitative section, semi-structured interviews and observation were conducted in real life situations, and the questions in the interview were reviewed by peers to establish the validity of the question. Willig (2008: 16) reported that obtaining feedback from participants, conducting interviews and observations in workplaces or in their natural setting are considered as useful strategy to validate qualitative methods. Moreover, prolonged engagement on-site, persistent observation, triangulation and peer debriefing are important strategies to prove validity in qualitative methods (Ridenour & Newman, 2008: 57-58). Likewise, Creswell (2007: 208-210) pointed out important opinions to increase the validity of qualitative methods. These are: peer review, prolonged engagement and persistent observation, use of multiple and different sources, methods and theories (i.e. triangulation), and involvement of respondents to view the credibility of findings and interpretations. That is why; peers review, triangulation and two phase observation were employed in this study. Moreover, two respondents were involved to review the validity of the results and interpretation of results in this study.

### **3.5.2 Reliability**

For quantitative section, Split-half Spearman Brown and Cronbach alpha internal consistency estimation methods were employed to establish the reliability of each item in the questionnaire. Split-half Spearman-Brown internal consistency estimation method is applicable for dichotomous items and Cronbach alpha is useful for multi-scale items or for those items which have five options (Cohen, Manion, & Morrison, 2007: 146-148). In this study, split-half method was used for items related to “Teachers’ Professional Learning” sub-scale and Cronbach alpha was for other sub-scales in the questionnaire.

Furthermore, reliability in qualitative research can be defined as “a fit between what researchers record as data and what actually occurs in the natural setting that is being researched, i.e. a

degree of accuracy and comprehensiveness of coverage” (Cohen, Manion, & Morrison, 2007: 120). According to Lincoln & Guba (1985), reliability in qualitative research are synonymous to “credibility”, “neutrality”, “consistency”, “applicability”, trustworthiness”, “transferability”, “dependability” and “conformability” (cited in Cohen, Manion, & Morrison, 2007: 140). For these writers, reliability in qualitative research can be checked through conducting the same observation at different time at different places and using semi-structured questions for interview. Thus, in this study, the same observation was conducted two times for five individual and one time for three individual at four schools and semi-structured interview were conducted with the same individual to get reliable data. Creswell (2007: 209) mentions that for qualitative approach reliability can be enhanced if the researcher obtains detailed written description from the interview and observation. For this reason, eight teachers were interviewed using audio recording and detailed written descriptions were taken during lesson observations.

### **3.6 Pilot study**

According to Bless & Higson-Smith (2000: 155), pilot study is defined as “A small study conducted prior to a larger piece of research to determine whether the methodology, sampling, instrument and analysis are adequate and appropriate” (cited in De Vos et al., 2005: 206). Hence, for quantitative section, pilot study was conducted to determine the strength and weakness of the questionnaire in terms of questioning format, wording, and order of items and to test the reliability of the questions. The pilot study is also helpful to see the face and content validity of the item (De Vos et al., 2011: 195). Moreover, according to Cohen, Manion, & Morrison (2007: 260) pilot study is important to check the clarity of the items, instruction and layout; to gain feedback on the validity and reliability of the instrument; to eliminate ambiguity in wording and to check the time needed.

The questionnaire was reviewed by Prof. Wynne Harlen (an international assessment professional, particularly, in science education) and 5 lecturers in Debre Markos University (one from English Department, one from Education and three from Psychology Department) before piloting. For the pilot study, a total of 15 teachers were selected from the target population based on their experience and completed the questionnaire. A lot of comments were forwarded related

to wording of sentences, vague sentences, item formats, elements that they believed to be incorporated or avoided, instruction and the face validity of the item. Besides, they strongly commented the need of translating the questionnaire to the subjects' mother tongue to get valid data. All of the comments were incorporated accordingly.

Furthermore, the reliability for each section of the piloted questionnaire was calculated using Split-half Spearman-Brown formula for the dichotomous items and Cronbach alpha for five-point Likert scale items. The results were as follows: "Teachers' Perception on Assessment for Learning" sub-scale ( $\alpha = 0.936$ ), "Teachers' Practice of Assessment for Learning" sub-scale ( $\alpha = 0.932$ ), "Teachers' Professional Learning" sub-scale (Spearman-Brown coefficient or  $r = 0.512$ ), "School Supports" sub-scale ( $\alpha = 0.886$ ), and "Possible Factors" sub-scale ( $\alpha = 0.506$ ). Based on the results of the pilot test, twenty items were removed that is nine items from perception sub-scale, four items from practice sub-scale, three items from Teachers Professional Learning sub-scale, two items from Schools' Support and two items from possible factors sub-scales. Moreover, the questioning format in the practice sub-scale was changed. Thus, changes were made accordingly. Finally, the questionnaire was translated into Amharic and administered for all science teachers in the selected schools.

The final questionnaire used in this study had 7 (4 closed-ended and 3 restricted open-ended) background items, 68 "Teachers' Practice of Assessment for Learning" sub-scale items, 38 items related to "Teachers' Perception on Assessment for Learning" sub-scale, 11 items related with "School Supports" sub-scale, 9 items connected to "Teachers' Professional Learning" sub-scale and 14 items related to the "Possible Factor" sub-scale (see Appendix D). The reliability coefficients of the different sub-scales in the final questionnaire were as follows: "Teachers' Perception on Assessment for Learning" ( $\alpha = 0.857$ ), "Teachers' Practice of Assessment for Learning" ( $\alpha = 0.908$ ), "Teachers' Professional Learning" (Spearman-Brown coefficient or  $r = 0.674$ ), "School Supports" ( $\alpha = 0.923$ ), and "Possible Factors" ( $\alpha = 0.843$ ). A reliability coefficient closer to one indicates that the scales have good internal consistency and both of the items in the scale are measuring the same thing (Daniel, 2004: 73-74). Moreover, Cohen, Manion, & Morrison (2007: 506) interpreted that a reliability coefficient of  $< 0.60$  is low reliable,  $0.60 - 0.69$  is marginally/minimally reliable,  $0.70 - 0.79$  is reliable,  $0.80 - 0.90$  is

highly reliable, and  $> 0.90$  is very highly reliable. Thus, the scales in this questionnaire have reasonably good reliability coefficient to carry out the study.

Furthermore, factor analyses were employed to reduce large number of items in the final perception, practice and possible factor sub-scales of the questionnaire to make it manageable prior to analysis. Particularly, Exploratory Factor Analysis (EFA) was used to refine and reduce a large number of individual sub-scale items to form a smaller number of factors that can be easily analyzed and interpreted (Cohen, Manion, & Morrison, 2007: 560-563). Moreover, it helps to increase the construct validity of the questionnaire. Thus, the participants' responses for each sub-scale were factor-analyzed with principal component methods of extraction and Varimax orthogonal rotation. Eigenvalues equal to one or more than one were extracted for each sub-scale to represent the amount of the total variance explained by each factor. Once factors have been extracted, Varimax orthogonal rotation was implemented to reduce the number of factors before interpreting the results. According to Cohen, Manion, & Morrison (2007:560-567), Varimax orthogonal rotation is the most commonly used technique to minimize the number of factors for easy interpretation of results. The factor loadings of the three sections of the questionnaire were presented below.

### **Teachers' practice of assessment for learning sub-scale**

To assess the factorability of the data and the adequacy of the sample, Bartlett's test and Kaiser Meyer-Olkin were used. Accordingly, the results from both tests look good. Kaiser Meyer-Olkin coefficient was 0.805 for "Teachers' Practice of Assessment for Learning" sub-scale which suggests that the data are suitable for principal component analysis. Likewise, a significant Bartlett's test at  $P= 0.000$  indicates there is a sufficient correlation between variables. Hence, it is possible to precede the factor analysis. Thus, based on the results of principal component analysis and Varimax orthogonal rotation method, the 68 items of "Teachers' Practice of Assessment for Learning" sub-scale were reduced to six factors which accounts for 49.82% of the total variance. Using a cutoff point 0.30, the six factors for "Teachers' Practice of Assessment for Learning" sub-scale were: Application of assessment evidences (which accounted for 10.7%), Collection of learning evidence (which accounted for 10.4%), Support



provided to engage students actively (which accounted for 9.4%), Interpretation and communication of evidence (which accounted for 7.6%), Learning contents assessed (which accounted for 6%), and Planning of formative assessment (which accounted for 5.8%) (see Table 3.2 below).

During factor analysis, five items (i.e. using of “tests”, “project works”, “laboratory works”, “students’ questioning”, and “open-ended questioning” were not included in the factor structure because they did not load greater than or equal to 0.03 on any of the six factors (Cohen, Manion & Morrison, 2007:426). The reliability for each factor was also calculated as follows: Application of assessment evidences ( $\alpha = 0.720$ ), Collection of learning evidence ( $\alpha = 0.878$ ), Support provided to engage students’ actively ( $\alpha = 0.820$ ), Interpretation and communication of evidences ( $\alpha = 0.905$ ), Learning contents assessed ( $\alpha = 0.808$ ), and Planning of formative assessment ( $\alpha = 0.782$ ). Table 3.2 below indicates the factors, the reliability, and the variance explained by each factor loading in “Teachers Practice of Assessment for Learning Practice” sub-scale (see the detail on Table 4.2).

**Table 3.2 Factor structure of teachers’ practice of assessment for learning sub-scale**

<b>Factors</b>	<b>Reliability</b>	<b>% variance explained</b>
Application of assessment evidences	0.720	10.70%
Collection of learning evidence	0.878	10.40%
Support provided to engage students’ actively	0.820	9.40%
Interpretation and communication of evidences	0.905	7.60%
Learning contents assessed	0.808	6%
Planning of formative assessment	0.782	5.80%
<b>% of Total variance Explained</b>		<b>49.90%</b>

**Teachers’ perception of assessment for learning sub-scale**

For this sub-scale, the Kaiser Meyer-Olkin coefficient was 0.850 which was good to precede principal component analysis. The Bartlett’s Test of Sphericity also is significant at  $P < 0.05$ . As

a result, the extraction method and the Varimax rotation of variables yielded three factors for “Teachers’ Perception of Assessment for Learning” sub-scale namely: Power of formative assessment to improve learning (accounts for, 21.33%), Involvement of students (accounts for, 12.16%), and Provision of feedback (accounts for, 11.22%). Totally the three factors accounts for 44.72% of the total variance in the sub-scale.

A correlation coefficient of 0.30 and above was used as cutting point for factor analysis. Accordingly, four items, “formative assessment is a series of tests administered to evaluate student’s learning”, “using one crucial assessment method allows students to improve their performance”, “formative assessment is an engaging and enjoyable process for students” and “creating competitive classroom environment among students improves learning” were not included in the factor structure because, they did not load greater than or equal to 0.03 on any of the three factors. Cronbach’s Alphas for each factor were as follows: Power of formative assessment to improve learning ( $\alpha = 0.877$ ), Involvement of students ( $\alpha = 0.708$ ), and Provision of feedback ( $\alpha = 0.729$ ). Table 3.3 presents the reliability coefficient, factor loading, and percent of the variance explained by each factor (see the factor loading of each item on Table 4.4).

**Table 3.3 Factor structure of teachers’ perception of assessment for learning sub-scale**

<b>No.</b>	<b>Factors</b>	<b>Reliability</b>	<b>% variance explained</b>
1	Power of formative assessment to improve learning	0.877	21.30%
2	Involvement of students	0.708	12.20%
3	Provision of feedback	0.729	11.20%
<b>% of Total variance Explained</b>			<b>44.70%</b>

### **Possible factor sub-scale**

To check the factorability of the data Kaiser Meyer-Olkin measure of sampling adequacy and Bartlett’s Test of Sphericity were also calculated with the help of SPSS for “Possible Factor” sub-scale. In this case the Kaiser Meyer-Olkin value is 0.789 and the Bartlett’s Test of Sphericity is significant at  $P < 0.05$ , therefore, factor analysis is possible. As a result, the principal

component analysis and the varimax rotation of variables yielded three factors namely: Support, Awareness and Resources related factors. Support related factors explains 22% of the variance, Awareness related factors explains 19.4% of the variance and Resources related factors accounts 14.6% of the variance. The total variance explained by the three factors was 56%. The reliability coefficient for Support, Awareness and Resources related factors were:  $\alpha = 0.805, 0.798$  and  $0.664$  respectively (see Table 3.4). The factor loadings for each item in each factor were presented in Table 4.3.

**Table 3.4 Factor structure of the possible factor sub-scale**

<b>No.</b>	<b>Factors</b>	<b>Reliability</b>	<b>% variance explained</b>
1	Support related factor	0.805	22.00%
2	Awareness related factor	0.798	19.40%
3	Resources related factor	0.664	14.60%
<b>% of Total variance explained</b>			<b>56.00%</b>

### **3.7 Method of data analysis**

The quantitative and qualitative data were analyzed separately given different nature of the data; however, after analysis the data were integrated to answer the themes or the basic research questions in this study. Next, methods of data analysis for quantitative and qualitative data are presented respectively.

#### **3.7.1 Method of analysis for quantitative data**

First descriptive statistics such as frequency and percentage were employed to describe the background characteristics of participants in the questionnaire. Then, the method of data analysis for each research question follows. The following illustrates how the quantitative data were analyzed for each research questions.

**Research question one:** *To what extent do teachers integrate assessment in their instruction to improve students' science learning?*

The data were collected using “Teachers Practice of Assessment for Learning” sub-scale. As mentioned in section 3.7, this sub-scale has 68 items including the 5 items which are rejected after factor analysis (see Appendix D). These items were also reduced into six factors (See Table 3.2). For this sub-scale the minimum, maximum and expected mean scores obtainable respectively are 63, 315 and 189. As a result, teachers are considered as they hardly ever practicing assessment for learning in their lesson if their score is less than the expected mean while a teacher with a score that is greater than or equal to the expected mean is considered as she/he regularly implements assessment for learning. Similarly, teachers' level of practice in the six factors was determined based on the expected and observed mean score of teachers in the factors. Likewise, mean, standard deviation and percentage of respondents were calculated for each item in each factor to assess their particular focus in the factors. Moreover, one sample t-tests were calculated to assess significant differences among factors.

For the sake of simple analysis and interpretation of percentage of respondents in each item, the five point Likert scale were grouped into three categories. These are: Strongly Agree and Agree were collapsed into Agree, Strongly Disagree and Disagree were collapsed into Disagree and the Neutral response was taken as it is. Finally, the percentage of respondents who scored below or above the expected mean was mentioned for the total sub-scale and each factor for universal conversation about the practice of assessment for learning in the selected schools.

**Research question two:** *What are the major factors that affect teachers' practice of assessment for learning in science subjects?*

Here, the “Possible Factor” sub-scale of the questionnaire was used to get the relevant information. The sub-scale has 14 items and three factor loadings (see Table 3.4). So, to assess the major possible factor that hinders teachers' practice of integrating assessment with their daily instruction, mean score and standard deviation were calculated for each item and factor. Moreover, percentage of respondents for each item was calculated by grouping the options of

very unlikely and unlikely into one category (unlikely) and very likely and likely into another category (likely). Meanwhile, the neither options was taken as it is. Besides, t-test for one sample case was conducted to assess significant differences among the three factors namely, resource, awareness and support related factors and to identify the most possible factor.

**Research question three:** *How do second cycle secondary schools science teachers perceive assessment for learning?*

To gather relevant information related to teachers' perception of assessment for learning 38 items including the excluded items after factor analysis were used. The sub-scale has three factor loadings (see Table 3.3). The minimum, maximum and expected mean score in this sub-scale is 34, 170, and 102 respectively. Thus, percentage of teachers who scored less than or greater than the expected mean in this sub-scale were described to assess whether teachers have negative or positive perception on assessment for learning. That is to say a teacher with less than a score of 102 in this sub-scale is considered to have negative perception toward assessment for learning whereas a teacher with a score that is equal to 102 and above is considered to have positive perception. Moreover, their perceptions for the three factors were judged based on the expected and observed mean score of teachers in each factor.

Besides, descriptive statistics such as mean, standard deviation, and percentage of respondents were calculated for each item and factor to determine teachers' level of perception on the different aspects of formative assessment. For simple data processing, the options of strongly disagree and disagree were collapsed into one category (disagree) and agree and strongly agree collapsed into agree. Meanwhile, neutral responses were taken as they are. One sample t-tests were also calculated to identify significant differences in teachers' perception of assessment for learning among the three factors.

**Research question four:** *To what extent do professional (pre-service and in-service assessment) trainings help science teachers to integrate assessment with their daily instruction?*

There are 9 items in the “Teachers’ Professional Learning” sub-scale of the questionnaire to gather pertinent information concerning teachers’ assessment training and its relevance to implement assessment for learning in science subjects. Thus, descriptive statistics such as frequency and percentage were conducted for each item to assess those teachers who received pre-service and in-service assessment trainings and to analyze the relevance of taking the training to implement assessment for learning in science subjects. Such analysis also helps to determine teachers’ awareness about their knowledge on formative assessment and their future opinion on taking trainings related to formative assessment.

**Research question five:** *What type of support schools provide for the effective practices of assessment for learning in science subjects?*

The quantitative data related to schools support provided for science teachers to implement assessment for learning were gathered using an 11 items sub-scale. As a result, frequency, percentage and mean were used to determine schools level of support provided for teachers to integrate formative assessment in their lesson and to upgrade their understanding about assessment for learning. It also helps to assess teachers’ level of perception on the importance of support provided by school principals and supervisors to integrate formative assessment in their instruction. In general, teachers who scored below 33 in this section are considered as no school support provided for them while scores equal to 33 and above were considered as they got intensive support from their school to implement assessment for learning.

**Research question six:** *Are there significant relationships between teachers’ perception and school supports with teachers’ overall practice of assessment for learning?*

Multiple linear regressions was used to assess the relationship between teachers’ perceptions on assessment for learning and school supports provided for teachers with teachers’ overall practice of assessment for learning. Here, the independent variables (predictors) were teachers’ perception and school supports provided for teachers and the dependent variable (criterion) was teachers’ combined practice of assessment for learning. Before conducting the analysis the preliminary assumption of linearity and multicollinearity were checked.

According to Daniel (2004: 176-182), in multiple regression, the independent variables should be highly correlated with the dependent variable and should have low correlation among themselves. The writer also added that the relationship between the two independent and dependent variable should be above 0.30. In this study, the correlation between teachers' perception and school support provided for teachers' overall practice of assessment for learning were 0.406 and 0.410 respectively, which indicates sensible relationship among the independent and dependent variable. Moreover, the relationship between the two independent variables was 0.28, which did not violate the primary assumption. The result in this study also showed tolerance value of 0.922, which indicates the amount of variance in the dependent variable not explained by the other predictor or independent variable. Therefore, the assumption of multicollinearity and linearity were fulfilled to perform the multiple regression analysis.

**Research question seven:** *How do the training and background characteristics (teaching experience, teaching subject, workload per week and class size) of teachers' affect their practice of assessment for learning?*

Here, separate one-way multivariate analysis of variance (MANOVA) was employed to investigate the effect of teachers' background characteristics and assessment trainings on the integration of formative assessment into their daily instruction. In the analysis, there are six dependent variables or factors of the composite practice sub-scale: planning of formative assessment (PFA), collection of learning evidences (CLE), interpretation and communication of evidences (ICE), support provided to engage students actively (SPESA), learning contents assessed (LCA), and application of assessment evidences (AAE). The independent variables were teachers' background characteristics (teaching experience, subject taught, workload per week, and average class size) and assessment trainings (pre-and in-service). The effect size of each independent variable was interpreted based on Cohen, Manion, & Morrison (2007: 521) guideline (i.e. 0.0 – 0.20, week effect; 0.21 – 0.50, modest effect; 0.51 – 1.00, moderate effect; and greater than 1.00, strong effect).

When the study has more than one dependent variable, multivariate analysis of variance is more preferable than repeated univariate analyses to reduce the rate of type I error and to assess the

combination effect as well as individual effects (David. 2010: 364-365). Hence, six separate MANOVA were conducted on the combined practice as a function of the six independent variables. In both cases, Wilks' Lambda multivariate test of significance was used to assess whether there are significance differences among the groups on a linear combination of the dependent variable. According to Pallant (2001: 290), Wilks' Lambda was the most commonly used multivariate test when the sample size is appropriate and preliminary assumptions are met. The significance level was tested at  $\alpha = 0.05$  for the combined dependent variable.

Following a significant result, separate univariate ANOVAs were done for each dependent variable with an adjusted alpha level to investigate significance differences across teaching experience, subject taught, class size, workload per week, pre-service and in-service assessment trainings. Pallant (2001: 229) recommended, the use of Bonferroni alpha level adjustment method to reduce the probability of making an inflated Type I error rate due to the use of multiple univariate tests. Thus, in the univariate tests the significance level was tested at  $\alpha = 0.0083$  (0.05 divided by the number of dependent variable). Finally, post hoc multiple comparisons test were used to assess significance mean differences between three or four groups, if the univariate tests showed significant results.

However, before using such parametric tests, the primary assumption of normality, linearity, homogeneity of variance-covariance, and equality of error variance were tested. Accordingly, the normality of the distribution was tested using skewness and kurtosis (z- value between -1.96 to +1.96), Shapiro-wilk test (p-value should be above 0.05) and histogram; homogeneity of variance-covariance matrices was tested using Box's test of equality of covariance matrices (p-value should be greater than 0.001); Bartlett's test of sphericity was used to assess the inter-correlation between the six dependent variables ( $p < 0.001$ ); and equality of error variance was tested by Levene's test of equality of error variances (p-value should be greater than 0.05). Based on these tests, the result indicated that there were no serious violations of the preliminary assumptions in both cases. Therefore, it is possible to use parametric tests such as MANOVA, ANOVA and Tukey post hoc comparison tests for this study.



### 3.7.2 Method of analysis for qualitative data

For qualitative part, the semi-structured interview questions and lesson observations were analyzed qualitatively using theme analysis. Thus, to extensively address the research questions, the qualitative data analysis in this section followed the steps and tasks proposed by (Creswell, 2007: 88-89; Cohen, Manion, & Morrison, 2007: 285-286; Willig, 2008: 55-59; Creswell, 2009: 21-22; De Vos et al., 2011: 411). These are: prepare, read and re-read notes that reflect the initial thoughts; identify and label significant statements; structure and group statements into a certain categories or themes; cluster themes into broader categories; and assemble the detail descriptions of the issue into the themes in the cluster. Accordingly, the following qualitative method of data analysis, particularly phenomenological data analysis steps were used in this study.

**Step1. Transcription:** During the interview and observation session the researcher developed detailed and original field notes. First, the recorded interview information was transcribed in the form of filed notes. Thus, in this step, the researcher organized comprehensive written transcript that reflects the initial verbal and non-verbal communication or expressions of participants. Then, the researcher reread the notes several times to obtain the general impression and credibility of the data and arrange the written transcript for analysis. Particularly, in the interview session, the researcher transcribed the interview data and then listened to the audio file again while comparing it to the transcript to ensure that she had not over looked any meaningful information.

**Step2. Identification of significant statements or phrases:** Here, the researcher developed a list of significant ideas from the written transcript. These statements are the real feeling or the direct expressions of respondents or quotations of ideas from the interview or observation sessions.

**Step3. Formulating themes or categories:** The identified significant statements were coded to generate categories or themes. The coded themes are basically related to the basic research questions in this study. According to Creswell (2009: 21), coding is the process of organizing the data into a certain group. Here, the researcher condensed what the participant has said into a unit of general meanings or the significant phrases grouped into larger themes or categories or

into the basic research question of this study because, according to De Vos et al. (2011: 411), if data are not categorized and named it creates confusion to analyze.

**Step4. Clustering themes:** Themes that share common meaning or relationships were clustered together. Here, it is important to ensure that the themes clustered at this stage make sense in relation to the initial thinking and observation. Thus, the researcher needs to move back and forth between the list of themes formulated and the significant statements on the transcript that generated the themes in the first place.

**Step5. Integrate the themes in the cluster into detail description of the phenomenon in the original data:** The detail description includes keywords and quotations from the interview and observation sessions. Finally, the results of the qualitative section were integrated to the quantitative part to answer the basic research questions in this study.

### **3.8 Ethical considerations**

Ethical issues are principles that guide the researcher to protect the rights of the participants who take part in the research. Lodico, Spaulding, & Voegtler, (2010:18) stated that ‘the traditional and often dominant issues that emerge when considering research ethics involve obtaining informed consent from participants, protecting them from harm, and ensuring confidentiality’. Thus, in this study, the researcher considered the following ethical issues during data collection, interpretation and dissemination. Firstly, the researcher showed an informed consent letter for respondents. The letter provides sufficient information for participants about the purpose of the study; participants’ role in the study; ways of selection; their right to ask questions, to participate or withdraw at any time; how the results will be; and the protection of their privacy to get consent from them whether they are volunteer to participate or not in this study (See Appendix C). As a result, all of the respondents decided to take part in this study. In the interview and observation part the researcher used coding for names and pseudonyms for individuals and places to protect identities. Moreover, before data collection the researcher has got ethical clearance certificate from University of South Africa and permission letter from East Gojjam Zone Education Department to conduct the research (see Appendix A and B).

### **3.9 Summary**

In this chapter, the research design and methods used were explained in detail. The validity and reliability of the data collection instruments were also established through pilot testing, peer review and factor analysis. Moreover, the respondents participated in the study more sincerely, once they understood the implications of giving their consent. The next chapter will encompass the results and discussion of results.

## **CHAPTER FOUR**

### **4. RESULTS AND DISCUSSION**

#### **4.1 Introduction**

Previous chapter explained the research design, sampling techniques, data collection instruments, pilot study, methodological norms (validity and reliability), data analysis methods, and ethical issues. Chapter four presents the results and discussion of results. As indicated in the methodological part, both quantitative and qualitative data were gathered using questionnaire, interview, and observation simultaneously. For quantitative approach survey design, whereas, for qualitative approach, phenomenological design were employed. Therefore, in this section both the qualitative and quantitative data are accessed and discussed in line with the basic research questions mentioned in section 1.2.

#### **4.2 Analysis of background data**

The background characteristics are discussed firstly to obtain gateway images about respondents in the selected schools before analyzing and discussing the results of the basic research questions in this study. Accordingly, the data about background characteristics of respondents were analyzed in the following table using frequency and percentage.

**Table 4.1 Respondents' background data**

Variable		Frequency	Percentage
Gender	Male	132	86.3
	Female	21	13.7
Educational Background	Diploma	2	1.3
	Degree	147	96.1
	Master's Degree	4	2.6
Teaching experience	Below 7 years	60	39.22
	8-15 years	49	32.03
	Above 16 years	44	28.8
Subject	Mathematics	44	28.8
	Biology	42	27.5
	Chemistry	38	24.8
	Physics	29	19
Teaching load per week	Below 10	33	21.6
	11-15	53	34.6
	More than 16	67	43.8
Number of Students	below 35	0	0
	35-50	45	29.5
	51-65	94	61.4
	Above 65	14	9.2

According to Table 4.1, 86.3% of the respondents are males, whereas 13.7% of them are females. The result shows, very low proportion of female science teachers in the selected schools which contrasts with the policy and strategies of the Education Sector Development Program IV for the year 2014/15 (FDRE, 2010: 20). Thus, attention should be given to increase the number of female science teachers in the second cycle secondary schools in the selected area, because females have also their own special contribution in the teaching learning process to improve science education.

Similarly, to achieve the intended learning objectives of science education, qualified teachers are required and it is a prerequisite for any activity in the teaching learning process. In line with this, most science teachers in the selected second cycle secondary schools (96.1%) had Bachelor's Degree, 2.6% held a Master's Degree and 1.3% of them are Diploma holders. Teachers' qualification or the knowledge, skill, and attitude they have about the subject matter they teach have its own prominent role to implement assessment for learning in science subjects.

Based on the length of teaching experience, teachers were grouped into four: teachers with 7 years of experience or below are 39.22 % of the respondents, 32.03% of the teachers have teaching experience between 8 and 15 years, and those teachers who have greater than or equal to 16 years of teaching experience are 28.8% of them. From this data, we can conclude that most of them have teaching experience below or equal to 7 years. Teachers' level of teaching experience also played a role in why and how they chose the various forms of assessment methods in their instruction.

With regard to the subject matter they taught, 24.8% of respondents were chemistry teachers, 28.8% of the respondents identified themselves as mathematics teachers, 27.5% were biology teachers, and physics teachers were 19% of the respondents. The way teachers apply different formative assessment strategies may differ in different subjects. Furthermore, most respondents (43.8%) have more than 16 periods per week (i.e. one period=40 minutes), about 34.6% of the respondents have teaching loads between 11-16 periods per week including laboratory works, and 21.6% have below 10 periods per week. If teachers were occupied by a number of periods in many classes, then it may create some challenges on teachers to share experiences with others; to see themselves; to observe models; and to engage themselves through action research to identify problems, changes and take interventions in their classroom practices.

As mentioned in Table 4.1, the average class size in the selected schools was between 51-65; (61.4% of the respondents agreed on it), 29.5% of them reported there were 35-50 students in one class and the average number of students was more than 65 in 9.2% of respondents' class. However, there was no class with average number of students below 35 in the selected school.

The observation result confirms it. In most of the observed classes, the number of students was between 45 and 60 which were difficult to actively involve each student in the lesson.

### **4.3 Analysis of the practice of assessment for learning in science education**

The instructional power of formative assessment to improve students' science learning is addressed by different well known researchers as mentioned in the literature review (see Section 2.3.2). Thus, the purpose of this part of the study is to present the results of the practice of assessment for learning or the integration of such powerful formative assessment in their daily instruction to improve students' learning and the hindering factors of such practice in second cycle secondary school science education, in East Gojjam Zone, Amhara Regional State, Ethiopia. The results for each basic research question were presented below.

**Research question one:** To what extent do teachers integrate assessment in their instruction to improve students' science learning?

#### **Teachers' practice of assessment for learning in science subjects**

Data collected through questionnaire, interview and observation related to teachers' practice of assessment for learning were analyzed and presented respectively. In the case of questionnaire, exploratory factor analysis was used to reduce the number of items in the "Practice of Assessment for Learning" sub-scale into six factors (see Table 3.2). The results of one sample t-tests also showed that all the six factor loadings of the practice sub-scale were significantly different from each other ( $p = 0.000$ ). Therefore, the results of the practice of assessment for learning in the selected school were presented in line with the six factors below.

The quantitative data in this study indicated that 93.43% of science teachers in the selected school score more than the expected mean in the provision of supports for students to engage them actively in the lesson. Moreover, 66% and 65.35% of science teachers score more than the expected mean in the application of assessment evidences to improve students' learning and

planning of formative assessment strategies as an integral part of their lesson preparation respectively, whereas, the qualitative data goes up against it. During the interview and observation sessions, there were not evidences that pointed to the teachers' use of different formative assessment strategies to engage their students actively in the lesson and use of assessment evidences to improve learning.

On the other hand, 67.32%, 56.21 and 48.37% of the teachers find themselves limited by different factors to implement different assessment for learning methods to collect learning evidences, interpret or identify learning gaps and communicate evidences in a way to improve students' learning, and assess higher order learning outcomes to equip their students with essential scientific skills and competencies respectively. It is similar with the qualitative result.

Generally, during the interview session, teachers stated that they did not regularly use different assessment for learning strategies as an integral part of their daily lesson because they fail to plan such strategies as an integral part of their daily lesson plan preparation and other related factors. As a result, they believed that they needed more training to understand the strategies and cycles in formative assessment and how to put them into practice to improve learning. Following the results of the quantitative and qualitative data for each factor loadings were presented in detail accordingly.

### **Support provided to engage students actively**

The descriptive statistics showed that from the six factors "Support provided to engage students actively" was practiced frequently than the other five factors in the practice sub-scale (a mean of 3.73). In general, teachers' responses to items in this factor seems to indicate that they provide regular support for their students to engage them actively in the lesson, but they also strongly agreed that they frequently encourage them to answer questions quickly. So, there is contradiction in teachers' response on the same issue, due to many factors as teachers reported on Table 4.3.

As indicated in Table 4.2, most science teachers (79% with a mean of 4.05 and 77.8% with a mean of 4.06) revealed that they regularly encourage every student in their class to ask questions



and to actively participate in the lesson. Moreover, considerable numbers of science teachers encourage students to share ideas (72.5% with a mean of 3.88), inspire every student's to answer questions (72.6% with a mean of 3.93), give home works (69.9% with a mean of 3.88), ask oral questions (65.4% with a mean of 3.84), persuade their students to take risks and listen to others ideas carefully (66.6% with a mean of 3.76), and create opportunity for their students to act on the feedback provided (65.4% with a mean of 3.75).

On the other hand, large number of teachers (81.1%) in the selected schools encourages their students to answer questions quickly. But, everyone did not think at the same speed or in the same way to be engaged in answering questions. Moreover, short waiting time during questioning did not allow students to think and build their thought to answer higher order questions and to get more explanation; rather such activities encourage rote learning and fast learners to answer simple facts.

The results of the classroom observation also showed that most science teachers in the selected school did not encourage their students to actively participate in the lesson. In all of the observed classes, except in one mathematics lesson (to some extent):

- Most of the students listen, read text books and take notes;
- Some fast students participate in answering questions and sometimes ask questions, if the expected answers were not forwarded; most teachers immediately answer the question and proceed to the next explanation.
- Students' were not given time to share ideas in group and to reflect it.
- Only volunteer students are encouraged to write their answer on the board (e.g. one chemistry lesson observation).

The interview result also verified it. During the interview session most teachers stated that they regularly used lecture methods and simple oral questions because of students' expectation about themselves as a student who passively receive information and the role of the teacher as a good reservoir and impart of knowledge. For example:

Mostly, I used teacher centered methods to explain important point in the lesson, because, most of my students expect me to clarify each point in the lesson to understand it better. If I did not explain it, they consider me as lazy teacher, who is careless and who does not worry about their learning. Moreover, they perceived the lesson that is not explained by the teacher as less important for their learning and ignore it. (Teacher B, Nov 19, 2014)

Besides, one mathematics teacher reported that he encourages his students to share ideas in group and to ask and answer questions during the lesson. The lesson observation also verifies it, but, he used such techniques to allow students to remember or better understand what they learn to score high marks on tests, rather than to reflect new ideas and thinking for future learning. He stated that:

After explaining the important points of the lesson, I mostly gave exercises from the text book to discuss in group (i.e. one to five grouping). I then ask oral questions randomly by calling their number and let them to ask questions which are not clear for them. If they ask questions I clearly explain the answer for them. But, still, I do not use other formative strategies to actively engage them in the lesson such as peer to peer questions, self-assessment, peer assessment, and peer feedback. (Teacher F, Nov 18, 2014)

### **Application of assessment evidences**

Regarding the “Application of assessment evidences” factor, most teachers showed inconsistency in their responses. This part of the practice sub-scale is the second loaded factor according to the results of the descriptive statistics (a mean score of 3.70, see Table 3.2). Even if, 66 % of teachers in the selected schools scored greater than the expected mean score in the factor, they also used assessment results for judgmental purpose mostly. As mentioned in Table 4.2, 80.4 % and 69.9% of science teachers agreed that they regularly used the collected assessment evidences to modify their teaching strategies and to plan what to teach next respectively. Furthermore, most teachers (70.6% with a mean of 3.92) frequently used assessment results to identify the gap of students understanding and 74.5% of them used it to advise their students on how to fill such gaps in their learning. Similarly, half of the science teachers (51.7 and 51.6%) suggest means for their students to plan their future learning and allow them to resubmit their work once they improved it respectively. In contrast to the above, 79.1% of respondents’ regularly used assessment results for the purpose of recording for final marks. Moreover, 62.6%, 60.2% and 67.4% of respondents’ described that they mostly used assessment

evidences to categorize students into different groups based on their results, to make the students aware about their achievement against other students' result and to approve students who score high in the assessment task respectively.

The interview result also confirmed that almost all science teachers in the selected schools were guided by the traditional use of assessment results, which have no value for future learning. Teachers were asked to describe: *For what purpose do you use assessment evidence?* For instance:

I used assessment results to classify students into different groups (high achiever, medium achiever and low achiever) and as a result to give tutorial accordingly. Moreover, I used it to create awareness among students about their level of understanding or rank against other students, because I believed that creating competitive environment between students is good to improve students' learning. Finally, I record the result on the mark sheet for final result to decide whether a student passes or fails in the subject. (Teacher E, Nov 14, 2014)

Some teachers used assessment evidences not only for recording and classifying students, but also they used it to improve their teaching pace during the lesson. For example:

Once I score the assessment tasks, I can easily recognize which of my students understood the lesson better and which did not. I then try to modify my teaching methods accordingly. For instance, if most of the students got low marks on the assessment task, I tried to revise the lesson learnt, decrease the pace of my explanation, repeat important points, give different examples and allow them to ask questions. (Teacher F, Nov 18, 2014)

In general, the results of this study clearly indicate that science teachers need intensive support to effectively use the data or the collected assessment evidence to adjust their instruction and to improve students' learning.

### **Interpretation and communication of evidences**

The third highest rated and diversely responded factor is the "Interpretation and communication of assessment evidences" (a mean of 3.39, see Table 4.2). In order of level of practice science teachers stated that the collected assessment evidences were interpreted regularly: against

learning objectives (59.5%, with a mean of 3.61), against assessment criteria or standards (58.1% with a mean of 3.58), compared to other class of students with the same grade level (52.3% with a mean of 3.5) and compared to other students result within the same class (51% with a mean of 3.44).

During the interview session, most teachers also expressed that they identify students' learning gap in the lesson by comparing their result with other students result in each assessment task.

One example:

I usually compared students' score in the assessment task within the class and with other class of students to know students' level of understanding in that lesson. Moreover, such relative analysis helps me to assess how much my lesson was successful compared to other lessons and teachers who teach the same grade level. Knowing students level of understanding in my subject (i.e. high achiever, medium achiever and low achiever) also helps me to form groups for further learning. (Teacher F, Nov 18, 2014)

However, interpreting assessment evidences against the result of other students (i.e. norm referenced evidence interpretation) does not allow teachers and students to clearly identify their learning gaps against learning objectives for the next steps in learning. Moreover, it does not allow students to put their effort to fill the gaps, rather it undermine the self-esteem and confidence of low achieving students.

Moreover, descriptive feedback was not given by most teachers in the selected schools. As mentioned on Table 4.2, 40.6% and 37.9% of teachers agreed that they regularly or often provide only marks and detailed answers for each assessment task along with marks respectively. Evidences from the interview also indicated that teachers regularly provide feedback to the students in the form of numerical scores and correct answers rather than giving constructive messages on how to improve their work in the future to facilitate their learning. However, this kind of feedback did not give direction for students to know where they are in relation to the learning objectives and success criteria to identify the gaps in their learning. From their expression, most of them connect feedback with the provision of correct answers, marks and provide detail answers for difficult questions. For example:

Regarding feedback, after students submit their assignment on the assigned day, I provide marks and do questions that are wrongly done. If it is quiz or test..., I score, write correct answers in front of the questions, and put their mark. Finally, I allow them to see it and ask question if there are difficult question for them. Yet, still I did not give suggestive feedback and time for students to do their works again or to discuss in group. (Teacher B, Nov 19, 2014)

In general, to improve students' learning through formative assessment, students result in any activity or task should be interpreted in line with the intended outcomes of the activities or success criteria and to some extent with their previous achievement. Thus, students' current learning progress and their efforts should be considered during interpretation of evidences and their communication.

### **Planning of formative assessment**

Science teachers in the selected school give less attention for the "Planning of formative assessment" and "Learning contents assessed" factor loadings in the combined practice sub-scale (a mean of 3.38 and 3.37 respectively, see Table 4.2). However, planning different assessment for learning strategies as an integral part of the daily instruction and communicating it clearly with students is the first step in formative assessment cycle.

From the six items which are loaded for the "Planning of formative assessment" factor, 56.2% and 53.6% of science teachers reported that they regularly identify learning objectives and assessment criteria and design better questions and questioning strategies in the planning of their lessons respectively. Similarly, 45.8%, 41.8% and 12.5% of the respondents plan how to share learning objectives and assessment criteria, examine students' prior knowledge in the subject (47.8% 35.3% and 17% of them), design student-centered assessment methods and tasks (41.8% 34.6% and 23.6%) and plan how and when to provide feedback (30.8%, 37.9% and 31.4%) frequently, occasionally and once in a blue moon respectively (see Table 4.2 below). Thus, the quantitative data indicates that the planning of formative assessment strategies as an integral part of the lesson preparation prior to collecting learning evidences was practiced by few science teachers in the selected schools.

Correspondingly, as to the expression of most interviewed teachers, their lesson preparation conventionally includes: rationales, learning objectives; unit contents; teacher's activity such as reviewing previous lesson, explaining important points, giving short notes, asking question...; students' activity such as taking notes, listening, answering questions...; evaluation (e.g. oral questions, classwork, home work from the text books); and summarization of main points in the lesson. Thus, the culture of lesson preparation in the selected schools was also discouraging to plan formative assessment strategies as an integral part of it. However, currently, the lesson format was changed regionally to include formative strategies as a part of their lesson preparation, but no one is aware of it and they do it accordingly. Two examples were included for this explanation as:

...still I did not plan formative assessment strategies such as self-assessment, peer assessment, self-reflection, peer questioning, student to student dialogue, sharing of learning objectives and assessment criteria, and when and how to provide descriptive feedback. The reason is that there is no such trend in our lesson preparation. (Teacher B, Nov 19, 2014)

Mostly, I include questions as one part of the lesson planning to evaluate students' level of understanding in each phase of the lesson. However, still I did not design the questions in line with the lesson objectives. Many of them are simple oral questions, because, most of my students are low achievers who passed the national examination through cheating. (Teacher A, Nov 20, 2014)

But, including formative assessment strategies as one part of their lesson preparation helps teachers and students to collect learning evidences related to students' knowledge, skills and attitude in the lesson and, as a result, to use such evidences as an input to improve students' learning and to adjust instruction. Moreover, it encourages higher order learning and active engagement of students during the lesson.

### **Learning contents assessed**

From the five items which loaded the "Learning contents assessed" factor, reasonable numbers of teachers (62.7%, 49.6% and 49%) often focus on assessing difficulties during the teaching learning process, knowledge of scientific facts and higher order thinking respectively. However, the application of problem solving skills and scientific reasoning ability were assessed

occasionally (37.3% and 35.9% of respondents respectively). And 28.1% and 26.8% of science teachers hardly ever assessed such higher order thinking in the selected schools.

The interview results also support it. Most interviewed teachers expressed their lack of focus on higher order thinking in the learning as well as in the assessment process due to students' different background. For instance:

Mostly, I asked factual questions that are short and precise. My question does not need much effort and calculation, if they understand the concept they can easily answer it. I did not ask questions that are difficult. As to me, in preparatory schools there are two extreme students, high achiever and low achiever. Only medium students are failed to pass the national exam. Thus, it is difficult to teach and assess higher order thinking in such classroom environment. For me, it is better to teach these students by classifying them into two groups to make all of them competent learners. (Teacher G, Nov 18, 2014)

...I mostly include objective type questions such as true/false, matching, completion, multiple choice and short answer questions in quizzes, tests, mid exams and final exams. Also, I give review questions from the text book in the case of assignments, class works and home works, since, students have different backgrounds and the time is too short to score subjective type questions and to show results for all students before the prescribed time. (Teacher C, Nov 18, 2014)

Lesson observations also verify it. In all of the observed lessons, divergent questions were not emphasized in the teaching-learning process, which encourage higher order thinking. Example of questions raised during lesson observation were: "what is biomass?", "how many types of succession are there?", "write the set of permissible quantum number for electron found in the outer most energy level of Potassium atom?", "when do metals expand?" etc., which did not let students to apply higher order thinking, problem solving skills and reasoning abilities in new issues. Moreover, as indicated in Table 4.2, only fast learners were involved to answer such questions. Even if some teachers ask higher order questions such as "how water is so important for living things?" they did not allow students to share and reflect ideas in group, rather they dominantly explain it throughout the lesson.

## Collection of learning evidences

Collection of learning evidences is the second step in formative assessment cycle next to planning. This factor addresses how different formative assessment strategies are integrated in the teaching learning process to collect information about students' knowledge, skills and attitude in that lesson to decide next steps in learning. However, compared to other factor loadings in the practice sub-scale "Collection of learning evidences" factor is the least practiced formative assessment phase in the selected schools (a mean score of 2.87).

Here, 67.32% of science teachers did not put into practice different assessment for learning strategies throughout their daily instruction to collect learning evidences for further learning in the selected schools. Particularly, large number of teachers hardly put into action: criteria and objective setting with students (2.23), self-reflection through drawing and concept mapping (2.48), quizzes (2.5), self-assessment (2.57), written feedback (2.57), peer feedback (2.65), peer-assessment (2.69), practical work (2.86), peer to peer questions (2.79) and students' reflection of ideas on the lesson learnt (2.99), which are the main components of assessment for learning to collect evidences.

Similarly, other assessment for learning methods which loaded this factor such as student-to-student dialogue (3.30), observation (3.19), oral feedback (3.17), presentation (3.13), teacher-to-student dialogue (3.05) and provision of written comments on how to improve their work (3.05) were practiced occasionally. While, 54.9 % of science teachers regularly ask self-evaluation questions at the end of their lesson (with a mean of 3.52), which is important to see the achievement of learning objectives, but it has little value to provide information for further learning.

Gathering learning evidences related to students' knowledge, skill and attitude using various formative assessment strategies should be a part of the instruction to identify learning gaps and to propose means for next steps in learning. Yet, the results of this study give the impression that teachers in the selected schools rarely collect this relevant learning information using such activities in their lesson to decide the next steps in learning (see Table 4.2 below).



**Table 4.2 Factor grouping of practice of assessment for learning items with factor loadings, mean, and percentage of teachers' response**

No.	Items	Factor loading	Mean	Rarely	Sometimes	Often
				%	%	%
	<b>Planning of Formative Assessment</b>	<b>5.8%</b>	<b>3.38</b>			
1	Identify learning objectives and assessment criteria	0.738	3.60	10.5	33.3	56.2
2	Plan how to share learning objectives and assessment criteria	0.675	3.46	12.5	41.8	45.8
3	Design student -centered assessment methods and tasks	0.638	3.26	23.6	34.6	41.8
4	Examine students prior knowledge in the subject	0.595	3.39	17.0	35.3	47.8
5	Plan how and when to provide feedback	.585*	3.03	31.4	37.9	30.8
6	Design better questions and questioning strategies	0.527	3.55	12.4	34.0	53.6
	<b>Collection of Learning Evidences</b>	<b>10.4%</b>	<b>2.87</b>			
7	Quizzes	0.366	2.50	49.1	43.8	7.2
8	Practical work	0.352	2.86	36.6	42.5	20.9
9	Presentation	0.377*	3.13	26.8	40.5	32.7
10	Self-assessment	0.718	2.57	41.8	34.6	23.5
11	Peer-assessment	0.771	2.69	49.6	32.0	18.3
12	Peer-to-peer questions	0.514	2.79	41.2	32.7	26.1
13	Peer feedback opportunities	0.728	2.65	44.5	35.9	19.6
14	Oral feedback	0.414	3.17	26.1	34	39.9
15	Written feedback	0.678	2.57	44.5	35.9	19.6
16	Set criteria and objective with students	0.669	2.23	62.8	18.3	18.9
17	Students' feedback reflection of ideas on the lesson learnt	0.683	2.99	29.5	39.2	31.4
18	Self-reflection using drawing, concept mapping, ...	0.666	2.48	55.6	27.5	17

No.	Items	Factor loading	Mean	Rarely	Sometimes	Often
				%	%	%
19	Student- to- student dialogue	<i>0.490*</i>	3.30	21.6	32.7	45.8
20	Teacher –to- student dialogue	<i>0.461*</i>	3.05	32.7	32.0	35.3
21	Observation	<i>0.454</i>	3.19	26.8	33.3	39.9
22	Provide written comments on how to improve their work	<i>0.431</i>	3.05	32.7	30.7	36.6
23	Self-evaluation questions at the end of the lesson	<i>0.369</i>	3.52	16.3	28.8	54.9
<b>Interpretation and Communication of Evidences</b>		<b>7.60%</b>	<b>3.39</b>			
24	compared to other students’ result within the class	<i>0.715</i>	3.44	16.4	32.7	51
25	against assessment criteria or standards	<i>0.683</i>	3.58	13.1	28.8	58.1
26	against their own previous performance	<i>.681*</i>	3.40	14.4	42.5	43.2
27	compared to other class of students with the same grade level	<i>0.639</i>	3.50	16.4	31.4	52.3
28	against learning objectives	<i>.580*</i>	3.61	10.5	30.1	59.5
29	provide the detailed answers for each assessment task along with marks	<i>-0.425</i>	3.06	34	28.1	37.9
30	provide only written marks for each assessment tasks	<i>.403*</i>	3.17	26.8	32.7	40.6
<b>Support Provided to Engage Students Actively</b>		<b>9.40%</b>	<b>3.73</b>			
31	Encourage students to share ideas	<i>0.736</i>	3.88	9.8	17.6	72.5
32	encourage every student to ask questions	<i>0.725</i>	4.05	7.8	13.1	79
33	engage every student to answer questions	<i>0.625</i>	3.93	5.9	21.6	72.6
34	encourage students to take risks and listen to others ideas carefully	<i>0.613</i>	3.76	11.8	21.6	66.6
35	advise students to assess their own work against learning objectives	<i>0.577</i>	3.54	16.4	26.8	56.9
36	Ask oral questions	<i>0.562</i>	3.84	5.9	28.8	65.4
37	Encourage class participation	<i>0.518</i>	4.06	4	18.3	77.8
38	encourage students to answer questions quickly	<i>0.516</i>	4.06	5.2	13.7	81.1

No.	Items	Factor loading	Mean	Rarely	Sometimes	Often
				%	%	%
39	Give class work	<i>0.504</i>	3.67	8.5	36.6	54.9
40	create opportunities for students to act on feedback provided	<i>.498*</i>	3.75	7.2	27.5	65.4
41	Give homework	<i>0.454</i>	3.88	5.9	24.2	69.9
42	provide examples of quality work that shows the standards required	<i>.422*</i>	3.53	14.4	34.6	50.9
43	advise students to assess others' work against learning objectives	<i>.391*</i>	3.18	26.2	35.3	38.6
44	repeat the learning objectives and criteria during the lesson	<i>0.387</i>	3.50	15.7	33.3	51
45	Provide group assignment	<i>0.377</i>	3.26	17.6	50.3	32.1
	<b>Learning Contents Assessed</b>	<b>6%</b>	<b>3.37</b>			
46	application of problem solving skills in new issues	<i>0.77</i>	3.10	28.1	37.3	34.6
47	Higher order thinking (i.e. analysis, synthesis, ...)	<i>.723*</i>	3.46	18.9	32.0	49
48	knowledge of scientific facts	<i>0.721</i>	3.44	18.9	31.4	49.6
49	scientific reasoning ability in new situations	<i>0.693</i>	3.13	26.8	35.9	37.3
50	Difficulties during the teaching-learning process	<i>.569*</i>	3.71	11.1	26.1	62.7
	<b>Application of Assessment Evidences</b>	<b>10.7%</b>	<b>3.70</b>			
51	Advise students about how to fill the gap in their learning	<i>0.745</i>	3.92	6.5	19	74.5
52	Modify my teaching strategies accordingly	<i>0.74</i>	4.04	5.2	14.4	80.4
53	Plan what to teach next	<i>0.737</i>	3.88	6.6	23.5	69.9
54	Approve students who score high result	<i>0.667</i>	3.83	9.8	22.9	67.4
55	Tell their achievement on a task against other students' result	<i>0.66</i>	3.66	13.1	26.8	60.2
56	Identify the gaps in students' understanding	<i>0.651*</i>	3.92	4.6	24.8	70.6
57	Categorize students into different groups	<i>0.644</i>	3.73	7.8	29.4	62.7
58	Orally suggest on how to improve their work	<i>0.639</i>	3.40	18.3	32.0	49.7

No.	Items	Factor loading	Mean	Rarely	Sometimes	Often
				%	%	%
59	Suggest means for students to plan their future learning	0.570*	3.46	20.9	27.5	51.7
60	Allow peer discussion on how to improve their work	0.516	3.42	16.4	36.6	47.1
61	Record assessment results	0.422*	4.10	5.9	15	79.1
62	Permit students to resubmit their work once they improved it	0.364*	3.46	17.6	30.7	51.6
63	Give written questions in group	0.354*	3.31	15.7	47.7	36.6

**Note:** \* In the factor loading column indicates items cross loaded but included in this factor

The observation result also confirms it. In the observed lessons, there were no student to student and teacher to student dialogues; self-reflection through ideas, drawings, concept mapping...; self-and peer assessment; provision of constructive feedback; peer to peer questions; and except one lesson observation, others even did not write their learning objectives on the board; which are the main components of formative assessment to collect learning evidences and as a result to integrate it into the lesson to improve students' learning.

In the interview, the researcher explored the use of formative assessment tools in their lessons to collect learning evidences. When I asked teachers about formative assessment tools they used in their own lesson all teachers appeared to share similar practices. The common assessment methods they employed were tests, assignments, mid exams, home work, and oral questions and for some class works. However, these assessments are not effectively integrated in their daily instruction, but they are given at the end of the lesson or chapter to check students' understanding. Moreover, evidences collected through such assessment methods cannot reflect the full range of learning goals to identify learning gaps; rather, it will be applicable for recording and reporting results.

Teacher A, for example, said that:

As I am a mathematics teacher, I regularly use oral questions, class works and home work to know students' existing knowledge about the chapter and to assess how much students understand the lesson learnt. It also helps me to identify high achiever, medium achiever and low achiever students; and it helps to give more attention for questions that are not well done by the students and to clarify difficult concepts for students. Moreover, after two and three weeks, I used tests, assignment and mid exam to evaluate students' understanding about the chapters and to collect marks for their final result. (Teacher A, Nov 20, 2014)

Some teachers viewed assessment and marks as two sides of the same coin and practiced accordingly, rather than embedding assessment with their daily instruction to collect learning evidences for the purpose of further learning. One example:

...I always give mark for any assessment task. For example, when I give group assignment ...I randomly call students to reflect what they do in the group. It helps me to identify those students who actively participate in the group and who did not, to give

marks accordingly. ....moreover, I give marks for quizzes, test, assignments, class attendances... after I do questions in the class and then I allow them to see it. (Teacher E: Nov 14, 2014)

Some teachers commented on students' lack of attention for assessment activities or tasks that have no marks. Most of them showed a preference for taking tests and mid exams in the traditional manner. One example:

... I did not give mark for class works, worksheet, and questioning, but I always threatened students by saying as all activities have marks in order to allow all of them to do the activity, because our students traditionally give more attention for activities that have marks. (Teacher G, Nov 18, 2014)

Moreover, during the interview sessions, teachers were asked to describe their experience in sharing learning objectives and assessment criteria for students, and implementing self-and peer assessment. However, there was not much evidence that pointed whether the teachers use these formative assessments or assessment for learning strategies regularly with their students to gather evidences. Most felt that large number of students in one class and limited instructional time were factors for them to effectively integrate such strategies into their lesson. That is why; most of them did not feel good to implement such assessment for learning strategies in their lesson to improve students' learning. Thus, they did not give clear direction for their students to have knowledge of where they are going, where there now, and which strategies will help them to achieve the learning objectives.

Some teachers described that involving students in assessing their own and others work and for other processes involved in assessment as vital to improve students' learning; yet, they find themselves limited by time and large content coverage. Example:

I want to tell the truth...Still I did not share learning objectives and assessment criteria for students and I did not implement self-and peer assessment, because the time given is very short and the content covered is too large to implement such assessment strategies. However, such assessment systems are very good to improve learning if I have time to employ them. (Teacher A, Nov 20, 2014)

One mathematics teacher used self-and peer assessment as a "time saver" for marking students work. He said that:

...sometimes I used self-and peer assessments to save time. If there is a class work, I give marks at least for five individuals and then allow other students to give marks for themselves based on the answers given and sometimes I allow them to give marks for their peers. However, I do it not for the sake of helping students to learn but to make all of them happy, because they may feel bad if I only score the works of some students. (Teacher F, Nov 18, 2014)

Except, tests, assignments, oral questions, home works, class works and mid exams, that are employed at the end of the lesson, other assessment for learning strategies such as sharing of learning objectives and success criteria, self-and peer assessment, descriptive feedback, better questions and questioning strategies etc., were not strong points for most of the teachers in the selected schools. Even, most of the participants realized that such words are new for them. Generally, most teachers seemed to still work in the view of traditional learning theories that formative assessment is a separate element that is not integrated with daily instruction but a tool that comes at the end of lesson to evaluate students' learning. Therefore, there is lack of continuous collection of students' learning evidences during the teaching learning process to adjust instruction and to fill the gaps in students' learning due to many factors.

**Research question two:** What are the major factors that affect teachers' practice of assessment for learning in science subjects?

Mean, standard deviation, and percentage of respondents for each item were computed to identify teachers' level of agreement on the possible factors that hinder the implementation of assessment for learning in science subjects in the selected schools. Based on the results of factor analysis, the possible factors of assessment for learning practices were reduced into three factor loadings. These are: Resources, Awareness and Support related factors (see Table 3.3). Descriptive statistics were calculated for the three factor loadings and the results were as follows: Resource related factors (a mean of 3.90), Awareness related factors (a mean of 3.64) and Support related factors (a mean of 3.54) respectively. Moreover, the one-sample t-test revealed that there is a statistically significant difference among the mean of the three factors loading of the possible factor scale. Thus, based on the results, the three factor loadings were presented below according to their possibility to be a factor of assessment for learning practice in science subjects in the selected schools.

### **Resource related factors**

Most science teachers reported that four items which loaded the “Resources” factor were more likely that hinder the practice of assessment for learning in science subjects in the selected schools; particularly, large number of students in one class (85.7% of respondents) and lack of available science resources (such as textbooks, lab rooms, lab equipment, demonstration sites...) (83.6% of respondents) delay their assessment for learning practice. Moreover, 76.4% and 51.7% of the respondents agreed on the impact of lack of instructional materials (e.g. teacher’s assessment guideline...) and summative assessment (e.g. mid exams, final exams, national exams...) on the integration of assessment for learning strategies in their daily instructions respectively.

During classroom observation, the researcher tries to observe the availability of instructional materials and the number of students in the eight classrooms. The result showed that: the number of students is on average 55, the seating arrangements are u-shaped, the chairs and desks are permanently connected for three students, and there is at least one text book in each desk. Moreover, in one of the newly established schools, there is armchair for each student which is easily movable and the number of students is on average 45, but the seating arrangement was facing toward the blackboard for lecturing. Moreover, two lab rooms were observed from the older and newly established schools. The result showed that compared to the newly established school, there were fair laboratory equipments in the older schools. However, there was no well-organized and equipped laboratory rooms observed in the two schools. Thus, accessing and organizing different resources and reducing the number of students in one class are vital, particularly to implement inquiry based science education through formative assessment in the selected schools to achieve the desired millennium goals.

### **Awareness related factors**

“Awareness” factors were also indicated by most science teachers as likely factors for assessment for learning practices in science subjects in the selected schools. 72.6%, 68.6%, 64.1% and 62.7% of teachers showed that shortage of instructional time, lack of professional



development activities (such as in-service assessment training, pre-service assessment courses, workshops, etc.), students' and teachers' negative perception on formative assessment are possible factors that affect the implementation of assessment for learning in science subjects in the selected schools respectively.

The classroom observation results also confirm it. During lesson observation, most teachers did not share learning objectives and assessment criteria with students, self-and peer assessment were not implemented, divergent and convergent questions were not asked, enough thinking time was not given during questioning. Moreover, most students passively listen and write what the teacher says, rather than actively engaging themselves through reflection, questioning and answering. These all happen due to teachers' lack of appropriate understanding about formative assessment strategies and students too.

### **Support related factors**

Five items which loaded "Support" factor were the third ranked possible factor of the implementation of assessment for learning in science subjects as mentioned by most teachers. Particularly, lack of support from principals and colleagues were charged as likely factor by 70.6% and 71.7% of respondents with mean of 3.7 and 3.65 respectively. In the same way, 66.7% (3.61), 64.1% (3.51), and 51% (3.24) of the respondents agreed that lack of support from supervisor, government mandates and pressure of national examination are likely factors respectively (see Table 4.3 below).

**Table 4.3 Factor groupings of the possible factor items with factor loadings, mean, and percentage of teachers' response**

No.	Items	Factor Loading	Mean	Unlikely	Neither	Likely
				%	%	%
<b>Support related factor</b>		<b>22%</b>	<b>3.54</b>			
1	Lack of support from school principals	.779	3.7	15.7	13.7	70.6
2	Government mandates on assessment issues	.740	3.51	21.6	14.4	64.1
3	Lack of support from supervisors	.736	3.61	17.6	15.7	66.7
4	Lack of support from colleagues	.721*	3.65	16.3	12.4	71.2
5	Pressure of national examinations	.613	3.24	28.1	20.9	51
<b>Awareness related factor</b>		<b>19.4%</b>	<b>3.64</b>			
6	Students negative perception on formative assessment	.837	3.61	19.6	16.3	64.1
7	Teachers' negative perception on formative assessment	.830	3.53	19	18.3	62.7
8	Lack of professional development activities	.785	3.63	20.3	11.1	68.6
9	Shortage of instructional time	.605	3.79	15.7	11.8	72.6
<b>Resources related factor</b>		<b>14.6%</b>	<b>3.9</b>			
10	Large number of students in one class	.693	4.39	7.2	7.2	85.7
11	Lack of available science resources (textbooks, lab rooms, lab equipment, demonstration sites...)	.607*	4.21	9.1	7.2	83.6
12	Impact of summative assessments (mid exams, final exams...)	.574	3.23	27.4	20.9	51.7
13	Lack of instructional materials (e.g. teacher's assessment guideline...)	.533*	3.95	12.4	11.1	76.4
14	Many number of periods per week /teaching load/	.502	3.7	21.6	12.4	66.1

**Note:** \*In the factor loading column indicates items cross loaded but included in this factor

Moreover, science teachers were interviewed to elaborate on what some of the challenges they observed in the integration of formative assessment strategies into their daily instruction. In answering this question, both teachers felt that there were challenges in the implementation. These are: limited time given for one lesson (40 minutes), students' different understanding level (high, medium, and low achiever), and large number of students in one class. Similarly, teachers added that bulky content coverage (particularly, grade eleven mathematics text book), students' negative perception on assessment (such as self-and peer assessment), teachers' lack of knowledge and skills and their negative perception on formative assessment strategies and its' implementation and their long experience on the use of summative assessment are other factors of implementing formative assessment strategies effectively in science subjects. Examples include:

There are challenges to implement formative assessment strategies effectively in our school. For instance, large number of students in one class does not allow us to integrate formative assessment effectively in our daily instruction, because it is difficult to score and give feedback for each student and to see the difficulties of students by approaching them. (Teacher A, Nov 20, 2014)

...different understanding levels of students (high, medium and low achievers), shortage of instructional time, and broad content coverage are major factors that hinder me in using formative assessment strategies effectively. (Teacher C, Nov 18, 2014)

Some teachers focused on students' lack of awareness on what was expected from them in the teaching-learning process and their negative perception on formative assessment, saying:

Students' background and experience in using self-and peer assessment and their level of understanding were the main factors that hinder the implementation of formative assessment. When I allocate tasks and permit them to assess themselves and their peers, they are not volunteer, rather they see it carelessly. (Teacher E, Nov 14, 2014)

Students mostly perceived assessment as a tool to collect marks in order to decide whether they pass or fail the required grade. For instance, if a student gets low mark on the given assessment task, he or she argues with/ begging me to add extra marks, rather than she/he provides the opportunities for me to show his/her the correct answers and how to do it. My students' have also negative perception on their role in the learning process. All of the students expect as everything is done by the teacher. (Teacher B, Nov 19, 2014)

Majority of the teachers commented on the quality of teachers' professional knowledge and skill on assessment, because such pedagogical knowledge is essential to implement student-centered assessment and learning process through understanding students' difficulty and needs in learning, asking questions that stimulate productive thinking, interpreting their responses in line with the learning objectives, and using it for the next steps in learning. One example:

Teachers' lack of knowledge and skills to use formative assessment methods, their negative perception about formative assessment strategies and long experience in using summative assessment are the possible factors to integrate formative assessment strategies in our daily instruction. (Teacher H, Nov 18, 2014)

In addition, one grade eleven mathematics teacher felt that lack of integration in the curriculum content and many number of holly days are a factors for implementing formative assessment. For instance, he stated that:

...curriculum contents are not integrated vertically and horizontally to allow students to do by themselves and it is too vast to cover within the prescribed time and many numbers of holly days do not allow us to implement active learning methods and formative assessment strategies effectively. (Teacher F, Nov 18, 2014)

Teacher G believed that lack of motivation among teachers is a main factor in implementing formative assessment as a part of teaching and learning process to improve students' science learning standards.

I perceived that lack of motivation among teachers and lack of attention given to teachers are hindering factors of assessment for learning practice. Mostly, teachers consider schools as places of leisure, rather than considering them as a place of work. Most of the teachers did not believe on their work. They do it, because there is no other option. This comes because of lack of motivations, incentives, supporting environment, low salary.... Public Medias also give low attentions for teachers' innovative works. Moreover, competent teachers and students are not selected from each school to let them to observe what others do in order to do their own innovative works in science and technology department. (Teacher G, Nov 18, 2014)

**Research question three:** How do second cycle secondary schools science teachers perceive assessment for learning?

As mentioned above (see Section 3.8.1) factor analysis were done to reduce the number of items in the teachers' perception sub-scale into three factor loadings. These are: Teachers' perception on the power of formative assessment to improve learning, students' involvement of, and the provision of feedback (i.e. formative or constructive feedback).

Descriptive statistics were computed for the three factor loadings and for each item in each factor to determine areas of formative assessment that teachers positively and highly perceived (see Table 4.4). One-sample t-test also was used to assess significant differences in the mean of the three factor loadings in the perception sub-scale. The result showed statistically significant differences between the means of the three factors in the perception scale ( $p=0.000$ ). The results were presented below.

### **Teachers' perception on the power of formative assessment to improve learning**

As indicated in Table 4.4, most science teachers reported higher level of agreement in each item in both factors. However, items loaded on teachers' perception of the "Power of formative assessment to improve learning" factor were highly perceived by most science teachers in the selected schools. Particularly, most teachers (92.2%, 91.5%, and 90.2%) agreed on the idea that formative assessment is an integral part of the teaching learning process, that helps them to easily identify students' problem in learning, and it allows students to know more about their own learning problems respectively.

Relating to its application, 80.4% of the teachers believed that formative assessment is vital to assess higher order thinking, to ensure all students to have an experience of success and it makes all of them competent learners (83%), and it improves every student's achievement in science subjects (82.3%). Likewise, most science teachers in the selected schools agreed that formative assessment is a means to capture students' attention and efforts (83.6% and a mean of 4.07), to start their study from day one (81% and a mean of 4.02), to develop positive self-esteem and confidence among all students (86.9%), and to make them independent learners (75.8%) by fostering their internal motivation to learn science subjects (79.7% and a mean of 4.03).

Similarly, reasonable numbers of teachers seem to realize the importance of mixing formative assessment with daily science lesson to improve learning, to assess the effectiveness of the teaching learning process and to decrease students' dropout and repetition in one class (86.9%, 80.4%; and 72.5%) than summative assessment respectively. Moreover, most of the science teachers reported that sharing of learning objectives and assessment criteria are vital to motivate students to learn (83.1%), to enable them to assess their own progress (83%) so that they improve their learning (86.3%). Thus, the result of this study gives the impression that most teachers in the selected schools have positive perception about the power of formative assessment to improve students' science learning.

However, evidences from teachers' response on item number 17, on the same table indicated that 86.3% of the respondents negatively perceived formative assessment as a process that follows the teaching-learning process which supports the behavioural learning perspectives on assessment. This means that formative assessment is viewed as a tool that is used to evaluate students' progress continuously in the lesson learnt rather than a means to improve their learning. Such contradictory responses of teachers on the same table indicate their misconception about formative assessment. The lesson observation and interview result also validate it. Even if, most teachers seem to have positive perception on the power of formative assessment on learning, the qualitative data revealed that many of these teachers did not effectively integrated formative assessment into their daily instruction.

### **Teachers' perception on the active involvement of students**

From the three factor loadings in teachers' perception of assessment for learning sub-scale, the second ranked factor was teachers' perception on the "Involvement of students" in the lesson. Most teachers (90.8%) reported high level of agreement (perception) on the importance of active involvement of students in the assessment process to improve their learning. Along with, as to the response of most teachers', the implementation of formative assessment allows such active engagement of students (90.2% of respondents) in the teaching learning process.

In addition, teachers agreed on the significant role of formative assessment strategies such as the provision of constructive feedback (89.6%), self-assessment (86.3%), peer assessment (77.7%), and questioning (84.3%) for active involvement of every student in the lesson to improve science learning. Reasonable number of teachers also highly perceived the advantage of giving long thinking time during questioning to engage every student throughout the lesson (77.1% with a mean of 3.81). Thus, science teachers have high level of positive perception about students' active engagement in the lesson to improve their learning in the selected schools. And they perceived formative assessment strategies as a means for such active engagement. However, what was evident from the actual practices was that those teachers participated in this study did not actively involve their students in the lesson through student-centered formative assessment strategies. The result indicates a mismatch between what teachers perceived and what they actually practiced in the classroom. This happened because of skill gap among science teachers in the selected schools.

### **Teachers' perception on the provision of formative feedback**

Teachers' level of agreement on the items that are related to the "Provision of feedback" factor was also high. Particularly, 95.5% of the respondents agreed on the purpose of using variety of assessment methods to get useful insights about students' progress. And they viewed formative assessment as the solitary means that used various assessment methods to gather relevant learning evidences (90.1%) and to inspire students' deep thinking in science subjects (87.6%). Moreover, they positively perceived the provision of formative feedback as one of the key component in formative assessment (87.6%), to identify gaps and to fill such gaps in students' understanding (86.3%), and to support or inhibit students' motivation to learn (54.9%) (see Table 4.4). In contrast, sensible number of teachers in the selected school believed on the provision of detailed correction of students' work (74.5%) and marks (64%) as effective way of feedback to improve science learning and to provide directions for students about their progress against the learning objectives (see Table 4.4 below). The qualitative data also showed that the majority of science teachers mostly provide mark for any of the assessment tasks, grant with correct and wrong answers, allow students to see the results, and finally record it. It looks like that teachers in the selected schools are dominantly challenged by the traditional way of judgmental feedback.

Thus, in this study teachers were not accustomed to delivering feedback in a constructive manner to identify students' learning gaps and to decide next steps for future learning.



**Table 4.4 Factor groupings of perception items with factor loadings, mean, and percentage of teachers' response**

No.	Items	Factor Loading	Mean	Disagree	Neutral	Agree
				%	%	%
	<b>Impact of formative assessment on learning</b>	<b>21.30%</b>	<b>4.08</b>			
1	Formative assessment helps students to develop positive self-esteem and confidence.	0.768	4.18	7.2	5.9	86.9
2	Formative assessment helps students to be independent learners.	0.735*	3.96	9.1	15	75.8
3	Formative assessment is vital to ensure all students to have an experience of success and makes all of them competent learners.	0.719	4.05	4	13.1	83
4	Formative assessment improves every student's achievement in science subjects.	0.717*	3.99	7.9	9.8	82.3
5	Transparency in learning objectives and assessment criteria enable learners to assess their own progress.	0.635	4.07	4.6	12.4	83
6	Formative assessment is vital to assess students' higher order learning (e.g. application, analysis, synthesis etc.) more than summative assessments.	0.625	4.01	9.8	9.8	80.4
7	Formative assessment empowers students to study from day one class.	0.623*	4.02	8.5	10.5	81
8	Sharing learning objective and assessment criteria to students is vital to improve learning.	0.613*	4.1	3.9	9.8	86.3
9	Formative assessment is important to capture students' attention and effort.	0.599	4.07	6.5	9.8	83.6
10	Sharing learning objectives and assessment criteria motivates students to learn.	0.599	3.94	9.1	7.8	83.1
11	Mixing formative assessment with daily lesson is useful to improve science education.	0.590*	4.1	6.5	6.5	86.9
12	Formative assessment helps students to know more about their own learning problems.	0.588*	4.24	5.3	4.6	90.2
13	Formative assessment fosters students' internal motivation to learn science subjects.	0.579*	4.03	6.6	13.7	79.7
14	Formative assessment reduces the rate of repetition and dropout more than summative assessment.	0.524	3.82	14.4	13.1	72.5
15	Formative assessment is an integral part of the teaching-learning process.	0.449*	4.31	4.6	3.3	92.2
16	Formative assessment is more vital to assess the effectiveness of teaching than final exams.	0.446	4.04	10.4	9.2	80.4

No.	Items	Factor Loading	Mean	Disagree %	Neutral %	Agree %
17	Formative assessment follows the teaching learning process.	0.429*	4.15	6.6	7.2	86.3
18	Formative assessment helps teachers to easily identify students' problems in learning.	0.427*	4.29	5.9	2.6	91.5
	<b>Involvement of students</b>	<b>12.20%</b>	<b>4.07</b>			
19	Active involvement of students in the assessment process is vital to improve learning.	0.751	4.28	2.7	6.5	90.8
20	Suggesting students on how to improve their work is effective feedback in learning.	0.716*	4.29	2	8.5	89.6
21	Using self-assessment in science subjects improves students' learning.	0.673	4.06	5.9	7.8	86.3
22	Using peer assessment in science subjects fosters students' learning.	0.625	3.85	8.5	13.7	77.7
23	Questioning is effective to engage every student in the lesson through thinking.	0.541	3.95	8.5	7.2	84.3
24	Formative assessment allows teachers to make their students active.	0.466*	4.22	5.3	4.6	90.2
25	Long waiting time during questioning is vital to engage every student in answering.	0.45	3.81	13.8	9.2	77.1
	<b>Provision of formative feedback</b>	<b>11.20%</b>	<b>4.05</b>			
26	Feedback is helpful to identify and fill the gaps in students' science understanding.	0.715	4.22	3.3	10.5	86.3
27	Teacher-student dialogue during feedback is effective in fostering science learning (i.e. dialogue refers to exchange of ideas).	0.685*	4.08	4.6	10.5	85
28	Using a variety of assessment methods provides useful insights about students' progress.	0.579	4.55	3.3	1.3	95.5
29	Feedback is a key component of formative assessment to improve learning.	0.571	4.22	4.6	7.8	87.6
30	Detailed correction of students work is effective way of feedback to improve learning.	0.497	3.86	12.5	13.1	74.5
31	Formative assessment uses a variety of assessment tools to gather evidence of learning.	0.452	4.31	4	5.9	90.1
32	Formative assessment inspires students' deep thinking in science subjects.	0.437	4.2	3.3	9.2	87.6
33	Scores provide direction for students about their progress against learning objectives.	0.413	3.61	13.8	22.2	64
34	Feedback is a powerful technique to support/inhibit students' motivation to learn.	0.391	3.42	23.5	21.6	54.9

**Note:** \*In the factor loading column indicates items cross loaded but included in this factor

The following points extracted from the interview transcript which provides insights about science teachers' perception on formative assessment. Some teachers considered formative assessment as a tool used to evaluate students' understanding in the teaching learning process to know how much they understand the day to day learning concepts. For example:

Formative assessment is an assessment that is used to evaluate students day to day lesson understanding and it includes, class work, homework, oral questions, assignments, tests... that are frequently given at the end of the lesson, whereas, summative assessment is used to evaluate their understanding at the end of the semester or course. (Teacher A, Nov 20, 2014)

Other teachers perceived formative assessment as a good tool to capture students' attention and effort throughout the lesson, because as to them, it is a continuous and unexpected process. One example:

...formative assessment is a day to day activity to know students' understanding about the lesson. It also helps students to be ready in each day; to be active participant in the lesson, and to be conscious in their learning, because they assumed that the teacher may give assessment tasks that have mark or ask them randomly. (Teacher C, Nov 18, 2014)

Some teachers perceived formative assessment as a more frequent assessment method that empowers students but they consider it a good murder for themselves. One example:

Formative assessment is a continuous process that helps students to know their level of understanding in the lesson and to have good understanding about the content. It also helps teachers to make healthy teaching-learning process and to categorize students into different groups based on their ability (i.e. to make 1 to 5 grouping). However, it is a very tiresome and difficult task for teachers. (Teacher B, Nov 19, 2014)

Some teachers have good understanding about the concept of formative assessment, but they interpret and use it wrongly. One example:

...particularly, formative assessment is a process that starts at the beginning of the lesson to assess students' day to day performance and to identify students' needs and to assist them accordingly. Thus, it is used to evaluate students' ability in a certain activity, to classify students into different groups, and to give tutorial according to their group and to give feedback for students to know themselves where they are. (Teacher H, Nov 18, 2014)

Some teachers were in doubt about the power of formative assessment to improve students' learning; they consider it as a day to day /continuous/ collection of marks, for example: Teacher E stated that:

If we did not assess students' understanding continuously, we did not know students' daily progress in a subject. ...however, if we implement formative assessment continuously, we are not teaching students to make them competent and to achieve the required competencies (skills, knowledge and attitude) in the subject. Our work becomes continuous assessment, which does not help students to be competent in a subject. Formative assessment by itself is not an active teaching method, but it helps us to modify our teaching methods, because when students results become low the teacher asks him/herself to change his/her methodology to make all students high achiever or to modify their results. (Teacher E, Nov 14, 2014)

Teacher F shared different perspective: He expressed formative assessment as an active teaching method and a way of evaluating the achievement of curriculum objectives. He stated that:

Formative assessment is used to evaluate the three domains of learning (i.e. the affective, cognitive and psychomotor domains) continuously and to identify students learning gaps. Thus, I used formative assessment as a teaching method for example, I provide class works, ask oral questions and give a chapter to discuss and present it in class simply for their learning without allotting mark. For me, formative assessment is one of the active teaching methods, but in our school teachers see it separately from instruction. (Teacher F, Nov 18, 2014)

Some teachers also expressed their lack of awareness about formative assessment strategies. And they feel that they needed appropriate continuous professional trainings to implement formative assessment strategies in their lesson. One example:

... I used formative assessment tools once a week to make all students competent to pass into the next grade and to make all of them good citizens. In addition to this, I used oral questions, group discussion, class activity etc. but I am not aware whether such activities are formative assessment methods or not. I simply implement it. (Teacher G, Nov 18, 2014)

**Research question four:** To what extent do professional (pre-service and in-service) or assessment trainings help science teachers to integrate assessment with their daily instruction?

As pointed out in Table 4.5, except 16.3% of the respondents, most science teachers took classroom assessment courses during pre-service trainings. And from those teachers, 71.9% of them have taken two assessment courses during this time. Besides, 80.4% of the teachers reported the values of pre-service assessment training in the classroom and its' importance to design formative assessment methods (58.2%). However, 41.8% of science teachers agreed that the assessment training they took during pre-service programmes has little value to design formative assessment methods.

Almost all the interviewed results also showed the inadequacy of the pre-services assessment training for implementing formative assessment effectively. For them, the training mostly focuses on summative assessment; particularly on the preparation and format of paper and pencil true/false, matching, short answer, multiple choice and essay items. The classroom observation also confirmed this idea.

Regarding to item 5, 50.3% of science teachers took in-service assessment training, whereas 49.7% of them did not take such training. 54.9% of them reported that the in-service training has little value in the classroom to implement formative assessment. Most interviewees who took in-service assessment trainings were also unhappy on the relevance of the assessment training to implement formative assessment effectively. Generally, even if, 65.4% respondents reported as they have good understanding to implement formative assessment, large number of teachers (83.7%) on the other hand agreed as they need more training about formative assessment to implement it effectively (see Table 4.5 below).

Some teachers strongly advocated the “jug and mug” education system, who assumed the teachers should be a packed vessel and the students are as an empty vessel to be filled by the teacher. One example:

I took one classroom assessment course. The course has little contribution in my teaching, compared to others who did not take such courses. But in my opinion greater emphasis should be given for subject areas rather than for professional course (such as active learning methods, continuous assessment methods...). Such courses have very little value in the teaching learning process to achieve the desired goals. Thus, the

government should give high attention and time to subject area rather than for educational courses. (Teacher G, Nov 18, 2014)

Most teachers on the other hand, commented on the non-relevance of the assessment training they received during pre-service trainings to implement formative assessment. For example:

I took one classroom assessment course, but the course focuses more on the preparation and administration of summative assessment, such as how to prepare paper and pencil tests, and exams questions using table of specification, how to find reliability...it does not give emphasis on formative assessment tools. The course is more of theoretical, which is difficult to change into practice. (Teacher B, Nov 19, 2014)

Some teachers also strongly comment on the need of professional training to renew themselves, because such pedagogical knowledge is essential to understand the fundamental principles and the kind of difficulties students might have in the teaching learning process and to use appropriate teaching methods in their lesson than high level of qualification in a subject matter.

One example:

...regarding in-service trainings, before two years, we took training on active learning and continuous assessment methods for a short period of time, but still there is no training. I strongly comment on the need of trainings or workshops related to formative assessment and active learning methods which are very important to up-date me. (Teacher A, Nov 20, 2014)

One teachers to some extent commented on the relevance of the assessment training he received during pre-service training, but, it is as impractical within a given period. One example:

I took one assessment course during pre-service training. The course helps me to know how to prepare assessment questions in relation to the learning objectives and contents. However, it is very difficult to implement it within 40 minutes. (Teacher C, Nov 18, 2014)

**Table 4.5 Science Teachers' Professional Learning and its' relevance to implement formative assessment**

No	Items	Yes		No	
		F	%	F	%
1.	I took classroom assessment course during pre-service training.	128	83.7	25	16.3
2.	I took more than one course dedicated to classroom assessment.	110	71.9	43	28.1
3.	The assessment method I learn during pre-service training had little value in the classroom	30	19.6	123	80.4
4.	I received adequate training on how to design formative assessment methods during pre-service training.	89	58.2	64	41.8
5.	I took in-service training which focuses on formative assessment.	77	50.3	76	49.7
6.	I learned assessment method during in-service trainings that has little value in the classroom.	84	54.9	69	45.1
7.	I received no training on formative assessment implementation.	43	28.1	110	71.9
8.	I need more training about formative assessment to implement it effectively.	128	83.7	25	16.3
9.	I understand how to implement formative assessment in my daily instruction.	53	34.6	100	65.4

Generally, the analysis of this finding revealed that science teachers in the selected schools need to be provided with the appropriate formative assessment training continuously on how to integrate different assessment for learning strategies into their daily instruction.

**Research questions five:** What type of support schools provide for the effective practice of assessment for learning in science subjects?

Table 4.6 points to teachers' level of agreement on school supports provided for them to implement formative assessment in science education. Science teachers indicated that they agreed on:

- Importance of principals and supervisors support to implement formative assessment in science education (60.2%, M =3.33);
- Promotion of collaborative work among science teachers (56.9%, M =3.31);
- Encouragement of mutual observation and sharing of experience within their school (64.7%, M =3.52);
- Provision of opportunity for teachers to tryout and evaluates new ideas (45.8%, M =3.14); and
- Acknowledgement of teachers who implement formative assessment effectively (41.9%, M =3.07).

However, science teachers showed their disagreement on the provision of:

- Opportunity for short-term trainings on classroom assessment (2.75);
- Support for mutual observation and sharing of experience with other schools (2.82);
- Motives for teachers to carry out research with one or more colleagues to improve classroom practice (2.92);
- Time for teachers to meet on regular basis and provides opportunity for them to report their work (2.93);
- Resource funds to fulfill science laboratory equipment (2.93); and
- Models of quality practice to further teachers' professional development (2.99).



**Table 4.6 School support provided for science teachers to implement assessment for learning strategies in science subjects in the selected schools**

No	Items	Mean	Disagree		Neutral		Agree	
			F	%	f	%	f	%
1.	School principals and supervisors support is very essential to implement formative assessment in science education.	3.33	47	30.7	14	9.2	92	60.2
2.	Our school promotes collaborative work among science teachers.	3.31	40	26.1	26	17	87	56.9
3.	Our school encourages mutual observation and sharing of experience within our school.	3.52	31	20.3	23	15	99	64.7
4.	Our school encourages mutual observation and sharing of experience with other schools.	2.82	65	42.5	32	20.9	56	36.6
5.	Our school gives time for teachers to meet on regular basis and provides opportunity for them to report their work.	2.93	59	38.6	31	20.3	63	41.1
6.	Our school gives opportunity for teachers to tryout and evaluates new ideas.	3.14	45	29.4	38	24.8	70	45.8
7.	Our school provides models of quality practice to further teachers' professional development.	2.99	54	35.3	35	22.9	64	41.8
8.	Our school tries to find resource funds to fulfill science laboratory equipment.	2.93	60	39.2	33	21.6	60	39.2
9.	Our school provides opportunity for short-term trainings on classroom assessment.	2.75	69	45.1	37	24.2	47	30.7
10.	Our school motivates teachers to carry out research with one or more colleagues to improve classroom practice.	2.92	57	37.3	40	26.1	56	36.6
11.	Our school acknowledges teachers who implement formative assessment effectively.	3.07	51	33.4	38	24.8	64	41.9

During the interview, teachers were asked to mention the supports provided for them from school principals and supervisors to implement formative assessment in science education and its relevance. Two examples:

Our school principal and supervisor always encourage us to give marks for each assessment tasks. Moreover, before two years ago, they coordinated and gave us a training which focuses on active learning methods and continuous assessment for one day. But, the training lacks continuity and it is given only for a short period of time to implement it in the classroom. (Teacher H, Nov 18, 2014)

As much as I know, still there is no support provided related to assessment from supervisor and principal. However, our department provides peer to peer observation opportunities and to prepare quizzes and tests together, if we teach similar classes. (Teacher C, Nov 18, 2014)

Thus, appropriate support from school leaders is vital for science teachers to integrate formative assessment strategies in their daily lesson and to use evidences for next steps in learning to improve science learning values in the selected schools.

**Research question six:** Are there significant relationships between teachers' perception and school supports with teachers' overall practice of assessment for learning?

Multiple regressions were employed to assess how much variance in the practice of assessment for learning can be explained by teachers' perception on assessment for learning and school supports provided for them (by the two independent variables). The result of the multiple regressions or the adjusted R square in Table 4.7 indicates that 25 percent of the variance in teachers' practice of assessment for learning (total practice) is attributed to the variance of the combined of the two independent variables (i.e. teachers' perception on assessment for learning and school supports provided for them). Moreover, F-test was employed to determine whether the result is statistically significant or not. The result showed that the correlation between the dependent variable (total practice sub-scale) and the combined of the two independent variable (teachers' perception and school supports provided for teachers) was statistically significant at  $F(2,150) = 26.37, p = 0.000$ .

**Table 4.7 The multiple linear regressions model summary and analysis of variance**

Model			Sum of Squares	df	Mean Square	F	Sig.
Multiple R	0.51	Regression	17002.98	2	8501.49	26.37	.000
R Square	0.26	Residual	48353.18	150	322.36		
Adjusted R	0.25	Total	65356.17	152			
SEM	17.95						

Note: *SEM*= Standard Error of the Estimate; *df* = Degree of freedom; *F* = *F*-test

T-test was calculated to identify the contribution of each independent variable in the regression model. The result showed that from the two independent variables, school supports provided for teacher to implement formative assessment makes the largest significant contribution to explain the dependent variable, when the variance accounted by all other variables in the model is controlled for (beta = 0.322), teachers' perception of assessment for learning was also made statistically significant ( $p = 0.000$ ) contribution (beta = 0.316) to the regression model (see Table 4.8 below).

**Table 4.8 Significance of the predictor variable**

Variable	B	SE	Beta	t	Sig
School Support	.660	.150	.322	4.4	.000
Total perception sub-scale	.525	.122	.316	4.32	.000

Note: *B* = Beta for unstandardized coefficient; *SE* = Standard Error; *t* = *t*-test

**Research question seven:** How do the training and background characteristics of teachers affect their practice of assessment for learning in science subjects?

Six separate one way MANOVA were conducted to assess the effect of teachers' background characteristics (teaching experience, teaching load per week, teaching subjects, and class size) and assessment trainings (pre-service and in-service) on the combined dependent variable (i.e. PFS, CLE, ICE, SPESA, LCA, and AAE). If the multivariate test for the combined dependent variable is significant ( $p > 0.05$ ) across each independent variable, separate univariate F tests or ANOVAs were employed for the six dependent variables with an adjusted alpha level ( $p$

=0.0083). However, before employing multivariate analysis of variance, the researcher confirmed the preliminary assumptions. There were no violations in both cases. Accordingly, the result showed that the inter-correlation between the six dependent variables are approximately moderate that allow to proceed MANOVA (Pallant, 2001:225). The correlations between the dependent variables were interpreted along with Cohen, Manion, and Morrison (2007: 521) recommendation. These writers interpret a coefficient of correlation ( $r$ ) less than 0.1 is weak, 0.1 – 0.3 is modest, 0.3 – 0.5 is moderate, 0.5 – 0.8 is strong, and greater than or equal to 0.8 is very strong.

Thus, in this study, modest correlation were shown between planning of formative assessment (PFA) with learning contents assessed (LCA) ( $r = 0.255$ ), with application of assessment evidences (AAE) ( $r = 0.267$ ), and with interpretation and communication of evidences (ICE) ( $r = 0.281$ ), while other dependent variables correlate each other moderately (see Table 4.9). Similarly, the significant level indicates there is sufficient inter-correlation between the six dependent variables. The Bartlett's Test of Sphericity is also statistically significant (chi square = 238.881,  $p < 0.001$ ), indicating sufficient correlation between dependent variables to use MANOVA. Thus, it is possible to proceed MANOVA to assess the effect of teachers' training and background characteristics on the combined dependent variable. Moreover, other preliminary assumptions such as normality, outliers, equality of variance-covariance matrices, and error variance were checked across the independent variables.

**Table 4.9 Inter-correlation among the six dependent variables**

	<b>PFS</b>	<b>CLE</b>	<b>SPESA</b>	<b>LCA</b>	<b>AAE</b>	<b>ICE</b>
<b>PFS</b>	-					
<b>CLE</b>	.387**	-				
<b>SPESA</b>	.341**	.529**	-			
<b>LCA</b>	.255*	.407**	.384**	-		
<b>AAE</b>	.267*	.398**	.535**	.361**	-	
<b>ICE</b>	.281**	.519**	.463**	.369**	.444**	-

Note: \* $P = 0.001$ , \*\* $P = 0.000$

## **Teachers' practice of assessment for learning across pre-service assessment trainings**

To assess the effect of pre-services assessment trainings on the combined practice of assessment for learning, MANOVA was conducted. The dependent variables were the six factors of the practice sub-scale. Pre-service assessment training was one of the independent variable. Thus, MANOVA asks whether these six dependent variables are associated with teachers' pre-service assessment training or not. There are two groups of teachers who received the training (83.7%) and who did not (16.3%). From 83.7% of teachers, 71.9% of them took more than one assessment courses during pre-service training.

Before preceding the analysis, all of the preliminary assumptions were checked across group. The result revealed that there were no serious violations of assumption. Thus, it is possible to conduct the multivariate analysis of variance on the combined dependent variables across the two groups.

Table 4.10 illustrated that teachers who received pre-service assessment training showed slightly higher mean score in the interpretation and communication of evidences, whereas teachers who did not took such training reported higher mean score in the planning of formative assessment, collection of learning evidences, and application of assessment evidence factors. Generally, those teachers who did not receive pre-service assessment trainings revealed slight difference in their mean score of the overall practice (mean = 33.27, SD = 5.23) than teachers who took the training (mean = 32.99, SD = 5.04). Thus, the results of the descriptive statistics showed the non-relevance of pre-service assessment trainings to integrate formative assessment strategies into daily instruction for further learning.

**Table 4.10 Mean scores and standard deviation for the combined dependent variables as a function of pre-service training**

Dependent variables	With pre-service training		Without pre-service training	
	Mean	SD	Mean	SD
PFA	19.88	3.97	20.44	3.33
CLE	47.44	10.22	48	10.58
SPESA	52.86	7.21	52.96	8.68
ICE	21.36	2.21	21.08	2.53
LCA	15.77	2.44	15.76	2.65
AAE	40.68	4.17	41.4	3.59
<b>Overall practice</b>	<b>32.99</b>	<b>5.04</b>	<b>33.27</b>	<b>5.23</b>

The results of one-way MANOVA indicated that there was no significant difference between teachers who took pre-service assessment training and who did not, *Wilks' Lambda* = 0.989,  $F(6, 146) = 0.281$ ,  $p = 0.945$ , *partial eta squared* = 0.011, in their overall practice of assessment for learning. The eta squared also showed that 1.1% of the variance in the combined practice sub-scale was explained by pre-service assessment training (see Table 4.11). It was a very small effect size. Even, teachers who took more than one assessment course during pre-service training did not showed significant difference in their overall practice of assessment for learning, *Wilks' Lambda* = 0.986,  $F(6, 146) = 0.34$   $p = 0.912$ , *partial eta squared* = 0.014. Thus, the *Wilks' Lambda* multivariate test suggested that the combined dependent variable or the overall practice of assessment for learning was not affected by the pre-service assessment training teachers took.

**Table 4.11 MANOVA for the combined dependent variable as a function of pre-service assessment training**

Sources	Multivariate test	Value	F	Sig.	Partial Eta Squared
Pre-service training	Wilks' Lambda	.989	.281	.945	.011

The results of the separate univariate analysis of variance also indicated that teachers who took pre-service assessment trainings did not show any significant differences in the six factors of the

combined practice. The results of the interview and lesson observation were also validating it. Thus, pre-service assessment courses have not any significant contribution for teachers to fill their knowledge and skill gaps to integrate formative assessment with their daily instruction to improve learning.

### **Teachers’ practice of assessment for learning across in-service assessment trainings**

Here, MANOVA was performed to see the effect of in-service assessment training on the combined six dependent variables. All the preliminary assumptions were satisfactory across the two groups. Teachers who took in-service assessment training consist of 50.3% and 49.7% of teachers are not involved in such trainings.

Table 4.12 depicts the observed mean score and standard deviation of teachers who took in-services assessment training and who did not, in the six dependent variables. The result indicates that teachers who did not took in-service assessment trainings showed minor difference in their average score on the application of assessment evidence, learning contents assessed, interpretation and communication of evidences, and collection of learning evidences, whereas teachers who received the training revealed minor difference on their average score on the planning of formative assessment and provision of support to engage students actively. However, the observed mean score of the two groups in the overall practice of assessment for learning did not show perceptible differences.

**Table 4.12 Mean scores and standard deviation for the combined dependent variables as a function of in-service training**

Dependent variables	With in-service training		Without in-service training	
	Mean	SD	Mean	SD
PFA	20.12	3.91	19.82	3.85
CLE	47.38	9.65	47.68	10.89
SPESA	52.92	7.52	52.83	7.41
ICE	21.30	2.18	21.32	2.34
LCA	15.70	2.43	15.83	2.52
AAE	40.69	4.14	40.91	4.04
<b>Overall practice</b>	<b>33.02</b>	<b>4.97</b>	<b>33.07</b>	<b>5.18</b>

Moreover, the results of MANOVA revealed that there were no statistically significant differences in the combined practice sub-scale across teachers who took in-service assessment trainings and who did not take such training,  $Wilk's\ lambda = 0.995$ ,  $F(6, 146) = 0.116$ ,  $p = 0.994$ ,  $Partial\ eta\ squared = 0.005$  (see Table 4.13). The results showed that only 0.5 percent of the variance in the total practice of assessment for learning is explained by in-service assessment trainings, which is a very small effect.

**Table 4.13 Multivariate analysis of variance for the combined practice as a function of in-service training**

Sources	Multivariate test	Value	F	Sig.	Partial Eta Squared
In-service training	Wilks' Lambda	.995	.116	.994	.005

Besides, the result of the separate univariate analysis disclosed that there were no statistically significant differences across teachers who took in-service assessment training and who did not receive such trainings in the six dependent variables. Generally, the results of the descriptive and inferential statistics showed that teachers who took in-service assessment trainings did not perform better in the overall practice of assessment for learning as well as in the separate factors, such as in their planning of formative assessment, collection of learning evidences, interpretation and communication of evidences, learning contents assessed, support provided to engage students actively in the lesson and application of assessment evidences.

The interview session also confirmed this result. Most teachers commented the non-relevance of the in-service assessment training they received to integrate different assessment for learning strategies into their daily instruction. As to them, the trainings were given for a short period of time even an hour workshops and the trainings include many issues. It does not emphasize on one concept (for example on assessment). Thus, the in-service assessment trainings given for second cycle secondary school science teachers in the selected school have no significant effect on the implementation of assessment for learning.



### **Teachers' practice of assessment for learning across different background characteristics**

Four separate MANOVA were conducted to investigate significant differences across teachers who have different background characteristics in the combined practice sub-scale (i.e. PFS, CLE, ICE, SPESA, LCA, and AAE). The dependent variables were the six factors in the practice sub-scale. Teaching experience, class size, workload per week and subject taught were the independent variables. The results of the preliminary assumptions of normality, linearity, homogeneity of variance-covariance and error variance were not violated for the six dependent variables across the independent variables.

### **Teachers' practice of assessment for learning across teaching experience**

Multivariate analysis of variance (MANOVA) was conducted on the combined dependent variables. Here, the independent variable was teachers' teaching experience. Based on their experience teachers were grouped into three and distributed as teachers with an experience of below or equal to 7 years (39.22%), between 8 to 15 years (32.03%), and above or equal to 16 years of teaching experience (28.8%). All of the teachers were included in the analysis, because there were no outliers in the distribution.

The results of the descriptive statistics (see Table 4.14) indicates that teachers with an experience of below or equal to 7 years and above or equal to 16 years showed better observed mean score in the collection of learning evidences, support provided to engage students actively, learning contents assessed, and application of assessment evidences than teachers with an experience of 8-15 years. On the other hand, teachers with 8-15 years of teaching experience scored higher mean in the interpretation and communication of evidences. Similarly, teachers with an experience above or equal to 16 years depict higher observed mean in their planning of formative assessment than other teachers in the selected schools.

**Table 4.14 Mean scores and standard deviation for a combined dependent variable as a function of teaching experience**

Dependent variables	Teaching experiences					
	Below or equal to 7 years		Between 8-15 years		Above or equal to 16 years	
	Mean	SD	Mean	SD	Mean	SD
PFA	19.93	3.98	19.22	3.55	20.84	3.97
CLE	49.50	10.80	43.16	9.85	49.70	8.41
SPESA	53.58	6.56	51.45	7.75	53.5	8.14
ICE	21.23	2.51	21.96	2.26	20.68	1.67
LCA	16.15	2.42	15.06	2.52	16.02	2.36
AAE	41.4	4.48	39.67	3.96	41.23	3.42
<b>Overall practice</b>	<b>33.63</b>	<b>5.13</b>	<b>31.75</b>	<b>4.98</b>	<b>33.66</b>	<b>4.66</b>

When we compare the overall practices of assessment for learning among the three groups of teachers, teachers with an experience of below or equal to 7 years (mean = 33.63, SD = 5.13) and teachers who have an experience of greater than or equal to 16 years (mean = 33.66, SD = 4.66) reported higher observed mean score than teachers with an experience between 8-15 years (mean = 31.75, SD = 4.98) (see Table 4.14).

Moreover, the Wilks' Lambda criterion revealed that the combined dependent variable was significantly affected by teachers' teaching experience,  $F(12, 290) = 1.88, p = 0.037$ , Wilks'  $\Lambda = 0.861$ ,  $\eta^2 = 0.072$  (see Table 4.15). The results of partial eta squared showed that 7.2% of the variance in the combined practice sub-scale was explained by teachers' teaching experiences. Even if, the effect size is relatively small, the multivariate test suggested that the overall practice of integrating formative assessment strategies into daily instruction to improve learning was statistically affected by teachers' teaching experience.

**Table 4.15 Multivariate analysis of variance for the combined dependent variable by teaching experience**

Sources	Multivariate test	Value	F	Sig.	Partial Eta Squared
Teaching experiences	Wilks' Lambda	.861	1.88	.037	.072

Following this significant multivariate test results, univariate ANOVAs were conducted separately to investigate the impact of teaching experience on the six dependent variables. Here, Bonferroni adjustment was made to reduce the rate of Type I error due to the six separate analyses. Thus, the significant level was tested at alpha level of 0.0083 (0.05/6). As indicated in Table 4.16, the univariate tests revealed that teaching experience significantly affects the collection of learning evidences than other factors in the combined practice sub-scale,  $F(2, 150) = 7.072, p = 0.001, Eta = 0.086$ . Even if, the partial eta squared showed relatively a small effect size, the result indicates that teaching experience has its own contribution on the use of different formative assessment strategies into their daily instruction to collect learning evidences for further learning. Thus, what, how, and why teachers' use different formative assessment strategies were influenced by their teaching experiences. However, no statistically significant teaching experience effects were observed on the five factors of the combined practice sub-scale (see Table 4.16).

**Table 4.16 Univariate analysis of variance for the six dependent variables across teaching experience**

Source	Dependent variables	SS	df	MS	F	Sig.	PES
Teaching experience	PFS	60.68	2	30.34	2.052	0.132	0.027
	CLE	1375.27	2	687.63	7.072	0.001*	0.086
	SPESA	146.93	2	73.47	1.333	0.267	0.017
	ICE	38.37	2	19.18	3.908	0.022	0.05
	LCA	36.09	2	18.04	3.043	0.051	0.039
	AAE	91.82	2	45.91	2.826	0.062	0.036

*Note: \* $p < 0.008$  (less than the adjusted alpha level); SS = Sum of squares; MS = Mean squares; PES = partial Eta squared*

Tukey post hoc test of multiple comparisons indicated that teachers with below or equal to 7 and above or equal to 16 years of teaching experience ( $M = 49.50, SD = 10.80$  and  $M = 49.70, SD = 8.41$ ) respectively, on average reported higher level of collection of learning evidence than teachers with between 8 to 15 years of teaching experience ( $M = 43.16, SD = 9.85$ ). The mean differences were significant at  $p = 0.003$  and  $0.005$  respectively (see Table 4.17). However, there were no statistically significant differences between teachers with below or equal to 7 years and teachers with above or equal to 16 years of teaching experience in their collection of learning

evidences even if the mean score of teachers with teaching experience above or equal to 16 years showed slight difference (see Table 4.17).

The result of this section seems to indicate that not only the effect of the length of teaching experience but also what teachers experienced might have some impact on the practice of assessment for learning in the selected schools.

**Table 4.17 Multiple comparison post hoc test (Tukey) for collection of learning evidences**

(I) Teaching experience	(J) Teaching experience	Mean Difference (I-J)	Std. Error	Sig.
below or equal to 7 years	8-15 years	6.34*	1.90	.003*
	above or equal to 16 years	-0.20	1.96	.994
8-15 years	below or equal to 7 years	-6.34*	1.90	.003*
	above or equal to 16 years	-6.54*	2.05	.005*
above or equal to 16 years	below or equal to 7 years	0.20	1.96	.994
	8-15 years	6.54*	2.05	.005*

Note: \* $p < 0.008$  (less than the adjusted alpha level)

### Teachers' practice of assessment for learning across different class size

To investigate the effect of class size on the combined practice of assessment for learning as well as on the individual dependent variable, multivariate analysis of variance was used. Thus, in this case the dependent variable was the combined practice, whereas the independent variable was class size. Teachers were clustered into three groups based on the average number of students accommodated in their classrooms. These are: teachers who teach class size of 35-50 (29.4%), 51-65 (61.4%) and above 65 (9.2%). Accordingly, the results of the descriptive statistics and multivariate analysis were presented below. The preliminary assumptions were not seriously violated among the three groups.

The descriptive statistics indicates that teachers who teach more than 65 students on average in one class scored a higher observed mean in their planning of formative assessment strategies and collection of learning evidences than teachers who teach average class size of between 35-50 and

51-65. Moreover, teachers who teach average class size of 51-65 reported higher average score in the learning contents assessed and application of assessment evidences than the other two groups of teachers. On the other hand, teachers who teach average class size of 35-50 showed a slight difference in the provision of support to engage students' actively in the lesson than others. The results of the observed mean score for the six dependent variables were unpredictable across teachers who teach different class size (see Table 4.18 below). Similarly, teachers who teach an average class size of more than 65 and 51-65 scored higher mean in their overall practice of assessment for learning than teachers who teach relatively small number of students in one class. Generally, the descriptive statistics revealed that the average number of students in one class has no consequence on the practice of assessment for learning in the selected schools.

**Table 4.18 Mean scores and standard deviation for the combined dependent variables as a function of class size**

Dependent variables	Class size					
	35-50		51-65		More than 65	
	Mean	SD	Mean	SD	Mean	SD
PFA	20.11	4.02	19.83	3.73	20.43	4.52
CLE	46.16	10.64	47.89	10.58	49.5	5.59
SPESA	54.40	7.10	52.06	7.61	53.43	7.04
ICE	21.22	2.13	21.33	2.34	21.43	2.24
LCA	15.40	2.44	16.02	2.53	15.21	1.97
AAE	40.60	4.02	41.04	4.12	39.79	4.12
<b>Overall practice</b>	<b>32.98</b>	<b>5.06</b>	<b>33.03</b>	<b>5.15</b>	<b>33.30</b>	<b>4.25</b>

The result of the Wilks' Lambda test also showed that there were no significant multivariate differences in the combined practice of assessment for learning among teachers who teach different number of students in one class,  $Wilks' \lambda = 0.886$ ,  $F(12, 290) = 1.513$ ,  $p = 0.119$ ,  $partial \eta^2 = 0.059$ . According to Table 4.19, only 5.9 % of the variance in the practice of assessment for learning was explained by class size (see Table 4.19).

**Table 4.19 Multivariate analysis of variance for the combined practice by class size**

Sources	Multivariate test	Value	F	Sig.	Partial Eta Squared
Class size	Wilks' Lambda	.886	1.513	.119	.059

The result of the separate univariate test also did not show statistically significant differences in the planning of formative assessment, collection of learning evidence, interpretation and communication of evidence, support provided to engage students actively, learning contents assessed and application of learning evidences across teachers who teach different class size. Thus, this result showed that the number of students in one class does not affect the practice of assessment for learning in the selected schools.

The classroom observation result also confirmed this output. The researcher observed eight classes which have different number of students, yet there were no difference observed among teachers in their interaction with students, questioning strategies, provision of feedback, sharing of learning objectives and assessment criteria and other strategies they employed. However, in the interview sessions almost all teachers agreed on: average number of students in one class as a major factor that hinders the implementation of assessment for learning strategies in science education in the selected schools. Therefore, this contradictory idea indicates teachers' fear of class size as a problem but not the practical or actual problem they faced during the lesson.

#### **Teachers' practice of assessment for learning across different teaching load per week**

MANOVA was performed on the six combined dependent variable to assess the multivariate effect of teachers' teaching load per week. The dependent variables in this analysis were the six factors of the practice sub-scale. Teaching load per week was the independent variable. Based on their teaching load, teachers were grouped into three (teaching load below 10 (21.6%), 11-15 (34.6%), and above 16 (43.8%). All of the preliminary assumptions were met among the three groups to precede the analysis. Thus, the results of descriptive statistics and multivariate analysis were presented below.

The descriptive statistics on Table 4.20 illustrates that teachers who teach below 10 periods per week reported a slightly higher mean score on the planning of formative assessment and interpretation and communication of evidences than the other two groups. Similarly, those teachers reported a higher mean score in the provision of support to engage their students' actively than teachers who have a teaching load 11-15 periods, but scored less than teachers who have a teaching load above or equal to 16 periods per week. However, teachers who have more than or equal to 16 periods per week perform better mean score in the collection of learning evidences, provision of support to engage students actively, and application of assessment evidences than other teachers who have low teaching load per week.

In the overall practice of assessment for learning, teachers who have a teaching load above or equal to 16 periods per week scored better mean than the other two groups. On the other hand, teachers who have teaching load below or equal to 10 periods per week scored higher observed mean in the overall practice than teachers who have a teaching load between 11 to 15 periods per week. Thus, the result seems to reveal that teaching load per week has no contribution in teachers' practice of assessment for learning in the selected schools.

**Table 4.20 Mean scores and standard deviation for the combined dependent variables by teaching load per week**

Dependent variables	Teaching load per week					
	Below or equal to 10		11-15		Above or equal to 16	
	Mean	SD	Mean	SD	Mean	SD
PFA	20.03	3.95	19.98	3.51	19.93	4.15
CLE	47.24	7.62	47.04	11.46	48.06	10.47
SPESA	53.24	6.83	51.23	8.32	54	6.84
ICE	21.55	2.28	21.36	1.99	21.15	2.46
LCA	15.73	2.4	15.64	2.58	15.88	2.44
AAE	40.58	3.44	40.49	4.18	41.15	4.31
<b>Overall practice</b>	<b>33.06</b>	<b>4.42</b>	<b>32.65</b>	<b>5.34</b>	<b>33.36</b>	<b>5.11</b>

Moreover, the results of multivariate test indicated that there were no statistically significant difference in the practice of assessment for learning across teachers who have different teaching loads or periods per week, *Wilks' lambda* = 0.958, *F* (12, 290) = .518, *p* = .902, *partial eta*

*squared* = 0.021 (see Table 4.21 below). The proportion of variance in the practice of assessment for learning that was explained by teaching load per week was 2.1%, which is a very small effect size. Therefore, teachers' teaching load per week has no any significant effect on the combined dependent variable or on teachers' overall practice of assessment for learning in the selected schools.

**Table 4.21 Multivariate analysis of variance for the combined dependent variable across teaching load per week**

Sources	Multivariate test	Value	F	Sig.	Partial Eta Squared
Teaching load per week	Wilks' Lambda	.0958	0.518	.902	.021

The univariate analyses were also computed across the six factor or dependent variables, but the result did not specify significance differences between teachers who teach different workload per week. The interview result also verified it; all of the interviewed teachers did not mention their workload per week as a factor that hinders their practice of integrating assessment in their instruction. Thus, it seems that teaching load per week did not significantly explain the variance in the practice of assessment for learning in the selected schools. Generally, from the result of multivariate test, we can say that teachers teaching load per week has no impression to integrate assessment for learning strategies into their daily instruction in the selected schools.

**Teachers' practice of assessment for learning across different teaching subjects**

One way-MANOVA was also used to assess significant differences among teachers who teach different subjects on the combined practice of assessment for learning. Here, the independent variable was subjects taught (Biology, Chemistry, Mathematics, and physics). Table 4.22 indicates the observed mean score and standard deviation of the six dependent variables for teachers who teach different subjects. The result demonstrated that mathematics teachers on average reported a higher level of mean score in their planning of formative assessment, collection of learning evidences, provision of support provided to engage students actively, and application of assessment evidences than biology, chemistry, and physics teachers. Also, on



average, physics teachers reported better observed mean score in the planning of formative assessment than other teachers.

When we compare the four subject teachers in their overall practice of assessment for learning, mathematics teachers reported higher observed mean score than the other three subject teachers. Similarly, biology teachers showed slight mean score differences in their overall practice than chemistry and physics teachers. Whereas, physics teachers reported low observed mean score in their overall practice than the other three subject teachers in the selected schools (see Table 4. 22).

The result of classroom observation also partially proved it; for example, in biology, chemistry, and physics classes teachers dominantly used lecture methods, and they sometimes ask simple oral questions, and those top students were involved to answer such questions. Specifically, in one chemistry lesson, only volunteer students were allowed to demonstrate their answers for requested questions on the blackboard. Whereas, in one mathematics class students sat in group (five students in one group), and the teacher gave class activities for them from text books. After 10 minutes of discussion the teacher rounded-up and checked their works. Besides, the teacher congratulated those groups who got correct answers and gave 3 minutes more for groups who did not come up with correct answer to do it again. Finally, the teacher randomly called students number from different groups to reflect what they did in groups. Thus, the descriptive result shows an inclination toward mathematics teachers in the overall practice of assessment for learning than other subject matter teachers in the selected schools.

**Table 4.22 Mean scores and standard deviation for the combined dependent variables by subject taught**

Dependent variables	Subject taught							
	Biology		Chemistry		Mathematics		Physics	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
PFA	19.86	4.23	19.42	4.06	20.16	3.35	20.55	3.89
CLE	47.07	9.92	46.74	10.32	50.7	10.22	44.41	9.89
SPESA	52.45	8.54	52.16	6.87	55.02	6.18	51.17	7.85
ICE	21.43	2.38	21.16	2.53	21.14	2.09	21.59	2.01
LCA	16.05	2.78	15.34	2.1	16.2	2.38	15.24	2.52
AAE	40.95	4.2	40.24	3.89	41.47	4.18	40.28	4.03
<b>Overall practice</b>	<b>32.97</b>	<b>5.34</b>	<b>32.51</b>	<b>4.96</b>	<b>34.12</b>	<b>4.73</b>	<b>32.01</b>	<b>5.03</b>

However, the multivariate test result showed that there was no statistically significant difference across different subject teachers in their combined practice of assessment for learning, *Wilks' Lambda* = 0.902,  $F(18, 290) = .84$ ,  $p = .652$ , *partial eta squared* = 0.034 (see Table 4.23 below). The partial eta squared indicated that 3.4% of the variance in the dependent variable was explained by the subject matter they teach. Thus, subject matter taught has no significant effect on the combined practice of integrating formative assessment strategies into their daily instruction.

**Table 4.23 Multivariate analysis of variance for the combined practice across subject taught**

Sources	Multivariate test	Value	F	Sig.	Partial Eta Squared
Subject taught	Wilks' Lambda	.0902	0.84	.652	.034

Univariate ANOVA was also computed across different subject teachers to assess their differences in the six factor loadings of the practice sub-scale, yet there were no statistically significance differences observed among them. The inferential statistics showed that the subject matter they teach has no significant role on teachers' practice of assessment for learning.

## **Summary**

In this section of chapter four, the data collected through questionnaire, semi-structured interview, and observation are presented to assess teachers' practice of assessment for learning and to identify possible hindering factors of their practice in the selected area. Moreover, the relationships between teachers' perception on formative assessment and school supports provided for the effective implementation of assessment for learning with their practice of assessment for learning and the effect of teachers' background characteristics and assessment training on the implementation of assessment for learning are presented in this section. The preliminary analysis of the data suggest that all of the teachers in the selected school experienced similar difficulties in putting formative assessment into practice effectively to improve students' learning. Some of the factors are: lack of available resources, awareness problems (lack of formative assessment related trainings), and lack of appropriate supports from school officials. The next section of this chapter discusses the key findings of this study in detail.

## **4.4 Discussion of Results**

The aims of this study were to examine science teachers' practices of assessment for learning in science subjects and to identify major factors that hinder such practices in second cycle secondary schools. This section of the chapter discusses major findings of the study regarding teachers' practice of assessment for learning, possible factors that hinder the practice of assessment for learning, teachers' perception on assessment for learning, school supports provided for science teachers to implement formative assessment, assessment trainings they received and its' relevance to integrate assessment for learning into daily instruction, the relationship between teachers' perception and school supports with their overall practice of assessment for learning and the impact of teachers' background characteristics and training on the combined practice of assessment for learning as well as on individual dependent variables.

#### **4.4.1 Teachers' practice of assessment for learning in science subjects**

Nowadays, formative assessment is viewed as a critical element in the teaching learning process to actively engage and encourage students in the process of constructing, modifying and deepening their knowledge (Keeley, 2008: 16-26; Harlen & Gardner, 2010:19-20) through developing positive self-esteem, confidence, motivation to learn, self-regulation skills and self-efficacy among students (Nicol & Macfarlane-Dick, 2006: 199-209; Irons, 2008: 27-50; Moss & Brookhart, 2009: 5-16; Harlen, 2010: 40-42). That is why ample research evidences also approved its decisive power, particularly in science education, to improve students' understanding of scientific ideas, reasoning ability, scientific competencies of gathering and using evidences, scientific attitudes, ability to communicate using appropriate scientific languages, and application of scientific problem solving skills in new situations (Bell & Cowie, 2002:80; Harlen, 2003: 7; Lauvas, 2008: 163; Keeley, 2008:19; Hasan & Ehsan, 2013:577).

Despite of this truth, the practice of integrating formative assessment with daily instruction to improve students' science learning standards in the selected school is very low. Mostly, science teachers implement formative assessment in its' traditional way. They use formative assessment to review learning over a period of time and to collect pieces of marks continuously rather than bearing in mind its learning value. However, research evidences showed that formative assessment is not just a tool that teachers use to evaluate learning at a particular time and to collect mark. Rather, it is a way of seeking, gathering, interpreting, communicating, and using evidences minute-by-minute throughout the instruction for the purpose of improving learning and it is a way of getting a complete picture of a student's progress.

Studies in the literature revealed that the practice of formative assessment is cyclic in nature and consists of the following phases: defining learning objectives and success criteria, collection of learning evidences, interpretation (identify the gap) of evidences, implementing interventions to close the gaps, and assessing the effectiveness of the interventions in closing the gaps (Harlen, 2003:20; Birenbaum et al., 2009:131).

Thus, sharing of learning objectives and success criteria, providing constructive or nonjudgmental written or oral feedback, engaging students in self-and peer assessment actively, and using effective questioning strategies and dialogue are important strategies to put the phases of formative assessment into practice effectively (Black & Wiliam, 1998: 17; ARG, 2002:7; Lee, 2006:44; CCEA, 2007:6-7; Hodgson & Pyle, 2010: 20; DCSF, 2010:6). However, the results of this study denoted that such basic features of formative assessment are far away from practice or awareness to be integrated with daily instruction to improve students' learning. Most of the teachers viewed assessment and instruction as two separate elements and practiced according to traditionalist perspectives. The results of this study related to the six factors of the practice of assessment for learning sub-scales were discussed below.

#### **4.4.1.1 Planning of formative assessment**

Planning different formative assessment strategies as an integral part of our lesson preparation is a prerequisite for any activity in the teaching-learning process to improve learning through active engagement of students in the learning to learn environment. It is the first step in formative assessment cycle. According to Assessment Reform Group (2002: 2), planning of assessment for learning strategies is one of the major principles to integrate assessment with daily instruction to improve students' learning. However, it was not the trepidation of most science teachers in the selected schools.

As mentioned on Table 4.2, half of the teachers reported that they frequently planned better questions and questioning strategies and learning objectives and assessment criteria in their lesson. While, considerable number of teachers reported that they occasionally plan to share learning objectives and assessment criteria; student-centered assessment methods such as self-assessment, peer assessment, peer to peer questions, self-reflection...; to provide feedback that will identifies next steps in learning; and to assess students' prerequisite knowledge. Yet, formative assessment is not an occasional practice; rather it is a continuous process throughout the instruction for the purpose of learning.

Similarly, the qualitative data revealed that planning of different assessment for learning strategies as an integral part of the lesson preparation was practiced hardly in the selected schools. Most interviewees articulated that they incorporate assessment methods such as classwork, assignments, home work, and oral questions in their lesson plan, but they did not incorporate other assessment for learning strategies for the sake of improving learning. Evidences showed that the endeavor of teaching, learning, and assessment closely entwined when lesson preparation includes various assessment for learning strategies. According to Harlen (2003: 22-23), careful thought and planning of various assessment for learning strategies is vital to gather learning evidences, to identify next steps, and to improve students' learning in science subjects.

Interviewed teachers also expressed that they did not plan the assessment tasks or questions corresponding to the learning objectives and real world tasks; rather, they simply set simple tasks or questions to evaluate students' level of understanding at the end of the instruction. Along with, Moss & Brookhart (2009: 102-104) specified that strategic higher order questions or tasks are not asked 'on the fly' rather than planned in relation to the learning objective. As to these writers, questions or assessment tasks that were carefully planned encourage classroom discussions, actively engage students in the learning how to learn skills, and increase teachers listening ability. Moreover, Wiliam (2008: 37) confirmed the effectiveness of carefully planned question to inspire higher order thinking among students and provide information for teachers to adjust instruction.

Besides, as observed in the lesson, most teachers frequently asked simple oral questions that came in their mind and gave homework from the textbooks at the end of their lesson. Thus, it seemed that the planning of different assessment for learning strategies as an integral part of daily instruction was very low in the selected schools. Moreover, the assessment tasks lacked alignment with the learning objectives and real world tasks in the selected schools. Consequently, failure in planning of different assessment for learning strategies as a part of lesson preparation would lead to failure in putting such assessment for learning strategies into practice to achieve higher order learning outcomes.

#### 4.4.1.2 Collection of learning evidences

One of the main phases in the integration of formative assessment into daily instruction was collecting evidences of students existing ideas, skills, knowledge, and ways of thinking in the context of an activity during the lesson. Thus, implementing different assessment for learning strategies was vital to partner teachers and students in the teaching-learning process to continuously and systematically gather evidences of learning for next steps in learning (Moss & Brookhart, 2009: 6). According to Swan (2005: 8), the assessment tasks that teachers' employed should enable students to demonstrate deep understanding of concepts and principles and give continuous evidence of their understanding and thinking for teachers and for themselves to close the gap between the existing understanding and the new understanding. Moreover, Dysthe (2008: 19-20), illustrated the importance of integrating different alternative assessment methods as a part of daily instruction to assess and gather evidences of higher-order science thinking, problem solving skills, and the application of knowledge into real-life contexts.

However, the result of this study indicates that most science teachers hardly integrated different assessment for learning strategies in their daily instruction to collect learning evidences for further learning in the selected schools. As illustrated on Table 4.2, large number of teachers rarely shared learning objectives and assessment criteria, and used self-assessment, peer assessment, self-reflection, quizzes, written feedback, peer feedback, practical work, and peer to peer questions.

The interview data also revealed that:

- Learning objectives and assessment criteria's were not shared, even some teachers expressed that they were not mindful for such activities;
- Self-and peer assessment, peer to peer questions, peer feedback, self-reflections were not implemented; some teachers used self-and peer assessment as a time saver for marking class works, home works...; and
- Written feedbacks that suggest future works were not provided on students' work.

Research evidences indicated that sharing of learning objectives and assessment criteria for students is vital to develop a learning to learn skills (Shepard, 2000: 11), to achieve learning objectives (Black & Wiliam, 1998: 23-24), to take responsibilities for their own learning (Fautley & Savage, 2008: 47), to be challenge seekers rather than challenge avoiders (Lee, 2006:46), to increase their' motivation to learn (Moss & Brookhart, 2009:24), to engage students actively in the lesson (Chappuis & Stiggins, 2002:5), to develop meta-cognitive skills (Keeley, 2008: 11-13), to foster positive classroom environment, and to build their self-esteem and confidence (CCEA, 2007:13).

However, none of the observed teachers clearly shared learning objectives and assessment criteria with students; except one teacher, who wrote learning objectives on the board and read it for students; but no one gave attention towards it. Regardless of this fact, different research works proved that even writing learning objectives and success criteria on the board or telling it for students is not sufficient to understanding it to achieve the desired learning objectives (Fautley & Savage, 2008:46-47; Moss & Brookhart, 2009:24-25; Havnes et al.,2012:22). Hence, students' need to have good understanding about learning objectives and assessment criteria to gather learning evidences, to interpret and identify next steps in in their learning and to be independent learners.

Moreover, research evidences revealed that students involvement in self-assessment, peer assessment, peer to peer questions, self-reflection, and peer feedback is crucial to make them more autonomous (Lee, 2006: 61; Nicol & Macfarlane-Dick, 2006: 205; Moss & Brookhart, 2009: 82), to help them to identify their own learning needs and next steps in their learning (Hodgson & Pyle, 2010: 4; CCEA, 2007:26), to increase their motivation and self-esteem (Lambert and Lines, 2001: 142), to develop higher level of commitment among them (Falchikov, 2005: 120-121; Lee, 2006: 61), to advance their social skills (MECY, 2008: 13), and as a result to increases the rate of their learning dramatically (Black & Wiliam, 1998a: 26; Nicol & Macfarlane-Dick, 2006: 205; Irons, 2008: 79).

However, the results of the qualitative data in this study confirmed that no one implemented such assessment strategies as a part of their instruction in the selected schools. Equally, Moss &



Brookhart (2009: 80) reported that most teachers use these skills as a “time-saver for grading tests”. This was also a problem for some teachers in this study. One interviewee expressed that he sometimes used self-and peer assessment for marking class works if he has no time to see the work of each student in the class. But, the use of self-and peer assessment is far from scoring and marking of students work as mentioned above.

Moreover, in most of the observed class:

- No student to student and teacher to student dialogue or sharing of ideas in the class;
- Check lists and anecdotal records were not used to observe students’ progress in the lesson;
- Most students simply listened and took notes during lecturing; and
- No time needed for involving students in assessing their own or others’ works or for other process involved in conducting assessment to improve learning.

Even if, teachers gave class works, home works, and individual exercise toward the end of the lesson, the tasks were given to the students to practice and consolidate what the teacher had just explained, not to collect evidences for further learning or to adjust next steps in learning. According to Harlen (2003: 22-23), such assessment data cannot entail next steps in learning and reflect the full range of learning objectives, rather it provides summative information. Moreover, such assessment methods did not give students evidences about their own thinking and growth during the lesson. As a result it did not give evidences to gain new perspectives on their potential to take actions in the future to learn science.

Evidences also showed that collection of learning evidences related to students’ skills, knowledge, and attitude relevant to the learning objectives is not the sole practice of teachers, but rather it is also the work of each student to achieve successful learning (Harlen, 2003:9) because, higher order learning is occurred through interaction between the student, the teacher, and the tasks in the social environment, “the learner both shapes and shaped by the community of practice” (James, 2006: 57). Therefore, teachers had to put into practice different assessment for learning strategies or tasks that engaged students actively and brought out the application of

higher order skills, knowledge, and thinking in a new situation to acquired continuous evidences for next steps in learning.

However, generally, the results of this section of the study showed that most of the teachers did not integrate a variety of assessment for learning methods as a part of their instruction to collect evidences for next steps in learning. Rather they used tests, home works, assignments, mid exams at the end of lessons, or chapters to collect marks. This might be due to lack of appropriate trainings on the learning value of different formative assessment strategies and schools intention on assessment.

#### **4.4.1.3 Interpretation and communication of assessment evidences**

Identifying learning gap and communicating this gap against learning objectives and assessment criteria is one of the main feature in formative assessment to improve students' learning, because it provides information about what a student can do or not do in relation to the intended learning objectives. The results in Table 4.2 revealed that more than half of teachers in the selected schools interpreted assessment evidences or identify learning gaps against learning objectives and assessment criteria. On the other hand, reasonable number of teachers reported that they regularly judged the results of students in the task in relation to the performance of other students within the class and with other classes. The results of the interview session also confirmed it. Most interviewed teachers identified students' level of understanding in the subject against the result of other students' and communicated the result with them and schools in such way.

However, assessment which focuses on competition among students rather than on personal improvement de-moralized low performing students to further next steps in their learning. Moreover, such type of norm-referenced interpretation of results does not clearly show students' progress and gap in learning against the intended learning objectives and it does not give clear direction for students what to do next. As mentioned in the literature review, evidence is elicited; performance is compared; and feedback is generated against learning objectives and success criteria to close the gap between students' current learning and desired objectives (CCEA, 2007: 8).

The results of this study also indicated that the types of feedbacks that science teachers mostly provided for their students were judgmental feedbacks. The following are teachers' responses during the interview session on the type of feedback they provided to students: scoring students' work and assigning marks out of the total, displaying the correct answer to help students to get the right answer to questions they got wrong, inform their achievement against other students, doing questions that students assumed difficult for them, and praising students who score high in the assessment task.

However, feedbacks in the form of marks and grades have little value on learning (Black & Wiliam, 2001: 8; Cheung et al., 2006: 216; Stobart, 2008:168; Fautley & Savage, 2008:42-45; MECY, 2008: 9; Heritage, 2010:13) because such type of feedback does not provide direction for next steps in learning (Kvale, 2007:65; Wiliam, 2008: 37; Stobart, 2008: 168). Rather, it creates competitive classroom environment (CCEA, 2007: 15) which undermine the self-esteem, confidence, and motivation of low achieving students to improve their learning in future (Lee, 2006: 56). On the contrary, effective feedback is the heart of formative assessment to develop self-regulated learners as mentioned in the literature part.

Besides, the qualitative data revealed that teachers' in the selected school mostly praised and appreciated those students who scored high in the assessment task. For instance, during observation most teachers said excellent, very good etc. those students who got the expected answer during oral questions. However, feedback which focuses on praise, reward, criticism and that lacked guidance has low impact on learning (TDU, 2009: 7; Bruno & Santos, 2010:112). Stobart (2008:160-168) considered such feedback which focuses on the self as a 'good murder'.

Therefore, feedback to be effective, it must help students to identify the gaps in their learning against learning objectives and assessment criteria and indicate next steps to fill the gaps in their learning (Lee, 2006:56; Heritage, 2010:4). Moreover, it should focus on the quality of students work not on the self. Black & Wiliam (2001:5) stated that "feedback to any pupil should be about the particular qualities of his or her work, with advice on what he or she can do to improve, and should avoid comparisons with other pupils".

Generally, Nicol & Macfarlane-Dick (2006:204), formulated guiding principle to make feedback to be powerful to improve students' learning. Accordingly, it should: be constructive (not judgmental), facilitate the development of reflection and self-assessment in learning, deliver high-quality information to students about their learning, encourage teacher and peer dialogue around learning, encourage positive motivational beliefs and self-esteem, provide opportunities to close the gap between current and desired performance, and give information to teachers that can be used to shape the teaching.

#### **4.4.1.4 Provision of support to engage students actively in the lesson**

It is one of the interventions in the assessment process to close the gaps in students' understanding. Currently, assessment for learning is seen as an active social process, particularly accomplished by the quality of teacher-student and student-student interaction in the learning context (Dythe, 2008:19; Heritage, 2010:8; Willis, 2011:402). This two-way exchange of information between teachers and students is the heart of formative assessment to improve students' learning. Thus, assessment tasks need to be collaborative, interactive and dynamic and students must be involved in the generation of problems and solutions (Shepard, 2000:12) because, according to Swan (2005: 4), higher order thinking and skills are developed from the social environment in which the individual live and actively practiced.

Furthermore, Keeley (2008:26-28) stated that encouraging students to use their prior knowledge and experience to predict and construct their explanation during scientific inquiry, creating a classroom culture of ideas not answers, developing student to students and teacher to student dialogue, encouraging students to take risks and listen other ideas carefully, using a variety of grouping arrangement, and encouraging continuous reflection are important elements to improve science learning through engaging students actively in the lesson. Thus, the role of the teachers in such situation is to reengineer or facilitate a learning to learn classroom environment to achieve the desired learning outcomes in science education.

The result of this study on Table 4.2 indicates that most science teachers regularly supported their students to engage them actively in their learning through the key elements of formative assessment. As to the response of most teachers, they allowed students to actively share ideas in group, to ask questions, to answer, to take risks, and to listen other ideas. Interviewees also expressed that there was one to five grouping in each class to allow discussion, to share ideas, and to do assignments in group and to help low achieving students to do more. However, most of them told that they did not use such mixed ability grouping in their instruction regularly for the purpose of learning because of students' negative and low expectation about themselves. Even if, when they gave assignments for the group, only few students dominantly performed the tasks.

The lesson observation also confirmed it. In most of the observed classes, teachers did not engage their students in the teaching-learning process, particularly in formative assessment. Except one mathematics lesson, most teachers dominantly used lecture method and students were not allowed to share ideas in group. Teachers invested very little effort to support their students to actively engage them in the lesson. It seemed that most teachers were very much concerned in covering the contents of the lesson on time, rather than engaging every student actively in the lesson to improve their understanding.

For this reason, most students passively listened, took notes, and read text books to follow teachers' explanation of contents, instead of actively engaging themselves in the lesson. However, the more students are involved in the collection of their learning evidences; interpretation of these evidences in terms of progress toward learning objective; decision of next steps in their learning and in the possible ways of their future learning, the more likely they will be able and want to achieve successful learning (Harlen, 2003: 19-25).

In addition, most teachers asked lower order oral questions which encourage memorization of facts and fast learners. When students did not provide the expected answers, teachers simply answered it and proceeded to the next explanation. In all of the observed schools, none of the teachers gave chance for students to discuss the questions raised in groups to allow every student to share ideas and to actively engage all of them in the lesson. However, as mentioned in the literature, questioning is one of the key strategies in formative assessment to engage every

student actively in the lesson through thinking, to assess students' prior knowledge (Wragg, 2001: 33), to communicate learning objectives (Moss & Brookhart, 2009: 29), and to develop a learning culture of open discussion or dialogue between students and students to teachers (Irons, 2008: 63).

Moreover, Rowe (1974) explained the importance of waiting time (thinking time) during questioning to increase the explanation of answers, students' confidence to respond in meaningful way in their own ability, and the number of students who knows the answer (cited in Black et al., 2003: 13; Fautley & Savage, 2008:38). Despite of this fact, most teachers in the observed class did not provide enough thinking time during questioning to engage every student through thinking and to get more explanation about the question. Similarly, large number of teachers on Table 4.2 reported that they regularly "encouraged students to answer questions quickly" which support fast learners and memorization of discreet facts and principles. These low levels of questions and the involvement of few students in answering such questions could keep the lesson going, but it was actually out of touch with the understanding of most of the class students.

Generally, even if 93.43% of teachers reported that they regularly or often provided support to engage students actively in the lesson, evidences from the qualitative data revealed that teachers had not developed the group work spirit with their students. The lesson was still all teachers driven. Students in the selected schools were not engaged actively to collect, interpret and use assessment evidences for their learning. Thus, the actual practice in the classroom showed that the provision of supports to engage students actively in the lesson was very low in the selected schools. Teachers became the only actors in the classroom, particularly in the assessment process, which had less value to achieve the desired learning objectives. Hence, there was a mismatch between what the teachers reported and what they actually practiced in the classroom. This indicated that teachers had the theoretical knowledge on the role of students that they had in learning, but they failed to put the theory into practice. This confirmed the need of further training on formative assessment strategies and the central role of students for any activity in the lesson.

#### 4.4.1.5 Learning contents assessed

It is widely acknowledged that science education should equip students' with problem solving skills, critical thinking, and scientific reasoning abilities, which enable them to be competent citizens of the country because the knowledge based economy in the 21<sup>st</sup> century needs critical thinkers and lifelong learners. On the contrary, the accumulation of factual knowledge, formulas and principles in science education is not vital to acquire the precondition for successful learning in the future. Consequently, assessment tasks need to be authentic (i.e. realistic, practical and challenging) to achieve these higher order learning outcomes in today's education system because authentic assessment tasks focus on students' problem solving and analytical skills, scientific reasoning abilities, ability to integrate and coordinate what they learn, creativity ability, and ability to work collaboratively (Murphy, 2006: 44-45).

Despite this intention, the results of this study indicates that science teachers mostly assessed lower levels of learning contents such as knowledge of facts, principles and formulas in the selected schools. According to Table 4.2 near to half of science teachers reported that they assessed learning difficulties, knowledge of scientific facts and higher order thinking (i.e. analysis, synthesis and evaluation). Relatively, large number of teachers on the other hand assessed problem solving skills and scientific reasoning ability of students in new situations occasionally and hardly ever.

However, assessment, which emphasis on the recall of isolated facts encourages shallow learning and memorization of scientific facts (Liu, Lee & Linn, 2011: 164). It does not deepen students' understanding of scientific knowledge, attitudes, and skills which are crucial in today's education (Harlen, 2003: 10) because, such factual scientific question do not allow students to fully understand the principle, the cause and effect relationship between variables and the application of scientific concepts in the real situations (Wragg, 2001: 63; Odom et al., 2011:2352).

The qualitative result also showed teachers' predominant focus on assessing memorization of simple learning outcomes. As said by most interviewed teachers, they did not design assessment

tasks in line with the learning objectives and real world tasks (see Section 4.4.1.1). Rather, they simply put straightforward oral or written questions to accommodate students who had different understanding levels (high, medium and low achievers) in science subjects and as a result to decrease attrition rate in the school. The results of the lesson observations also validated it. During lesson observation, most teachers asked factual questions such as simple facts, formulas, principles, which discourage divergent thinking's. As a result, the question-answer dialogue became superficial; one in which all thoughtful involvements suffers (Black & Wiliam, 2001:8).

Due to this fact, the assessment tasks used in the selected schools lacked alignment with the learning objective and real world tasks to be authentic assessment to assess higher order thinking in science subjects. It is true that failure in planning a variety of assessment for learning strategies as a part of their lesson would lead to failure in assessing higher order learning outcomes in the selected school. Thus, it seemed that the teaching learning process (instruction) in the selected schools mainly focused on content coverage and rote learning rather than conceptual understanding. Consequently "coverage without understanding is pointless and may even harmful, as pupils will learn that they cannot successful at learning" (Lee, 2006:52).

So far, good assessment practices gives emphasis to learning with understanding and application and on the entire conceptual network of the learners, not on memorization of facts and principles (Cizek, 1997: 6-7; Shepard, 2000:8). Hence, questions or assessment tasks should be challenging, cause deep thinking, provoke discussion, explore the full range of learning objectives, and buildup from previous learning (Lee, 2006: 50). Wragg (2001:32) also emphasized the importance of open ended higher order questions to deepen students understanding. In consequence, teachers need to use effective questioning techniques that keep all students engaged in higher order thinking and gauge the understanding of the whole class instead of just selected students (Wiliam, 2008:37).

Briefly, based on the results of this study, we can say that the role of the student in the selected schools became working hard to listened, took notes, and read text books correctly to memorize and repeat scientific facts, concepts, principles, and formulas to score high marks in science



subjects to pass to the next grade and the national examination. This reality was also manifested in higher education's and became challenging.

#### **4.4.1.6 Application of assessment evidences**

This factor concentrated on how the collected evidences from different assessment tasks were used to address students' needs or next steps in learning and adjust instruction. As mentioned in the review part, the process of using the collected assessment evidences to advance learning is the hallmark of formative assessment that makes it different from other modes assessment. However, the results of this study indicated that the application of the collected assessment evidences to improve students' learning was very low in all of the selected schools. As Table 4.2 illustrated, there was clear discrepancy in teachers' response regarding the application of assessment evidences in the learning process. On one hand, large number of science teachers agreed that they regularly applied the collected assessment evidences to modify their teaching strategies, identify gaps in students understanding, advise students how to fill the gaps in their learning, and plan what to teach next.

Though, it seemed that most teachers in the selected schools were appropriately using assessment evidences to improve students' science learning. In contrast, they also frequently used the collected assessment evidence to record marks for final result; approve students who scores high in the test, quiz, mid exam...; categorize students into different groups (high achiever, medium achiever and low achiever); and tell their achievement on a task against other students result. The fact was that, formative assessment was aimed for internal modification, but not to select, rank and group students accordingly. As indicated in the literature, even if teachers give suggestive comments which are operationally helpful for students' work and modify their teaching methodology accordingly, the provision of marks and classification of students into different groups have negative effect on it because students ignore comments when marks or other judgmental ways are provided together (Black & Wiliam, 1998:13).

This might be happened because of teachers' lack of appropriate knowledge and skills about formative assessment strategies and its purpose in learning. In the interview session, most

teachers also expressed that they used assessment evidences to classify students into different groups, to record marks for final result and to make aware students about their results against other students. The observation also confirmed it. However, as mentioned in the literature such activities undermine the self-esteem, motivation and confidence of low achieving students. It empowers low achieving students to attribute their failure towards lack of ability rather than effort to do more in the future.

In general, what was evident in this study was that the majority of science teachers did not effectively integrate formative assessment strategies into their daily instruction to improve students' learning. Particularly, the results of the qualitative data clearly indicated that science teachers relied predominantly on the traditional form of formative assessment. This was unfortunate because it narrowed the instruction and limited students' engagement in the teaching learning process to achieve high science education standards.

Hence, the results of this study were consistent with studies conducted in UK and some countries of Africa including Ethiopia. As to them, teachers formative assessment practice focused on paper and pencil assessment tasks which encouraged rote and superficial learning, over-emphasized on marks and grades, focused on sorting and ranking of students, focused on judgmental feedback which served social and managerial purposes, worked with an insufficient picture of students' learning needs, gave priorities to record marks rather than using them as an input for future learning, and gave less attention for students previous performance (Lambert & Lines, 2001:133-135; Black & Wiliam, 2001: 4-6; Black et al., 2003: 10-11; Ottevanger, Akker, & Feiter, 2007: 19; Kvale, 2007: 65; Black & Wiliam, 1998 as cited in Fautley & Savage, 2008:92; Fisseha, 2010: 103-105).

As a result, it was a dream to say assessment for learning was put into practice effectively in science subjects in the selected schools to assess higher order thinking; to increase students motivation for learning, to develop self-regulated learners, to engage every students actively in the assessment process, to collect learning evidences (i.e. evidence of students' knowledge, skill and attitude), to identify learning gaps and to fill such gaps using appropriate intervention mechanisms and as a result to improve their science learning.

#### **4.4.2 Factors that hinder the implementation of assessment for learning in science subjects**

Ample research evidences in chapter two, section 2.3.4.1 mentioned the possible factors that hinder the practice of assessment for learning in science subjects. Such as: teachers' view of science learning and assessment, the pressure of external high stake tests, fear of resources and time pressure, lack of professional development opportunities in pre-service and in-service trainings, teachers self-efficacy on their ability, lack of intensive support and students past experience on assessment. The result of this study is not different from these issues. Thus, the discussions were done based on the three factor loadings below.

##### **4.4.2.1 Resources related factors**

Table 4.3 indicates that resource related factors are the first ranked factors than awareness and support related factors to hinder the practice of assessment for learning in science subjects. Large number of teachers believed that lack of resources in the selected schools was the major possible factors to implement assessment for learning in science subjects. Particularly, class size, lack of available science resources (textbooks, lab rooms, lab equipment, demonstration sites...), lack of instructional materials and teaching load per week were mentioned as the major factors of their practice. Most interviewees also disclosed large number of students in one class and broad content coverage (particularly grade 11 Mathematics text book) were major possible factor of assessment for learning practices in these schools. The observation result also confirmed, lack of well-organized lab rooms and equipment and number of students in one class as a major factor for implementing assessment for learning in the selected schools to improve the learning standard of students in science subjects.

The results of this study is parallel with the work of other researchers mentioned in the literature, who together stated the organization and the availability of appropriate instructional material, well-equipped classrooms and laboratories, and reduced class size are vital to implement inquiry based science education in secondary schools within a safe environment. Firestone et al. (2004: 66) mentioned, the value of a wide variety of instructional materials, equipment and

technological tools to implement student centered learning strategies. Other many researchers also specified the impact of resources availability, layout and quality on formative assessment practice in science education (Bell & Cowie, 2002: 22; Lasonen, Kempainen, & Raheem, 2005: 40-41; Wiliam, 2006: 287; Bines & Woods, 2007: 10; MoE, 2010: 19-27).

For this reason, Ministry of Education in Ethiopia promoted different strategies to establish school clusters and resource centers at a certain area; as a result, teachers in the clustered schools were strongly supported and used reference materials, models and equipment jointly in order to enhance their competences and to achieve higher order learning outcomes in science education (MoE, 2002: 51-52). Despite this policy intention, lacks of available resources were the major fear of science teachers in the selected schools to integrate formative assessment as a part of their pedagogy to improve students' learning.

#### **4.4.2.2 Awareness related factors**

According to Table 4.3, large number of science teachers supposed “shortage of instructional time” as possible factors that affected the implementation of assessment for learning in science subjects. Most of the interviewed teachers also agreed on this idea. Most of them told that the time allotted for one period (i.e. 40 minutes) to cover large contents prevented them from implementing formative assessment strategies for the rationale of improving students' learning.

The result is comparable to the statement of Walport et al. (2010: 43), “it is impossible to achieve visible learning outcomes, if time and other resources are limited and that the consequence is teaching to the test”. Moreover, Harlen (2003: 38) added the importance of time to share their experiences, to discuss on the barriers and enablers of their formative assessment practice, and to observe other model classrooms. Besides, she stated fear of time pressure as one of the main factor for most teachers not to integrate formative assessment in their instruction.

On the other hand, the result of this study is contrasting to the work of (Liu & Carless, 2006:281; CCEA, 2007: 26; Irons, 2008: 27-32 and 80), which emphasis on the importance of implementing self-and peer assessment as an integral part of instruction to save time and resource and to reduced teachers' workload.

Teachers lack of knowledge and skill about formative assessment strategies and their long experience in using summative assessment were also considered as a factor in the selected school to implement assessment for learning. For instance, many respondents on Table 4.3 revealed that lack of professional training on formative assessment as a factor that hindered their practice of assessment for learning in science subjects. Evidences also showed the greatest impact of professional trainings which focused on formative assessment on teachers' teaching practice and students' achievement (William; 2006: 287; Moss & Brookhart, 2009: 9-13; Stiggins, 2009: 421; Harlen, 2010:127; Hill, 2011: 348) because such type of training is vital to identify and use a variety of assessment for learning strategies, collect learning evidences and interpret the evidences against the stated objectives, and to take actions to fill the gaps in learning. Thus, the practice of formative assessment depends largely on teacher's professional knowledge, skill and experience (Bell & Cowie, 2001: 548; Harlen, 2010: 101) to achieve higher learning standards (see details on Section 4.4.5).

Similarly, on the same table teachers' negative perceptions on formative assessment was also mentioned as a factor that hindered the practice of assessment for learning in the selected schools. That was why; most of the observed teachers dominantly used the traditional learning and assessment systems. Literature also showed the impact of teachers' negative perception about formative assessment to implement it effectively to improve students' learning (Shepard, 2000: 5; Moss & Brookhart, 2009: 13-15), because, as said by Stobart (2008:144), it leads teachers into dilemmas in assessment for learning practices (see Section 4.4.3).

Another awareness factor that hinders the implementation of assessment for learning was students' negative perception (see Table 4.3). According to interviewees, most students expected them as good listener and note taker and teachers as the only source of knowledge in the lesson to be successful in their learning. However, as to the current learning paradigm, students are not passive receiver of information or empty vessels to be filled by the teacher rather they are an active agents in the teaching learning process (Harlen, 2003: 34; Black et al., 2003:14).

Moreover, as to the response of most teachers, students' viewed formative assessment as a pre-packed system employed for their privilege to collect a piece of marks rather than an active

strategy to improve their learning. For this reason, most students' gave more attention for assessment tasks that have marks. The observation results also verified it. Despite this fact, formative assessment was designed and practiced primarily to improve students' learning by actively engage them in the assessment process (Anderson, 1998:5; ARG, 1999: 2; Shepard, 2000: 6; Chappuis & Stiggins, 2002: 1-2; Singh et al., 2012:198). Assessment is a process of collaborative work in which information about learning flows between teacher and student; it is not something that teachers do to students (Lambert & Lines, 2001: 138; Wragg, 2001: 68). Rather, students need to be actively involved in the assessment process to achieve learning goals (Lambert & Lines, 2001: 141).

The results of this study also supports previous findings by Yigzaw (2013:1490), who stated that teachers' and students' perceptions of continuous assessment had a great impact on their assessment practices. Thus, school principals, supervisors and higher officials need to plan awareness creation programs not only for teachers but also for students about the purpose and nature of formative assessment strategies and their role on it to improve science learning standards.

#### **4.4.2.3 Support related factors**

The results of this study indicated that support related factors are the third ranked factors than the resource and awareness related factors to affect the practice of assessment for learning in the selected schools. Most teachers perceived that lack of support from colleagues, school principals, supervisors, and government mandates on assessment issue as a major factor that hindered the practice of assessment for learning in science subjects (see Table 4.3). The interview result also confirmed it.

As stated in the review part, the effective implementation of any educational changes in secondary school depends on the effectiveness of support from principals, supervisors and department heads. Specifically, Harlen (2003: 39) emphasized the importance of school principal and supervisor support for teachers to implement inquiry based formative assessment in science learning (see Section 4.4.4 for more detail).

### **4.4.3 Teachers' perception on assessment for learning**

As mentioned in the literature, the perceptions that every teacher holds are a result of their personal experiences and assumptions. Therefore, probing teachers' perception about assessment for learning is important in the sense that it provides an indication of how different assessment for learning strategies are being used or misused because teachers' perception on assessment affected their assessment practices (Moss & Brookhart, 2009:13-18). The results of this study showed that science teachers in the selected schools seem to held misconception about assessment for learning. This was evident by their response inconsistency to items related to the power of formative assessment on learning and the provision of constructive feedback to fill the gaps in students' learning and what they actually do in the class. The details were presented below.

#### **4.4.3.1 Teachers' perceptions on the power of formative assessment to improve learning**

According to Table 4.4 the majority of science teachers' reported as they have high level of positive perceptions on the power of assessment for learning to improve students' learning. Large number of teachers agreed on that formative assessment was an integral part of the teaching learning process. Moreover, they agreed on the power of formative assessment to identify students' problem in learning, to improve every students science achievement, to empower students to study from day one, to capture students attention and effort, to develop students positive self-esteem and confidence, to motivate students to learn, to ensure all students to have an experience of success and to make all of them competent learners, and to reduce the rate of repetition and dropout than summative assessment.

On the other hand on the same table most teachers perceived that "formative assessment follows the teaching learning process." This is synonymous with the inspiration of traditional assessment theorists, who viewed assessment and instruction as two separate elements. As indicated in the literature, teachers who strongly advocated the view of behavioural learning theories mostly developed negative perception about formative assessment. They mostly viewed formative assessment as a way of estimating how much learning objectives a student has acquired and

making judgment about the change in the observable behavior of the student (Bell & Cowie, 2002: 19; Dysthe, 2008: 19). However, formative assessment is multidirectional, contextualized, integrated, and social activity which is used to collect learning evidences and to identify the gaps in students' understanding to decide the next steps in learning (Bell & Cowie, 2002: 18-19; Harlen, 2003: 20-25; Heritage, 2010:8).

Most of interviewees similarly perceived formative assessment as tests, assignments, oral questions, class works, home works, mid exams... that teachers gave at the end of the lesson or a topic to evaluate students' level of understanding continuously, to classify them accordingly and to record their marks for final result. But, as different evidences indicated in the literature formative assessment is far from continuous or day to day evaluation of students' understanding.

According to Bell & Cowie (2002: 8), assessment is “considered as formative assessment if some action to improve learning during the learning was involved.” Therefore, what makes any particular assessment formative is not the specific assessment tool employed continuously but how the information gathered from the tool is used to improve learning and to adjust instructional strategies toward the learning goals. Similarly, Assessment Reform Group characterized formative assessment as “the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there” (ARG, 2002:2). However, if formative assessment tools were administered at the end of a lesson, the major purpose would be to evaluate students' understanding of subject matter and not necessarily to help students and teachers to get evidences for further learning. Such inconsistency in teachers' perception about the power of formative assessment to improve learning indicates their lack of theoretical knowledge and consistent practice of assessment for learning in science subjects.

#### **4.4.3.2 Teachers' perception on the active involvement of students**

Currently, learning is viewed as an active process in which the learner constructs and reconstructs meaning for him or herself. Due to this fact, most science teachers in this study also positively perceived the importance of active involvement of students in the lesson through



self-assessment, peer assessment, questioning, and effective feedback to improve their learning (see Table 4.4). Moreover, they believed on the provision of long waiting time during questioning to engage every student in thinking and answering higher order questions. Despite this fact, most of the observed teachers did not provide such chances for their students in the actual classroom practices to engage them actively in the lesson. They mentioned shortage of time and broad content coverage as the main factors.

As indicated in the literature, active involvement of students in the learning process is the heart of formative assessment to improve their science learning because, formative assessment is a collaborative process between teachers and students and among students (Bell & Cowie, 2002:116). Formative assessment provides different opportunities for learners to reflect their ideas, to evaluate their understanding and others and to understand valuable practices in the classroom community (Black & Wiliam, 2006: 13-15; Willis, 2011:402). Also teachers who participated in this study perceived formative assessment as the means to actively engage their students in the lesson.

However, in all of the observed schools, formative assessment strategies were not embedded as an integral part of the instruction (as an active teaching method) to improve students' learning, except simple oral question. Even, the questions were also not used for learning rather for evaluation purpose. Therefore, there is disparity what teachers perceived and what they actually practiced in the classroom. It indicated that teachers had gaps in their theoretical knowledge as well as in their actual practice. Teachers mentioned a number of reasons that might contribute to failure to grasp the appropriate knowledge and skills about formative assessment strategies and to translate the theoretical knowledge into practice (see Section 4.4.2 above).

#### **4.4.3.3 Teachers' perception on the provision of formative feedback**

As indicated on Table 4.4 almost all of science teachers believed on the importance of using a variety of assessment method to get comprehensive evidence about students learning. And formative assessment was the one that allowed them to use such variety of assessment methods. Similarly most teachers believed feedback as a dialogue between teachers and students to

identify gaps and to fill it and to improve learning. Yet, teachers in the selected schools failed to put such variety of assessment methods into practice to improve students' learning.

Even though they considered feedback as a key component of formative assessment to improve learning, they also negatively perceived it as detailed correction of students work. Moreover, they believed that giving scores or marks provided direction for students about their progress against learning objectives (see Table 4.4). Thus, feedbacks in the selected schools were considered as provision of detailed correct answers and marks. The qualitative data also confirmed it. However, when feedback is viewed as a transmission of simple facts and correct answers to students, learning improvement become absent (Underhill, 2006:166; Scholtz, 2007: 43) and teachers workload increase year by year as students number and class size increases (Bell & Cowie, 2002: 16-17; Nicol & Macfarlane-Dick, 2006: 200).

Generally, this study revealed that there was a clear detachment between how teachers' perceived formative assessment or assessment for learning strategies in science instruction and how they actually put it into practice in the selected schools. Therefore, "deep changes both in teachers' perception of their own role in relation to their students and in their classroom practice is important to implement classroom assessment changes effectively" Black et al., (2003:13). Similarly, Harlen (2003: 34-36) and CERI (2005:48), revealed the requirement of a change in teachers' perception and beliefs towards science education, students, students' learning, and assessment to implement formative assessment in science subjects successfully.

#### **4.4.4 Assessment trainings teachers received and its relevance to implement assessment for learning**

Teachers' professional quality is a pre-requisite for any activity in the teaching learning process to bring quality education. Regarding it, the Assessment Reform Group declared that teachers' professional development was one of the fundamental principles to put assessment for learning into practice effectively (ARG, 2002:2) because professional learning is one of the key elements to ensure meaningful and sustainable change in teachers' assessment practice (Moss & Brookhart, 2009: 13; Harlen, 2010: 100; Gardner, 2010: 8).

The results of this study indicated that most science teachers took more than one classroom assessment course during pre-service training, and nearly half of them agreed on its value in the classroom practice. Similarly, half of teachers took in-service training which focused on formative assessment, but it has little value in the classroom practice. On the other hand, large number of teachers reported two contradictory ideas. On one hand, most teachers reported that they took adequate pre-service assessment training which focused on formative assessment; on the other hand, they reported that they did not understand how to implement formative assessment and as a result they commented that they needed more training on formative assessment to implement it effectively (see Table 4.5).

Such inconsistency in teachers' response on the items related to their professional training showed that there was a clear gap in teachers' knowledge and skill to implement formative assessment effectively to improve students' learning. This might be because of lack of appropriate and deep professional trainings in the area. Due to this fact, most science teachers frequently used a series of tests which were simple to score and assess students' memorization of facts (Kubitskey et al., 2006: 108; Moss & Brookhart, 2009:9).

The results of the interview also indicated that many teachers were not satisfy with the nature of assessment training they received during pre-service and in-service trainings and they felt that the assessment training did not give them sufficient understandings and skills related to formative assessment strategies needed to integrate in their daily instruction to improve students' learning. They also felt that the assessment training they received during pre-services programs mostly focused on summative assessment. Moreover, the in-service trainings were given occasionally, it was days workshops which do not focus only in one issue (example on assessment), and it was not practical to help teachers to implement formative assessment effectively. These findings were consistent with the findings of other researchers. For instance, study by Moss & Brookhart (2009:9) indicated that most teachers implemented formative assessment strategies blindly and so did students because they were ordered to adopt a program prescribed and presented by outside experts in a one-shot workshop. For these reason, most

teachers used tests which are easy to score instead of using other assessment for learning strategies.

Therefore, to bring effective change in the assessment practices both pre-service and in-service assessment trainings should focus on formative assessment strategies and its learning values (Bell & Cowie, 2001: 548; Akerson, McDuffie, & Morrison, 2006: 160; Wiliam, 2006: 287). Moreover, the in-service trainings should be given continuously over a period of time rather than a day workshops (McWaters & Good, 2006: 53) and involve teachers actively and cooperatively in the development of new procedures or practices or trainings (Harlen, 2010: 101-105).

Due to this fact, currently, Ministry of Education in Ethiopia revised teachers' pre-service training curriculum and introduced various types of in-service trainings to improve the quality of teachers in implementing continuous assessment effectively (MoE, 2005: 11-14). However, adding a set of new techniques and principles to the existing practice or curriculum is not enough to demand effective assessment changes (Harlen, 2010: 102). Findings from USA by Khattri et al. (1998:144), as cited in (Yung, 2006:3), noted that "new approaches to student assessment alone are not sufficient to improve teaching and learning. Rather, the principles and ideas underpinning assessment reform must be clearly defined and understood at all levels".

Research showed that teachers who are engaged in appropriate professional learning can implement assessment for learning successfully (Mansell, James, & ARG, 2009: 19); involve their students actively in the assessment process (Chappuis & Stiggins, 2002: 43); share their experiences and create a community which feel comfortable (Kubitskey et al., 2006: 114-115); help to take responsibility for what happened in their classrooms; and not blame external circumstances to accept and implement the change (Swaffield, 2009: 9).

Thus, to effectively implement formative assessment in the selected schools, science teachers need to get appropriate training on how to plan formative assessment strategies in their daily instruction; share learning targets and assessment criteria's in understandable ways; gather evidences using different assessment strategies (observation, listening, questioning...); develop the skill of self-and peer assessment among students; interpret assessment evidences in relation

to learning goals or criteria; provide descriptive, timely, clear and task specific feedbacks; and show students where they are in relation to the goal and what they should do next to close the gap (Harlen, 2003: 39; Jones, 2005: 3; Moss & Brookhart, 2009:19; Stiggins, 2009: 421; Baird, 2011:344). Generally, according to Harlen (2010: 102), teachers' deep understanding is important for new practice to be successful unless "it is all too easy for new practices to fade away."

#### **4.4.5 School supports provided for teachers to implement assessment for learning**

Evidences showed that school officials' support was vital to implement any educational change effectively. According to Ottevanger, Akker, & Feiter (2007: 20), the provision of school support for teachers is a key element to the success of formative assessment. The findings of this study also confirmed it. As indicated on Table 4.5, most teachers believed on the importance of principals and supervisors support to put formative assessment into practice effectively, particularly in science education. Similarly, they reported that their school encouraged them to work together and to observe and share experience each other within the school.

In contrast, most teachers (see Table 4.5) believed as they were not well supported to:

- Observe and share experiences with other school;
- Meet on regular basis and provides opportunity for them to report their work;
- Tryout and evaluates new ideas;
- Observe models of quality practice to further their professional development;
- Get opportunity for short-term trainings on classroom assessment; and
- Carry out research with one or more colleagues to improve classroom practice.

However, such opportunities built and enhance teachers' capacity to implement assessment for learning effectively in science education (Black et al., 2004: 20-21; CERI, 2005: 5-7; Ottevanger, Akker, & Feiter, 2007: 20; Swaffield, 2011:438; Hill, 2011: 350). Research finding by Mansell, James, & ARG (2009: 22) as well proved that teachers who were supported by school leaders through collaborative classroom focused inquiry activities became more

successful to promote assessment for learning than other teachers. Moreover, these researchers added the positive impact of sharing of ideas and occasional meeting on the practice of assessment for learning. Despite this fact, the result of this study indicated that the extent of school principals and supervisors support for science teachers to implement assessment for learning in science education was inadequate.

As to most respondents, even schools did not acknowledge teachers who implemented formative assessment effectively; rather they encouraged rushed curriculum coverage, collection of pieces of marks, and teaching to the test. Mostly, schools give recognition for teachers who dominantly used lecture methods and students too. That is why many teachers found difficulties in bringing their practice in line with the purpose (learning value) of formative assessment, because teachers' different opportunities to build their capacity largely depends on school structures, cultures, and leadership (Swaffield, 2009:9). Hence, school principals and supervisors needed to have a good understanding about assessment for learning and built commitment to the vision amongst the teachers and students to achieve the desired objectives in science education. School leaders in the selected schools had also poor performance to find resource funds to fulfill science laboratory equipment (see Table 4.5). It was supported by the interview result.

Thus, principals and supervisors in the selected schools should support and provide opportunities for science teachers to upgrade their professional knowledge and skills in formative assessment through discussion with their peers, observing best practices, visiting other schools, assessing their own practice with action research, participating in conferences and workshops, reviewing other related research works (Harlen, 2003: 38; Akerson, McDuffie, & Morrison, 2006: 161; Stiggins, 2009: 421; Bennett & Gitomer, 2009: 55; Harlen, 2010: 101-116); and providing opportunities to get professional courses founded on formative assessment in pre-service and in-service trainings (ARG, 2003: 1; Harlen, 2003: 39; MoE, 2005: 13; ARG, 2005 cited in Montgomery, 2010: 64) Similarly, other researchers mentioned ongoing regular conversations with teachers about formative assessment as the most powerful means to develop professional competence (Moss & Brookhart, 2009: 13; Harlen & Gardner, 2010:26). Generally, such assessment change (assessment for learning) took time and required continuous attention from

policy makers, school leaders, researchers, and teachers themselves until it became a part of school culture.

#### **4.4.6 The relationship between teachers' perception and school supports with practice of assessment for learning**

The result of this study indicates there is statistical significance relationship between teachers' overall practice of assessment for learning and the combined independent variables (i.e. teachers' perception and school supports provided) at  $F(2,150) = 26.37, p = 0.000$ . As to the results of multiple linear regression 25% of the variance in teachers' practice of assessment for learning were accounted or explained by the two independent variables (see Table 4.7). Moreover, from the two independent variables school supports contribute the largest variance in the combined practice of assessment for learning than teachers' perception.

The results of this study is similar to Hill's (2011: 347-360) study in New Zealand Secondary Schools, which showed the relevance of school principals' and senior staffs support for teachers to effectively implement assessment for learning in mathematics and science education. Similarly, other researchers explained, the powerful effect of teachers' perception and belief about how students learn and knowledge is constructed on their classroom practice, particularly on the implementation of formative assessment in science education (Harlen, 2003: 34-36; Black et al., 2003:13-14; Firestone, Monfils, Schorr, Hicks, & Martinez, 2004: 67; Shepard, 2000:5; Yung, 2006: 3; Cheung et al., 2006: 210; Weaver, 2006: 380; Moss & Brookhart, 2009:13-18; Gulikers, Biemans, Wesselink, & Wel, 2013:122). Therefore, school principals and supervisors support and teachers' perceptions about assessment for learning have a significant contribution in the overall practice of assessment for learning in the selected schools.

#### **4.4.7 The impact of training and background characteristics of teachers' on the practice of assessment for learning**

The following discussion highlights how teachers background characteristics (including teaching experience, teaching load per week, class size, and subject taught) and the pre-service and in-

service assessment trainings teachers took significantly impacts the combined practice of assessment for learning in science subjects in the selected schools.

Two separate multivariate analyses of variances were employed to assess significance differences in teachers' overall practice of assessment for learning across teachers who received pre-service and in-service assessment trainings and who did not. However, the result of Wilks' Lambda test showed that there is no statistical significance difference between teachers who took pre-service assessment training and who do not took such type of training in their combined practice assessment for learning,  $F(6,146) = 0.281, p = 0.945$  (see Table 4.10). Pre-service assessment training accounted only 1.1% of the variance in the combined practice of assessment for learning, which was a very weak effect size. Thus, pre-service assessment training has no significant effect on the combined practice of assessment for learning as well as on the individual practice factors (variables).

Moreover, teachers who took in-services assessment training did not show statistically significance differences in their combined practice of assessment for learning than teachers who did not involve in such type of training,  $Wilks' Lambda = 0.995, F(6, 146) = 0.116, p = 0.994$  (see Table 4.12). Eta squared also indicated, a very small or insignificant effect. Only 0.5% of variance in the combined practice sub-scale was explained by the in-services assessment training. Thus, the in-service assessment training given in the selected schools has no effect on the implementation of assessment for learning.

The result of this study is opposed to other research findings mentioned in the literature. For example, evidences showed that professional learning which focused on formative assessment has a greatest positive impact on teachers' teaching practice and students' achievement (William; 2006: 287; Moss & Brookhart, 2009: 9-13; Stiggins, 2009: 421; Harlen, 2010:127; Hill, 2011: 348). According to these researchers, such professional learning is vital to identify and use a variety of assessment methods, to gather and interpret evidences in line with the learning objectives, and to take actions and fill gaps in future learning.



However, such differences between the current study and other studies may come because of the inadequacy of assessment trainings teachers received in the selected schools. As most interviewee indicated, most of the pre-service assessment training they took predominantly focuses on summative assessment, and the in-services trainings are too short and rarely given and mostly, it has little value in the classroom practice.

Moreover, four separate MANOVA were performed to assess the effect of teachers' background characteristics (teaching experience, teaching load per week, class size, and subject taught) on the combined dependent variable. Accordingly, the results of the Wilks' Lambda depicted the significant effect of teachers' teaching experience on the overall practice of assessment for learning,  $F(12, 290) = 1.88, p = 0.037, \text{partial eta squared} = 0.072$  (see Table 4.14). It explained 7.2% of the variance in the combined practice of assessment for learning.

Moreover, the results of the separate univariate analyses of variance indicated that this significant difference lays on the collection of learning evidences factor of the composite practice sub-scale,  $F(2, 150) = 7.072, p = 0.001, \text{partial eta squared} = 0.086$  (see Table 4.15). That is, teaching experience affects the use of different assessment for learning strategies to collect learning evidences for further learning. On the other hand, teaching experience has no significant effect on the planning of formative assessment, interpretation and communication of evidences, support provided to engage students actively, learning contents assessed, and application or use of assessment evidences.

To compare the mean score between the three groups in the collection of learning evidences Tukey post hoc multiple comparisons test was employed. The result indicated that teachers with below or equal to 7 and above or equal to 16 years of teaching experiences showed statistically significant mean score differences in their practice of collecting learning evidences for further learning than teachers with 8 to 15 years of teaching experience (see Table 4.16). But, there were not significant differences between teachers who have an experience below or equal to 7 years and those who have an experience above or equal to 16 years in the compound practice of assessment for learning sub-scale as well as in the collection of learning evidences.

The result of this section not only indicates the effect of experience, but it may have also some relation with the introduction of the new paradigm in the assessment system (i.e. assessment for learning). Since, the reform is recent in Ethiopia; teachers with below or equal to 7 years of teaching experience are entirely experienced with the new assessment reform (formative assessment) than summative one. These teachers have new experience than other teachers in the assessment aspect, and they were taught in the revised curriculum. Therefore, their opinion about formative assessment most probably was not affected by the traditional testing practices. Similarly, teachers with teaching experience above or equal to 16 years have long experience in the two forms of assessment, and as a result, their experience may help them to compare the effect of the two forms of assessment on students' learning. That is why they performed better than other teachers.

On the other hand, teachers with teaching experience between 8 to 15 years are permitted to use summative assessment dominantly for a short period of time and within a short period of time the emphasis shifts toward formative assessment. Consequently, their short period of experience in summative assessment does not allow them to see its impact on students' learning compared to that of formative assessment.

As indicated in the literature, teachers accept any change in education if they are the part of the change. Broadfoot (2008: 224) stated that to cure from disease, "the first step is to recognize that there is a problem. As such recognition develops, so the search for a cure will become more urgent and widespread." Thus, first teachers' should have experience to recognize the occurrences of the problem in their practice to accept the change and to be cured from the disease.

According to table 4.18, class size does not have statistically significant effects on the combined practice of assessment for learning in science subjects. The results of multivariate test indicates, there was no statistically significant differences in the combined practice of assessment for learning among teachers who teach different class size,  $Wilks' \Lambda = 0.886$ ,  $F(12, 290) = 1.513$ ,  $p = 0.119$ ,  $Partial \eta^2 = 0.059$ . Class size accounted only 5.9% of the variance

on the combined practice of assessment for learning. Moreover, the difference was not significant in all of the six factors in the practice sub-scale.

This result contradicts other research result which acknowledges the impact of large class size for teacher to engage in dialogue with students (Nicol & Macfarlane-Dick, 2006: 211). On the other hand, it supports researchers who advocate the importance of formative assessment strategies such as peer-assessment, self-assessment, peer to peer dialogue, provision of constructive feedback, and sharing of learning objectives and assessment criteria if there is large class sizes. Therefore, personalising learning for each of the student in the class is important, when there is large class size.

Even if, science teachers reported teaching load as one factor, the multivariate result showed that there was no statistically significant differences in the practice of assessment for learning across teachers who have different teaching loads per week, *Wilks' Lambda* = 0.958,  $F(12,290) = 0.518$ ,  $p = 0.902$  (see Table 4.20). Teaching load per week accounts only 2.1% of the variance in the combined practice of assessment for learning. Thus, multivariate test revealed that teaching load per week has no significant effect on the combined practice of assessment for learning in the selected schools. Moreover, there were no significant differences in the six factors or individual dependent variables across teachers who have different work load per week. The results of classroom observation and interview session also confirmed that.

The result goes along with other research works. According to Liu & Carless (2006: 281); CCEA (2007: 26); Irons (2008: 32 and 80) workload does not have impact to integrate formative assessment strategies into daily instruction. As to them, workload decreases, when self-assessment, peer assessment and constructive feedbacks implemented effectively.

Wilks' Lambda result illustrated that subjects taught (i.e. Mathematics, Chemistry, Biology and Physics) have no statistically significant impact on teachers' overall practice of assessment for learning,  $F(18, 290) = 0.84$ ,  $p = 0.652$  (see Table 4.22). The partial eta squared also showed a very small effect size (i.e. 3.4%). Therefore, teachers who teach different subjects have no significant difference in the composite practice of assessment for learning as well as in each of

the six factors of the practice sub-scale. The combined practices of formative assessment strategies as an integral part of daily instruction were similar across the four subject matter teachers.

However, to some extent, the results of the descriptive statistics, interview, and classroom observation revealed that mathematics teachers practiced assessment for learning a little better than other teachers in the selected school. Black and Wiliam (2004:17) have also got significant differences among teachers of science, mathematics, and language arts in their formative assessment approach. Yet, such similarities in the current study may be due to the subjects' are more related in many aspects and the teachers as well.

## **4.5 Summary**

In this chapter results and discussion of results regarding the basic research questions were done. Both the quantitative and qualitative data were presented and integrated consistently with the research questions and discussed in line with the literature review. Next chapter, chapter five presented the conclusions, summary, and recommendation of this study based on the results.

## **CHAPTER FIVE**

### **5. FINDINGS, CONCLUSIONS AND RECOMMENDATIONS**

#### **5.1 Introduction**

The aim of this research was to investigate teachers' practice of assessment for learning and identify major factors that hinder the practice of assessment for learning in science education at East Gojjam second cycle secondary schools, Amhara Regional State. The study has five chapters including the findings, conclusions and recommendations part and it is briefly summarized as follows.

Chapter 1 addressed the background of the study, knowledge gaps in the field, basic research questions, significance of the study, overview of the theoretical framework and literature review, definition of basic concepts, and general methodological frameworks.

In chapter 2, the underlying theories of assessment for learning i.e. the detailed theoretical view of behaviourism, cognitive constructivism, and social constructivism perspective on assessment and learning were presented. Moreover, literature review on the strategies of formative assessment, research evidences on the pedagogical power of formative assessment on students' learning, achievement, motivation, self-esteem, confidence, self-efficacy, and self-regulation, particularly in science education, good qualities of formative assessment, and factors that hinder the practice of assessment for learning are discussed in detail.

Furthermore, the research design; sampling and sampling technique; data collection instruments such as questionnaire, semi-structured interview, and observation; pilot testing; methodological norms that is the validity and reliability of instruments; and ethical considerations in the study are explained in detail in chapter 3.

In chapter 4, both the results of the quantitative and qualitative data were presented, integrated, and discussed in line with the basic research questions. Here, Chapter 5 illustrates the findings, conclusion and recommendation of this study. Moreover, the summary, contributions, and limitation of this study are presented in this chapter.

As mentioned above, to answer the research questions asked both quantitative and qualitative data were gathered using questionnaire, semi-structured interview, and observation and analyzed using appropriate methods. Thus, the conclusions, recommendations, and summary of this study were done accordingly.

## **5.2 Findings and conclusions**

Science education is believed to be a means to achieve the millennium development goals in Ethiopia to be a middle income country in 2025. And formative assessment is an engine to conduct inquiry based science education to make all students scientifically literate. As mentioned in the literature review, many researchers recognized the role of formative assessment in improving students learning, particularly in science education (see Chapter Two, section 2.3.2.1). However, to make the impact of formative assessment fruitful, science teachers need to understand the strategies and phases of formative assessment and its' learning value. Having this knowledge, teachers can integrate authentic assessment tasks into their daily instruction continuously to improve students' learning.

The findings and conclusion of this study related to the practice of assessment for learning and factors that hinder the implementation of assessment for learning in the selected schools are presented below.

**Research question one:** *To what extent do teachers integrate assessment in their instruction to improve students' science learning?*

Teachers in the selected schools predominantly focused on increasing students' academic achievement or on the collection of marks to decrease attrition rate rather than on improving

students' learning to learn skills. As a result of this intention, teachers mostly involved with continuously administering and scoring more simple assessment tasks such as assignment, quiz, test, homework, mid exam... throughout the school year as a means of evaluating students' learning and collecting marks. However, assessment, particularly formative assessment, is far from continuous scoring of different assessment tasks.

In general, the outcome of the first research question of this study indicated that teachers' overall practice of assessment for learning or the integration of different authentic assessment tasks with their daily instruction to improve students' science learning was very low in the selected second cycle secondary schools. Both the quantitative and qualitative results used below to conclude for each factor loading of the overall practice of assessment for learning in the selected schools.

### **Planning of formative assessment**

In the selected schools science teachers' give low attention to incorporate different assessment for learning strategies in their lesson preparation for the purpose of students' further learning. More than half of teachers reported that they regularly design better questions and questioning strategies and identify learning objectives and assessment criteria in their lesson plan (see Table 4.2). However, the qualitative data suggested that most teachers did not plan different formative assessment strategies as part of their lesson preparation. Most of the interviewed teachers expressed that they did not plan to share learning objectives and assessment criteria, did not include peer and self-assessment as part of their lesson, did not plan when and how to give constructive feedback for their students to improve students' science learning standards in the selected schools. It was also suggested by teachers' response inconsistency in the rest of the items in the questionnaire. Even, many teachers did not plan and conduct classroom questions in the way that might help students to learn.

Currently, new lesson preparation format was introduced regionally to improve learning standards through assessment, but most teachers incorporate tests, assignments, classwork, homework, and mid exams in their plan to collect marks and to consolidate what they taught. Thus, planning of different formative assessment strategies as an integral part of the lesson

preparation to improve students' learning is not a matter of most science teachers and even for school officials in the selected schools, they all follow the traditional culture of lesson planning. As a result, the teaching-learning process becomes superficial which focuses on content coverage.

### **Collection of learning evidences in the selected schools**

The outcome of this factor loading revealed that teachers did not effectively implement different assessment for learning strategies to collect learning evidences for further learning. Rather they regularly used tests, class works, home works, and assignments towards the end of the lesson to collect marks for the purpose of summative assessment.

Both the quantitative and qualitative data suggested that:

- Learning objectives and assessment criteria's were not clearly shared for students;
- Self-and peer assessment, practical work, peer to peer questions, peer feedback, students reflection of ideas on the lesson learnt and self-reflection using drawings, and concept mapping were not implemented;
- Divergent questions were not forwarded both from teachers and students during the lesson;
- Feedbacks were not delivered in a descriptive or constructive manner; and
- Student to student dialogue, teacher to students dialogue were not implemented.

Thus, students' active involvement became suffered in the teaching-learning process to collect evidence for their learning to improve their understanding in future. Moreover, it has negative implication to assess the more holistic picture of students' performance in the lesson.

### **Interpretation and communication of assessment evidences**

The identification and filling of students' learning gaps against the learning objectives in the selected schools and communication of assessment evidences too were very low. There is



contradictory teachers' response on items related to the interpretation and communication of evidences in the quantitative data. Half of teachers in the selected school reported that they regularly identify learning gap against learning objectives and assessment criteria. In contrast, they also reported that they regularly compared students result with other students within the same class and other class students to identify learning gap and communicate it with students accordingly.

The qualitative data also approved that the collected evidences through tests, home works, class works, assignments, and mid exams were interpreted against the performance of other students within the same class and other class of students. Teachers rarely identify learning gaps against the learning objectives, students' pervious performance and assessment standards to decide next steps in learning. Moreover, most teachers in the selected school communicate the evidences with students using marks, correct and wrong answers which have no value to improve students' learning. Rather, such type of judgmental feedback undermine the self-esteem, confidence, effort and working ability of low achieving students in the selected schools. Yet, as stated in the literature review (see Chapter Two, Section 2.3.2.2), ample research evidences agreed that formative assessment is more powerful to increase learning gains, particularly for low achieving students.

### **Supports provided to engage students actively**

Active engagement of students in formative assessment is the key element to improve their learning. In the new learning paradigm, students are at the center for any activity in the lesson to achieve the competencies required in the 21th century. Thus, students need to be involved actively in the lesson through peer assessment, self-assessment, reflection of ideas, provision of descriptive feedback, questioning, answering, dialogue, identifying gaps, and planning of next steps in learning, because such activities help students to develop learning to learn skills, to be motivated to learn, to be self-regulated learner, to develop positive self-esteem and confidence, to develop a belief on their effort rather than on lack of ability, and finally, to improve their learning.

The result of the current study, pertaining to teachers' action to engage their students actively in the lesson, particularly the quantitative data (see Table 4.2 of Chapter Four), divulged that most science teachers:

- Encourage every students to actively participate in asking questions, answering and sharing of ideas;
- Encourage students to take risks and listen to others ideas carefully;
- Provide examples of quality work in the lesson;
- Ask oral questions, give class work, and home works;
- Repeat learning objectives and assessment criteria during the lesson to shape the learning direction; and
- Advise students to assess their own work against learning objectives.

Whereas, the qualitative data suggested that most teachers in the selected schools did not engage every student actively in the lesson to improve their learning.

The observed and interviewed practice showed that:

- Students were not allowed to share ideas during a lesson;
- Teachers predominantly used lecture method in their lesson;
- Even some teachers did not see what students do at the backside, some students chat, laugh, and disturb nearby students;
- Students were not encouraged to ask questions, but top students randomly ask question without getting the chance; and
- The majorities of students passively listen and take notes what teachers and some students say in the lesson.

Thus, we can conclude that most teachers have a theoretical knowledge on the provision of support to engage students actively in their lesson, but fail to put it into practice. The results of the qualitative data revealed that only a few fast students are participated dominantly to answer teachers' oral questions. It was also supported by the quantitative data (see Table 4.2, item

number 38) most of science teachers reported that they frequently “encourage students to answer questions quickly.” In general, most of the observed lessons are teacher driven, except one mathematics lesson which tried to make the lesson student driven to some extent. Such activities, in turn troubled the development of independent, confident, self-regulated, and creative science learners who are vital for today’s knowledge based society.

### **Learning contents assessed**

Based on the results of the current study, concerning the learning aspects that teachers mostly assessed, the following conclusions are made:

In accordance with Table 4.2, science teachers mostly assessed students’ learning difficulties in the lesson and knowledge of scientific facts. The qualitative data also supported that. Most of the interviewees and lesson observations attested that teachers predominantly focused on assessing scientific facts, formulas, and principles which encourages memorization of facts and rote learning. In addition, the lesson observed showed that learning difficulties of students were not assessed to help them to learn more and to engage them actively in the lesson. Thus, the assessment tasks lacked alignment with learning objectives, students’ need, and real world tasks to achieve the intention of the current education system. Most teachers attribute it with students’ different understanding level and shortage of instructional time.

From these we can conclude that the implementation of authentic assessment methods to assess students’ problem solving skills, creativity, scientific reasoning ability, and application of new ideas in the real works in the selected schools was very low. As a result, it inhibits students to apply or demonstrate their knowledge and skills in real-life context or problems.

### **Application of assessment evidences**

The application of the collected assessment evidences to plan next actions in learning is one of the basic elements in the cycle of formative assessment to improve students’ learning. And it is the vital manifestation of formative assessment that makes it different from other modes of

assessment and it is called assessment for learning. However, the outcomes of this study related to the application of assessment evidences for further learning seems very low. In general, the use of formative assessment as a tool for learning was very limited in the selected schools.

According to Table 4.2, most science teachers in the selected schools used the collected assessment evidences or results to modify their teaching strategies, identify the gaps in students learning, plan what to teach next, and to advise students how to fill the gaps in their learning.

In contrast, on the same table large number of teachers frequently used assessment evidences to:

- Record it for final results;
- Categorize students into different groups (i.e. higher, medium, and lower achievers);
- Approve students who scores high; and
- Tell their achievement on a task against other students.

Moreover, as to the qualitative data priorities were given to rank and classify students and to record results for final mark to decide whether the students pass or fail the required subjects. In the selected schools, formative assessment was not used for sake of improving students' learning and to adjust instruction. Thus, students were disappointed from using assessment evidences to identify their learning gaps and to take actions for their future learning.

Generally, for research question number one, there was a clear difference between what the teachers reported in the questionnaire with what they did and what they actually practiced in the classroom. What was evident from this study was that science teachers in the selected schools did not effectively embed formative assessment into their daily instructions to improve students' science learning, but rather to collect marks for final decisions. As a result, such assessment activities negatively affect the quality of science education in the selected schools to achieve the long-term vision of the country.

**Research question two:** *What are the major factors that affect teachers' practice of assessment for learning in science subjects?*

The quantitative data revealed that the major possible factors of teachers' practice of assessment for learning were as follows (as mentioned on Table 4.3):

- Large class size;
- Lack of available science resources (such as reference materials, lab rooms, lab equipment, and demonstration sites);
- Lack of instructional material (e.g. teachers' assessment guideline);
- Shortage of instructional time;
- Lack of support from colleagues and from school principals;
- Lack of professional development activities (such as in-service, pre-service trainings...);
- Lack of support from supervisors;
- Students' and teachers' negative perception on formative assessment; and
- Government mandates on assessment issues.

In addition, the qualitative data suggested that large content coverage and lack of integration between contents particularly of grade 11 mathematics subject, teachers' motivation and lack of media coverage for teachers' innovative work as possible factors that hinder the implementation of assessment for learning in science subjects. Data from lesson observation also confirmed that lack of well-organized and equipped laboratory rooms as a possible factor. Thus, all of the factors mentioned above hinder the effective integration of assessment for learning into daily instruction to increase science learning standards in the selected schools in East Gojjam Zone, Amhara Regional State, Ethiopia.

***Research question three:*** *How do second cycle secondary schools science teachers perceive assessment for learning?*

Regarding teachers' perception about assessment for learning, the current findings (see Table 4.4), revealed that most science teachers seemed to have positive perception about the instructional power of formative assessment and the active engagement of students in the assessment process to improve science learning standards.

In contrast, most science teachers' perceived:

- Formative assessment as a process that follows the teaching learning process;
- Detailed correction of students work is effective way of feedback to improve learning; and
- Scores or marks provide direction for students about their progress against learning objectives.

Moreover, most of the interviewed teachers perceived formative assessment as variety of tools that continuously used at the end of the lesson to evaluate students' understanding and to collect marks. The lesson observation also confirmed that. Most teachers practiced formative assessment accordingly. Thus, we can conclude that even if it seems that most teachers have positive perception on assessment for learning; their response instability on the items and the qualitative data disclosed their negative perception. As a result, assessment for learning was no implemented effectively to improve students' science learning in the selected schools. This clear discrepancy happened because of teachers' lack of appropriate knowledge and skills about formative assessment strategies and its role in learning. Generally, teachers who participated in this study perceived formative assessment as variety of tools used to collect a piece of marks at the end of the lesson continuously. Thus, school administrators, teacher training institutions, and policy makers should evaluate the appropriateness of the assessment trainings given to increase science teachers' competency to implement formative assessment to improve learning.

***Research question four:*** *To what extent do professional (pre-service and in-service) trainings help science teachers to integrate assessment with their daily instruction?*

For quality education, teachers' professional training is essential. For this reason, Ministry of Education in Ethiopia gives high attention for professional courses during pre-service and in-service training programmes. Moreover, great efforts were put to revise the previous curricula in line with the current learning paradigm, particularly active learning and continuous assessment methods were given higher attention to improve the quality of science learning.

The outcome of this research question in this study revealed that the majority science teachers took assessment trainings during pre-service and in-services programmes, but the assessment trainings they received had little value to integrate different formative assessment strategies into their daily instruction to improve students' science learning in the selected schools.

As can be seen in Table 4.5,

- Most science teachers took more than one assessment course during pre-services trainings and they believe that it has value in the classroom;
- Half of the subjects took in-services assessment trainings which has little value to implement formative assessment effectively; and
- Nearly every one of participants of this study commented that they need more training related to formative assessment strategies.

Moreover, the qualitative data suggested that the assessment training they took during pre-service and in-service training had little value to implement assessment for learning in science education. As most interviewed teachers, the training predominantly focused on summative assessment and summative nature of formative assessment. Moreover, particularly, the in-service assessment trainings are not given continuously, mostly they are a day long workshops incorporated in other trainings. Therefore, we can conclude that the assessment trainings teachers received in the selected schools was not adequate to integrate different assessment for learning strategies into their daily instruction in science subjects to improve learning.

***Research question five:*** *What type of support schools provide for the effective practices of assessment for learning in science subjects?*

The findings of this research question demonstrated lack of intensive support from school principals, supervisors and colleagues for science teachers to implement formative assessment effectively. Thus, students' science learning in the selected schools might not be as good as what it would be because of lack of exhaustive support from schools, particularly on the effective

implementation of formative assessment (see Table 4.6). The qualitative data also supported that. Most interviewed science teachers expressed that most school leaders and students mostly preferred and supported teachers who used dominantly teachers-centered methods rather than those teachers who implement student-centered teaching methods. This consequently affects teachers' practice of formative assessment to improve students' learning.

***Research question six:*** *Are there significant relationships between teachers' perception and school supports with teachers' overall practice of assessment for learning?*

The results of the multiple regression indicated that there is a significant relationship between teachers' perception and school supports with the practice of assessment for learning. 25% of the variations in the practice of assessment for learning were accounted by the two independent variables. And, large variation in the practice of assessment for learning was explained through the supports provided by school principals, supervisors, and colleagues than teachers' perception on assessment for learning. Thus, teachers' perception and school supports provided affects the implementation of formative assessment in the selected schools.

***Research question seven:*** *How do the training and background characteristics of teachers affect their practice of assessment for learning?*

The findings of this study showed that there is a statistically significant difference in the combined practice of assessment for learning across teachers who have different teaching experiences (see Table 4.14). Teachers' with below or equal to 7 years and above or equal to 16 years of teaching experiences implemented assessment for learning in their daily instruction more than teachers with 8 to 15 years of teaching experience. Teachers' long experience in the use of both summative and formative assessment in their lesson matters the implementation of assessment for learning in science subjects. Likewise, the newly trained teachers with the revised curricula are found to implement assessment for learning in science subjects in the selected schools. Particularly, the long and less experienced teachers showed relatively high level of practice in the collection learning evidences factor, rather than on the planning of formative assessment, interpretation and communication of evidences, support provided to engage students



actively, learning contents assessed, and application of assessment evidences. Here, what teachers' experienced have also value on the overall practice of assessment for learning to improve students' learning.

On the other hand, the results of the separate MANOVA revealed that class size, teaching load per week, and subject taught have no significant impact on the overall practice of assessment for learning in the selected schools. In addition, the pre-service and in-service assessment trainings teachers took have less value to integrate formative assessment strategies into their daily instruction, the trainings mostly focused on the summative nature of formative assessment. This, in turn affects the quality of science instruction in the selected schools. Thus, appropriate trainings should be given for science teachers to effectively implement formative assessment to improve students' learning, because teachers' professional knowledge and skill gap affects the quality of the teaching-learning process and students' achievement in the selected schools.

In the next section of chapter 5, contributions, recommendations, limitation, and summary of the study are presented.

### **5.3 Contributions of the study**

Integrating formative assessment strategies into daily instruction to improve learning is a current trend in the field of assessment. Previously, it was dominated by summative assessment perspectives, which focused on the evaluation of students' learning or understanding at the end of the lesson. For this reason, most educational experts, teachers, and students in Ethiopia understand formative assessment as tool that is used to evaluate learning objectives continuously and still practiced accordingly. School officials also recommended and acknowledged teachers to use tests, class works, assignments, home works, and mid exams at the end of the lesson or chapter to evaluate learning and collect pieces marks continuously to decrease attrition rate. Moreover, as to the knowledge of the researcher, most previous studies in Ethiopia focused on summative assessment and summative nature of formative assessment. Hence, the practice of assessment for learning in the context of Ethiopia is in its infant stage. Even, as mentioned in the

literature part, the practice of assessment for learning was still challenging worldwide, because of its' newness and other many related factors.

Thus, this study which aimed at investigating science teachers' practice of assessment for learning and identifying underpinning factors of such practice can contribute a new dimension to the field of assessment and strengthen the theories of formative assessment, particularly in the context of Ethiopia. Even if this study was not conducted without any limitation, however, the researcher put great effort to use both the quantitative and qualitative approaches to deliver detailed and valid information about the practice of assessment for learning in the selected schools. Accordingly, this study revealed a number of conclusions that inform policy makers, education experts, researchers, school leaders, teachers, and students about the integration of different assessment for learning strategies into daily instruction and possible factors that hinder this practice in the selected schools.

Thus, this study is useful to clarify the pedagogical power of formative assessment on learning, strategies of integrating different formative assessment methods into daily instruction, major factors hindering the practice of assessment for learning, teachers' perception about assessment for learning, the adequacy of school leaders' support, and the relevance of pre-service and in-service assessment trainings to implement assessment for learning. The findings of this study are also helpful for other researchers who want to conduct comprehensive investigation in the area.

## **5.4 Recommendation**

### **5.4.1 Recommendation for improving assessment for learning practice**

Based on the findings of this study, the following issues need attention to implement assessment for learning effectively in science subjects to improve students' learning in East Gojjam Zone second cycle secondary schools:

- The results revealed that second cycle secondary school science teachers need more and relevant training on:

- Formative assessment strategies and their purpose in learning, for instance, how to share learning objectives and assessment criteria, provision of descriptive feedback, active involvement of students through self-and peer assessment, and questioning strategies and dialogues;
  - The phases of formative assessment to integrate assessment with the daily instruction, such as , planning of formative assessment strategies as a an integral part of the lesson, collection of learning evidences, interpretation of evidences, and use of evidences for future learning.
- As a result, short term trainings, workshops, and seminars in the area of formative assessment for science teachers should be given on a regular basis so that they can have a deeper understanding of formative assessment strategies and its purpose in learning. To improve science learning standards in the selected schools, building the competencies of teachers to improve their assessment knowledge, skills and attitude should be prioritized, because realizing good practice of assessment for learning in second cycle secondary schools is a critical aspect to advance educational standards.
  - Based on the findings, this study suggested that a school support is needed to foster the implementation of assessment for learning in the selected schools. Thus, school administrators should construct strong linkage with nearby universities or other potential organization or work with other stakeholders and ensure that teachers are provided with relevant in-service training or workshops on assessment on regular basis to make sure that they maintain the current assessment strategies for learning skills.
  - This study also proposed that the assessment trainings given during pre-service and in-service programs predominantly focused on summative assessment and summative nature of formative assessment. Thus, policy makers should evaluate the relevance of the current assessment training (both the in-services and pre-services assessment trainings) programmes to implement the new assessment paradigm to improve learning.

- Furthermore, school principals should encourage teachers to observe other model teachers within or in other schools, provide opportunity for teachers to show if there are innovative works or best practices, conduct regular discussions with teachers about formative assessment, encourage them to conduct action research, provide access to use available materials in the community, encourage collaborative works among staffs, provide awareness trainings for students about assessment and their role in learning, motivate teachers, and avoid supervisory approaches, rather use participatory discussion methods with teachers.
- Based on the results of this study, the current teaching-learning system in the selected schools promoted memorization of facts, formulas, and procedures because of the nature of assessment that students are subjected to and the way the subject matter is presented. Most teachers in the selected schools focused on teaching to the test, collecting marks, and rush to cover contents. Thus, policy makers, school administrators, and teachers themselves should give great attention on the assessment tasks designed to be authentic to strengthen the development of problem solving skills, scientific reasoning abilities and innovative works instead of encouraging rote memorization.
- Assessment guidelines should be prepared not to collect marks using different assessment methods but to collect learning evidences regarding students' knowledge, skills and attitude in the subject for next steps in learning.
- The result of this study also indicated that there is science resource scarcity in the sampled schools, thus to implement authentic assessment that equip students with necessary skill and knowledge for the world work, well-equipped and organized laboratory rooms and reference materials are needed. Thus, the Ministry of Education, Regional Education Bureau, Zone Education Department, and School Administrators should give due attention to find ways to fulfill science laboratory equipments and to organize them in appropriate manner.

- Teachers in this study also prescribed the time fixed for one period (40 minutes) as a major hindrance to integrate assessment in their instruction, hence, school principals should arrange the time schedule by embedding two or three periods into one to lengthening teachers life in the classroom and to reduce wastage of time.

#### **5.4.2 Recommendation for future research**

Based on the critical analysis of the main findings of this study, it may be useful to conduct research in the following focus areas:

- To gain deeper insights about the practice of assessment for learning and hindering factors in the selected schools, an action research study, which will encourage teachers to look at their own practice of formative assessment, could be conducted.
- In addition, future studies should be conducted on exploring the extent to which the collected assessment evidences are used to improve learning and adjust instruction.
- A study to investigate the perceptible gap between the teachers' theoretical knowledge and actual classroom assessment practices should be conducted. This study has recognized that there is a difference between what teachers perceived formative assessment (for example, their perception on the power of formative assessment to improve learning, and active involvement of students in the lesson) is and what they actually practice in the classroom.
- A study is needed to investigate students' perception of themselves and a teacher. The expectation of students about themselves as a student and a teacher may play a role in student's participation in the lesson, particularly in the assessment process because, if students expected themselves as a passive listener and note takers and teachers as the only source of knowledge, it is impossible to implement the new assessment perspective.

- School principals', supervisors', and other staffs' perception of a good teacher may also influence the way they approach and support teachers in the teaching learning process, particularly to implement assessment for learning. Thus, undertaking a study on this issue may be useful to improve learning standards.
- Students' perception of the nature and power of formative assessment on their learning, particularly on their role in the assessment process to improve their learning is also a potential area for research.
- It might also be valuable to more deeply examine teachers' perception and their practice of assessment for learning strategies such as: provision of constructive feedback; sharing of learning objectives and assessment criteria; better questioning strategies and dialogues; use of student-centered assessment tasks such as self-assessment, peer assessment, self-reflection; and their perception on how knowledge is constructed.
- In addition, there might be a need for substantial empirical research on the power of formative assessment to improve students' learning (low, medium and high achievers) and to examine the relationship between the current assessment practice and students' learning.

## **5.5 Limitation of the study**

Even if this study revealed important findings that can be used as an input for both policy makers and actual classroom practice, such findings are too limited to give a comprehensive picture of the current practice of assessment for learning and underpinning factors of these practices in science subject across all second cycle secondary schools in Amhara Regional state. This study is only limited in East Gojjam Zone Second Cycle Secondary Schools. The data sources are also limited only to science teachers in the selected schools, other subject teachers, students, other staff members, parents, school principals, and supervisors were not included. Document reviews were not included also to get more valuable data. Therefore, further research which includes large areas, different audiences, various data sources, and different grade level is recommended.

## 5.6 Summary of the study

The aim of this study was to gain an insight into second cycle secondary school science teachers' practice of assessment for learning, as well as to identify the possible factors that hinder the practice of assessment for learning in these schools. Based on this the following research questions were forwarded:

- To what extent do teachers integrate assessment in their instruction to improve students' science learning?
- What are the major factors that affect teachers' practice of assessment for learning in science subjects?
- How do second cycle secondary school science teachers perceive assessment for learning?
- To what extent do professional (pre-service and in-service) trainings help science teachers to integrate assessment with their daily instruction?
- What type of support schools provide for the effective practices of assessment for learning in science subjects?
- Are there significant relationships between teachers' perception and school supports with teachers' overall practice of assessment for learning?
- How do the training and background characteristics (teaching experience, teaching subject, work load per week and class size) of teachers' affect their practice of assessment for learning?

To answer these research questions, both quantitative and qualitative data were collected from science teachers in the selected schools using questionnaire, semi-structured interview, and lesson observation. Totally there are 18 second cycle secondary schools in East Gojjam Zone. The schools were clustered into two based on their year of establishment, after that, 5 schools from each cluster were randomly selected. All teachers in the selected schools were made to fill in the questionnaire but from the total of 186 teachers, only 156 teachers returned the questionnaire. Finally, 153 of the questionnaire were included in the analysis and the rest three questionnaires were rejected because of their response partiality. The questionnaire has six

sections (see Appendix D). The Practice, Perception, and Possible Factor sub-scales (sections) of the questionnaire were factor analyzed using exploratory factor analysis method to reduce the number of items into manageable factors. Moreover, eight purposively selected science teachers were included in the interview and observation session. The observation was made two times for five teachers and for the rest three teachers the observation was conducted once because of inconvenience.

The quantitative data were analyzed using descriptive statistics (such as frequency, percentage, mean, and standard deviation) and inferential statistics (such as one sample t-test, multiple regression, separate MANOVA, and univariate ANOVA). The significant level was tested at  $\alpha = 0.05$ . However, in the case of univariate tests adjusted alpha level was used (i.e. 0.0083) to decrease the rate of Type I error. Tukey post hoc multiple comparisons was used to assess significance differences between group means after statistically significant differences were observed using univariate ANOVAs. All the preliminary assumptions were checked before the analysis. Whereas, the qualitative data were analyzed using phenomenological data analysis steps and tasks proposed by (Cohen, Manion, and Morrison, 2007: 285-286; Creswell, 2007: 88-89; Willig, 2008: 55-59; Creswell, 2009: 21-22; De Vos et al., 2011: 411) (see Section 3.7.2). Different methods were also employed to maintain the methodological quality of the instruments.

The analyses of the quantitative and qualitative data led to a number of relevant research results which addressed the aim of this study. These were: the practice of assessment for learning or making formative assessment an integral part of the daily instruction to improve students' science learning in the selected school was very low. Most science teachers' in these schools hardly plan different assessment for learning strategies as a part of their lesson, collect learning evidences using different assessment for learning techniques, identify learning gaps for future learning, provide feedback in descriptive or constructive manner, provide support to engage students actively in the lesson, assess higher order learning outcomes, and use assessment evidences for next step in learning.



Teachers reported, lack of well-equipped and organized laboratory rooms and reference materials, large class size, shortage of instructional time, lack of school support, lack of appropriate professional development activities, lack of instructional materials, students negative perception of themselves and of formative assessment, teachers' lack of knowledge, skill and negative perception about formative assessment, and large content coverage as the possible factors to implement assessment for learning.

Even if teachers seemed to have positive perception about assessment for learning strategies, the actual practice and their report inconsistency revealed that most science teachers perceived formative assessment as a variety of tools used to evaluate students' understanding at the end of the lesson or chapter and to collect pieces of marks to decrease students' attrition rate.

School principals and supervisors support were also inadequate in the selected schools. Moreover, the pre-and in-service assessment trainings have less contribution to integrate formative assessment into daily instruction, because they lacked focus on the learning value of formative assessment strategies.

In addition, this study revealed a significant relationship between teachers' perception and school supports with the overall practice of assessment for learning. The two variables explained 25% of the variance in the overall practice of assessment for learning. From this, large contribution was accounted by school supports. Whereas, teaching load per week, class size, and subject taught except teaching experience have not significant effect on the overall practice of assessment for learning. Moreover, the pre-service and in-service assessment trainings given did not contribute significantly for science teachers to practice assessment for learning. Thus, appropriate professional trainings and supports should be provided for science teachers in the selected schools to build their positive perception about formative assessment strategies and skills to put it into practice effectively to improve students' learning. Finally, based on these findings recommendation were forwarded to improve the practice of assessment for learning in the selected schools and for further investigations in the area.

## REFERENCE

- Adair-Hauck, B., Glisan, E. W., Koda, K., Swender, E., & Sandrock, P. (2006). The Integrated Performance Assessment (IPA): Connecting Assessment to Instruction and Learning. *Foreign Language Annals*, 39, 359-382.
- Adula, B. H., & Kassahun, M. T. (2010). Qualitative Exploration on the Application of Student-centered Learning in Mathematics and Natural Sciences: The case of Selected General Secondary Schools in Jimma, Ethiopia. *Ethiopian Journal of Education and Science*, 6, 41-58.
- Akerson, V., McDuffie, A. R., & Morrison, J. A. (2006). Performance Assessment Tasks as A Stimulus for Collaboration Among Preservice and Inservice Teachers. In M. McMahon, P. Simmons, R. Sommers, D. DeBaets, & F. Crawley (Eds.), *Assessment in Science: Practical Experience and Education Research* (pp. 159-171). USA.
- Alkharusi, H. (2008). Effects of Classroom Assessment Practices on Student Achievement Goals. *Educational Assessment*, 13, 243-266.
- Allen, R. (2007). *Priorities in Practice: The Essentials of Science, Grades 7-12: Effective Curriculum, Instruction and Assessment*. USA: Association for Supervision and Curriculum Development (ASCD).
- 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*, 5-16.
- AREB. (2010). *Strategies and Guideline for Curriculum Implementation*. Bahir Dar.
- Assessment Reform Group. (1999). *Assessment for Learning: Beyond the Black Box*. England: University of Cambridge.
- Assessment Reform Group. (2002). *Assessment for Learning: 10 Principles: Research-Based Principles to Guide Classroom Practice*. Nuffield Foundation.
- Assessment Reform Group. (2002). *Testing, Motivation and Learning*. Cambridge: University of Cambridge Faculty of Education.
- Assessment Reform Group. (2003). *The Role of Teachers in the Assessment of Learning*. University of Cambridge.
- Baird, J. A. (2011). Does the learning happen inside the black box?., *Assessment in Education: Principles, Policy and Practice*, 18, 343-345.

- Bass, K. M., & Glaser, R. (2004). *Developing Assessment to Inform Teaching and Learning*. Los Angeles: The Regents of the University of California.
- Bell, B., & Cowie, B. (2001). *The Characteristics of Formative Assessment in Science Education*. Hamilton, New Zealand: John Wiley & Sons, Inc.
- Bell, B., & Cowie, B. (2002). *Formative Assessment and Science Education*. New York: Kluwer Academic Publishers.
- Bennett, R. E., & Gitomer, D. H. (2009). Transforming K-12 Assessment: Integrating Accountability Testing, Formative Assessment, and Professional Support. In J. J. Comming, & C. Wyatt-Smith (Eds.). New York: Springer.
- Bines, H., and Woods, E. (2007). *Country Profile Commissioned for the EFA Global Monitoring Report 2008, Education for All by 2015: Will We Make It? Cross Country (Eritrea, Ethiopia, Rwanda and Tanzania) Case study*. UNESCO.
- Birenbaum, M., Kimron, H., Shilton, H., and Shahaf-Barzily, R. (2009). Cycles of Inquiry: Formative Assessment in Service of Learning in Classrooms and in School-based Professional Communities. *Studies in Educational Evaluation*, 35, 130-149.
- Black, P., & Harrison, C. (2004). *Science Inside the Black Box*. London: National Foundation for Educational Research Nelson.
- Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy and Practice*, 5, 7-74.
- Black, P., & Wiliam, D. (1998). Inside the Black Box-Raising Standards Through Classroom Assessment. *Phi Delta Kappan*, 80, 139-148.
- Black, P., & Wiliam, D. (2001). *Inside the Black Box: Raising Standards Through Classroom Assessment*. London: King's College London School of Education.
- Black, P., & Wiliam, D. (2006). Assessment for Learning in The classroom. In J. Gardner (Ed.), *Assessment and Learning* (pp. 9-25). London: SAGE Publication.
- Black, P., & Wiliam, D. (2006). Developing a Theory of Formative Assessment. In J. Gardner (Ed.), *Assessment and Learning* (pp. 81-100). London: SAGE Publication.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for Learning: Putting it into Practice*. New York: Open University Press.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working Inside the Black Box: Assessment for Learning in the Classroom. *Phi Delta Kappan*, 86, 8-21.

- Bloxham, S., & Boyd, P. (2007). *Developing Effective Assessment in Higher Education: A Practical Guide*. New York: Open University Press.
- Boud, D. (2009). How can Practice Reshape Assessment? In G. Joughin (Ed.), *Assessment, Learning and Judgement in Higher Education* (pp. 29-44). Wollongong (Australia): Springer.
- Boud, D., & Falchikov, N. (2006). Aligning assessment with Long-term Learning. *Assessment and Evaluation in Higher Education*, 31, 399-413.
- Boud, D., & Falchikov, N. (2007). Developing Assessment for Informing Judgement. In D. Boud, & N. Falchikov (Eds.), *Rethinking Assessment in Higher Education: Learning for the Longer Term* (pp. 181-191). New York: Routledge.
- Broadfoot, P. (2008). Assessment for Learners: Assessment Literacy and the Development of Learning Power. In A. Havnes, & L. McDowell (Eds.), *Balancing Dilemmas in Assessment and Learning in Contemporary Education* (pp. 213-224). New York: Routledge.
- Brook, V. (2002). *Assessment in Secondary School: The New Teacher's Guide to Monitoring, Assessment, Recording, Reporting and Accountability*. USA: Open University Press.
- Brookhart, S. M. (2008). Feedback that Fits. *Educational Leadership*, 65, 54-59.
- Bruno, I., and Santos, L. (2010). Written Comments as a Form of Feedback. *Studies in Educational Evaluation*, 111-120.
- Buxton, C. A. (2006). Assessment in Support of Contextually Authentic Inquiry. In M. McMahon, P. Simmons, R. Sommers, D. DeBaets, & F. Crawley, *Assessment in Science: Practical Experience and Education Research* (pp. 173-182). USA: National Association for Research in Science Teaching.
- Buxton, C. A. (2006). Assessment in Support of Contextually Authentic Inquiry. In M. McMahon, P. Simmons, R. Sommers, D. DeBaets, & F. Crawley (Eds.), *Assessment in science : practical experiences and education research* (pp. 173-182). USA: National Science Teachers Association.
- Cauley, K. M., & McMillan, J. H. (2010). *Formative Assessment Techniques to Support Student Motivation and Achievement*. Virginia: Heldref Publications.
- Centre for Educational Research and Innovation (CERI). (2005). *Formative Assessment: Improving Learning in Secondary Classrooms*. Paris: Organisation for Economic Co-operation and Development (OECD).

- Centre for Educational Research and Innovation (CERI). (2008). Assessment for Learning: Formative Assessment. *OECD/CERI International Conference "Learning in the 21st Century: Research, Innovation and Policy"*. Organisation for Economic Co-operation and Development.
- Chappuis, S., & Stiggins, R. J. (2002). Classroom Assessment for Learning. *Educational Leadership*, 60, 40-43.
- Cheung, C., Cody, A., Hirota, I., Rubin, J., Slavin, J., Coffey, J., et al. (2006). Moving Beyond Grades and Scores: Reconsidering Assessment feedback. In M. McMahon, P. Simmons, R. Sommers, D. DeBaets, & F. Crawley (Eds.), *Assessment in Science: Practical Experience and Education Research* (pp. 207-218). USA: National Association for Research in Science Teaching.
- Chin, C. (2006). Classroom Interaction in Science: Teacher Questioning and Feedback to Students' Responses. *International Journal of Science Education*, 28, 1315-1346.
- Cizek, G. J. (1997). Learning, Achievement and Assessment: Constructs at a Crossroads. In D. P. Gary, *Handbook of Classroom Assessment: Learning, Achievement and Adjustment* (pp. 2-29). New York: Academic Press.
- Comming, J. J., & Wyatt-Smith, C. (2009). Framing Assessment Today for the Future: Issues and Challenges. In J. J. Comming, & C. Wyatt-Smith (Eds.), *Educational Assessment in the 21st Century: Connecting Theory and Practice* (pp. 1-16). New York: Springer.
- Council for the Curriculum Examinations and Assessment. (2007). *Assessment for Learning for Key Stage 3*. Northern Ireland Curriculum.
- Cowie, B. (2012). Focusing on the Classroom: Assessment for Learning. In B. J. Fraser, K. G. Tobin, & C. J. McRobbie (Eds.), *Second International Handbook of Science Education* (pp. 679-690). London: Springer.
- Cowie, B., Moreland, J., & Otrrel-cass, K. (2013). *Expanding Notions of Assessment for Learning: Inside Science and Technology Primary Classrooms*. Netherlands: Sense Publishers.
- Creemers, B., Kyriakides, L., & Sammons, P. (2010). *Methodological Advances in Educational Effectiveness Research*. USA: Routledge.
- Creswell, J. W. (2007). *Qualitative Inquiry and Research Design: Choosing Among Five Approaches* (2nd ed.). United Kingdom: Sage Publication Ltd.
- Creswell, J. W. (2009). *Research Design: Qualitative, Quantitative and Mixed Methods Approaches*. United Kingdom: Sage Publication Ltd.

- Creswell, J. W. (2012). *Educational Research: Planning, Conducting, and Evaluating quantitative and qualitative Research* (4th ed.). USA: Pearson Education, Inc.
- Crisp, G. T. (2012). Integrative Assessment: reframing Assessment Practice for Current and Future Learning. *Assessment and Evaluation in Higher Education*, 37, 33-43.
- Daniel, M. (2004). *Doing Quantitative Research in Education*. London: Sage Publications Ltd.
- David, C.-C. (2010). *Quantitative Psychological Research: The complete Student's Companion* (3rd ed.). New York: Psychology Press.
- De Vos, A., Strydom, H., Fouche, C., & Delport, C. (2005). *Research at Grass Roots: For the Social Science and Human Service Professions* (3rd ed.). Pretoria: Van Schaik.
- De Vos, A., Strydom, H., Fouche, C., & Delport, C. (2011). *Research at Crass Roots: For the Social Science and Human Service Professions* (4th ed.). Pretoria: Van Schaik.
- Department of Children, Schools and Famlies. (2010). *Learning Targets in Science: The National Strategie*. Crown copyright.
- Dunn, K., E. and Mulvenon, S.,M. (2009). A Critical Review of Research on Formative Assessment: The limited Scientific Evidence of the impact of Formative Assessment in Education. *Practical Assessment, Research and Evaluation*, 1-11.
- Dysthe, O. (2008). The Challenge of Assessment in a New Learning Culture. In A. Havnes, & L. McDowell (Eds.), *Balancing Dilemmas in Assessment and Learning in Contemporary Education* (pp. 15-28). New York: Routledge.
- Dysthe, O., Engelsen, K. S., Madsen, T. G., & Wittek, L. (2008). A theory-based discussion of assessment criteria: The balance between explicitness and negotiation. In A. Havnes, & L. McDowell (Eds.), *Balancing Dilemmas in Assessment and Learning: Contemporary Education* (pp. 121-132). New York: Routledge.
- Eshetu, A., Dilamo, O., Tsfaye, A., & Zinabu, G. (2009). Science and Mathematics Secondary Education in Ethiopia. *An Option paper presented at the Technical Workshop on Science and Mathematics Secondary Education in Africa (SEIA) Follow-up Program*. Tunis.
- Falchikov, N. (2005). *Improving Assessment Through Student Involvement: Practical Solutions for Aiding Learning in Higher and Further Education*. New York: Routledge.
- Falchikov, N. (2007). The place of peers in learning and assessment. In D. Boud, & N. Falchikov (Eds.), *Rethinking Assessment in Higher Education: Learning for the Long Term* (pp. 128-143). New York: Routledge.

- Falchikov, N., & Boud, D. (2008). The Role of Assessment in Preparing for Lifelong Learning: Problems and Challenges. In A. Havness, & L. McDowell (Eds.), *Balancing Dilemmas in Assessment and Learning in Contemporary Education* (pp. 87-100). New York: Routledge.
- Farenga, S. J., Joyce, B. A., & Ness, D. (2006). Adaptive Inquiry as the Sliver Bullet: Reconciling Local Curriculum, Instruction and Assessment Procedures with State-Mandated Testing in Science . In M. McMahon, P. Simmons, R. Sommers, D. DeBaets, & F. Crawley (Eds.), *Assessment in Science: Practical Experience and Education Research* (pp. 41-51). USA: National Association for research in Science Teaching.
- Fautley, M., & Savage, J. (2008). *Achieving QTS: Assessment for Learning and teaching in Secondary Schools*. Great Britain: Learning Matters Ltd.
- Federal Democratic Republic of Ethiopia (FDRE). (2010). *National Science, Technology and Innovation Policy: Building Competitiveness Through Innovation*. Addis Ababa.
- Firestone, W. A., Monfils, L. F., Schorr, R. Y., Hicks, J. E., & Martinez, M. (2004). Pressure and Support. In W. A. Firestone, L. F. Monfils, & R. Y. Schorr (Eds.), *The Ambiguity of Teaching to the Test: Standards, Assessment, and Educational Reform* (pp. 63-89). London: Lawrence Erlbaum Associates.
- Fisher, D., & Frey, N. (2007). *Checking for Understanding: Formative Assessment Technique for Your Classroom*. USA: Association for Supervision and Curriculum Development (ASCD).
- Fisseha, M. (2010). The Role of Assessment in Curriculum Practice and Enhancement of Learning. *Ethiopian Journal of Education and Science*, 5, 101-113.
- Fraenkel, J. R., & Wallen, N. E. (2009). *How to Design and Evaluate Research in Education* (7th ed.). New York: McGraw-Hill.
- Gamlema, S.M., and Smith, K. (2013). Student Perceptions of Classroom Feedback. *Assessment in Education: Principles, Policy and Practice*, 20, 150-169.
- Gardner, J. (2006). Assessment for Learning: A Compelling Conceptualization. In J. Gardner (Ed.), *Assessment and Learning* (pp. 197-205). London: SAGE Publication Ltd.
- Gardner, J. (2010). Developing Teacher Assessment: An Introduction. In J. H. Gardner, *Developing Teacher Assessment* (pp. 1-13). New York: Open University Press.
- Gioka, O. (2008). Teacher or Assessor? Balancing the Tensions Between Formative and Summative Assessment in Science Teaching. In A. Havnes, & L. McDowell (Eds.), *Balancing Dilemmas In Assessment and Learning in Contemporary Education* (pp. 145-156). New York: Routledge.

- Goodrum, D., Hackling, M., & Rennie, L. (2001). *The Status and Quality of Teaching and Learning of Science in Australian Schools*. Department of Education, Training and Youth Affairs.
- Graue, M. E. (1993). Integrating Theory and Practice Through Instructional Assessment. *Educational Assessment, 4*, 283-309.
- Graue, M.E., and Smith, S.Z. (2004). Instructional Innovation and Changing Assessments. In T. A. Romberg, *Standards-Based Mathematics Assessment in Middle School: Rethinking Classroom Practice* (pp. 25-45). New York: Teacher College Press.
- Greenstein, L. (2010). *What Teachers Really Need to Know About Formative Assessment*. USA: ASCD.
- Gulikers, J. T., Bastiaens, T. J., & Kirschner, P. A. (2008). Defining Authentic Assessment: Five Dimensions of Authenticity. In A. Havnes, & L. McDowell (Eds.), *Balancing Dilemmas in Assessment and Learning in Contemporary Education* (pp. 73-86). New York: Routledge.
- Gulikers, J.T.M., Biemans, H.J.A., Wesselink, R., and Wel, M. (2013). Aligning Formative and Summative assessment: A Collaborative Action Research Challenging Teacher Conception. *Studies in Educational Evaluation, 39*, 116-124.
- Hackmann, D. G. (2004). Constructivism and Black Scheduling: Making the Connection. *Phi Delta Kappan, 697-702*.
- Harlen, W. (2003). *Enhancing Inquiry Through Formative Assessment*. USA: Exploratorium.
- Harlen, W. (2006). The Role of Assessment in Developing Motivation For Learning. In J. Gardner (Ed.), *Assessment and Learning* (pp. 61-80). London: SAGE Publication.
- Harlen, W. (2010). Professional Learning to Support Teacher Assessment. In J. Gardner (Ed.), *Developing Teacher Assessment* (pp. 100-129). New York: Open University Press.
- Harlen, W., & Crick, R. D. (2003). Testing and Motivation for Learning. *Assessment in Education: Principles, Policy and Practice, 10*, 169-207.
- Harlen, W., and Gardner, J. (2010). Assessment to Support Learning. In J. Gardner (Ed.), *Developing Teacher Assessment* (pp. 15-28). New York: Open University Press.
- Harris, C. J., McNeill, K. L., Lizotte, D. J., Marx, R. W., & Krajcik, J. (2006). Usable Assessments for Teaching Science Content and Inquiry Standards. In M. McMahon, P. Simmons, R. Sommers, D. DeBaets, & F. Crawley (Eds.), *Assessment in science : practical experiences and education research* (pp. 67-88). USA: National Science Teachers Association.



- Hasan, M., & Ehsan, A. (2013). Present Situation of Teaching Elementary Science in the Primary School of Dhaka City. *American Journal of Educational Research*, 1, 576-582.
- Havnes, A. (2008). Assessment Dilemmas in Contemporary Learning Culture. In A. Havnes, & L. McDowell (Eds.), *Balancing Dilemmas in Assessment and Learning in Contemporary Education* (pp. 3-14). New York: Routledge.
- Havnes, A., Smith, K., Dysthe, O., and Ludvigsen, K. (2012). Formative Assessment and Feedback: Making Learning Visible. *Studies in Educational Evaluation*, 21-27.
- Henriques, L., Colburn, A., & Ritz, W. C. (2006). Developing Assessment Items: A How-to Guide. In M. McMahon, P. Simmons, R. Sommers, D. DeBaets, & F. Crawley (Eds.), *Assessment in science : practical experiences and education research* (pp. 15-30). USA: National Science Teachers Association,.
- Heritage, M. (2007). Formative Assessment: What Do Teachers Need to Know and Do. *Phi Delta Kappan*, 140-145.
- Heritage, M. (2010). *Formative Assessment and Next-Generation Assessment Systems: Are We Losing an Opportunity?* Washington, DC: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Heritage, M. (2011). Formative Assessment: An Enabler of Learning. *Evidence-Based Education*, 18-19.
- Herman, J. (2013). Formative Assessment for Next Generation Science Standards: A Proposed Model. *Invitational Research Symposium on Science Assessment*.
- Higgins, R., Hartley, P., and Skelton, A. (2002). The Conscientious Consumer: Reconsidering the Role of Assessment Feedback in Student Learning. *Studies in Higher Education*, 53-64.
- Hill, M. F. (2011). 'Getting Traction': Enablers and Barriers to Implementing Assessment for Learning in Secondary Schools. *Assessment in Education: Principles, Policy and Practice*, 18, 347-364.
- Hobbs, M., Brown, E., & Gordon, M. (2009). Learning and Assessment with Virtual Worlds. In C. Spratt, & P. Lajbcygier (Eds.), *E-Learning Technologies and Evidence-Based Assessment Approaches* (pp. 55-75). USA: IGI Global.
- Hodgson, C., & Pyle, K. (2010). *A Literature Review of Assessment for Learning in Science*. National Foundation for Educational Research.
- Hofstein, A., Mamlok, R., & Rosenberg, O. (2006). Varying Instructional Methods and Assessment of Students in High School Chemistry. In M. McMahon, P. Simmons, R.

- Sommers, D. DeBaets, & F. Crawley (Eds.), *Assessment in Science: Practical Experience and Education Research* (pp. 139-148). USA: National Association for Research in Science Teaching.
- Improving Quality of Science Teacher Training in European Cooperation (IQST). (2008). *Assessing Science for Understanding: A Constructivist Approach*.
- Irons, A. (2008). *Enhancing Learning Through Formative Assessment and Feedback*. New York: Routledge.
- James, M. (2006). Assessment, Teaching and Theories of Learning. In J. Gardner (Ed.), *Assessment and Learning* (pp. 44-60). London: SAGE Publication.
- Jones, C. A. (2005). *Assessment for Learning*. London: Learning and Skills development Agency.
- Jones, L., & Somekh, B. (2005). Observation. In B. Somekh, & C. Lewin (Eds.), *Research Methods in the Social Sciences* (pp. 138-145). London: Sage Publication, Ltd.
- Keeley, P. (2008). *Science Formative Assessment: 75 Practical Strategies for Linking Assessment, Instruction and Learning*. United Kingdom : Corwin Press.
- King, M. D. (2006). Assessment in Support of Conceptual Understanding and Student Motivation to Learn Science. In M. McMahon, P. Simmons, R. Sommers, D. DeBaets, & F. Crawley (Eds.), *Assessment in Science: Practical Experience and Education Research* (pp. 31-39). USA: National Association for Research in Science Teaching.
- Kola, A. J. (2013). Importance of Science Education to National Development and Problems Militating Against Its Development. *American Journal of Educational Research*, 1, 225-229.
- Kubitskey, B., Fishman, B. J., Margerum-Leys, J., Fogleman, J., Brunvand, S., & Marx, R. W. (2006). Professional Development and Teacher Learning: Using Concept Maps in Inquiry Classrooms. In M. McMahon, P. Simmons, R. Sommers, D. DeBaets, & F. Crawley (Eds.), *Assessment in Science: Practical Experience and Education Research* (pp. 107-118). USA: National Association for Research in Science Teaching.
- Kvale, S. (2007). Contradiction of Assessment for Learning in institutions of Higher Learning. In D. Boud, & N. Falchikov (Eds.), *Rethinking Assessment in Higher Education: Learning for the Longer Term* (pp. 57-71). New York: Routledge.
- Kvale, S. (2008). A Workplace Perspective on School Assessment. In A. Havness, & L. McDowell, *Balancing Dilemmas in Assessment and Learning in Contemporary Education* (pp. 197-208). New York: Routledge.

- Lambert, D., & Lines, D. (2001). *Understanding Assessment: Purpose, Perception, Practice*. London and New York: Routledge.
- Lasonen, J., Kemppainen, R., and Raheem, K. (2005). *Education and Training In Ethiopia: An Evaluation of Approaching EFA Goals*. Institute for Educational Research.
- Lauvas, P. (2008). Changing Assessment Practice in Norwegian Higher Education: From Where to Where? In A. Havnes, & L. McDowell (Eds.), *Balancing Dilemmas In Assessment and Learning in Contemporary Education* (pp. 158-168). New York: Routledge.
- Lee, C. (2006). *Language for Learning Mathematics Assessment for Learning in Practice*. England: Open University Press.
- Lee, C. (2006). *Language for Learning Mathematics Assessment for Learning in Practice*. New York: Open University Press.
- Libman, Z. (2010). Alternative Assessment in Higher Education. *Studies in Educational Evaluation*, 62-68.
- Liu, N., and Carless, D. (2006). Peer Feedback: The Learning Element of Peer Assessment. *Teaching in Higher Education*, 279-290.
- Liu, O. L., Lee, H.-S., & Linn, M. C. (2011). An Investigation of Explanation Multiple-Choice Items in Science Assessment. *Educational Assessment*, 16, 164-184.
- Lodico, M. G., Spaulding, D. T., & Voegtle, K. H. (2010). *Methods in Educational Research: From Theory to Practice* (2nd ed.). USA: Jossey-Bass.
- Lopez-Fernandez, O., & Molina-Azorin, J. F. (2011). The use of mixed methods research in the field of behavioural sciences. *Qual Quant*, 45, 1459-1472.
- Louis, C., Lawrence, M., & Keith, M. (2007). *Research Methods in Education* (6th ed.). USA: Routledge.
- Louis, C., Lawrence, M., & Keith, M. (2007). *Research Methods in Education*. New York: Routledge.
- Manitoba Education, Citizenship and Youth. (2006). *Rethinking Classroom Assessment with Purpose in Mind: Assessment for Learning, Assessment as Learning, and Assessment of Learning*. British Columbia: Crown Press.
- Mansell, W., James, M., & ARG. (2009). *Assessment in Schools. Fit for Purpose? A Commentary by the Teaching and Learning Research Programme*. London: Economic and Social Research Council, Teaching and Learning research Programme.

- Martinez, J.F., Stecher, B., and Borko, H. (2009). Classroom Assessment Practice, Teacher Judgments, and Students Achievement in Mathematics: Evidence from the ECLS. *Educational Assessment, 14*, 78-102.
- McDowell, L. (2008). Students' Experiences of feedback on Academic Assignments in Higher Education. In A. Havnes, & L. McDowell (Eds.), *Balancing Dilemmas in Assessment and Learning in Contemporary Education* (pp. 237-249). New York: Routledge.
- Mebea, F. (2008). *The Challenges of Implementing Continuous Oral Assessment in EFL Classes with Reference to Debre-Birhan Teachers' College: A Thesis Presented to the Department of Foreign Languages and Literature*. Addis Ababa.
- Mentkowski, M. (2006). Accessible and Adaptable Elements of Alverno Student Assessment - as- Learning: Strategies and Challenges for Peer Review. In C. Bryan, & K. Clegg (Eds.), *Innovative Assessment in Higher Education* (pp. 48-63). New York: Routledge.
- Ministry of Education. (1994). *Education and Training Policy* (1st ed.). Addis Ababa: ST. George Printing Press.
- MoE. (2002). *The Education and Training Policy and Its Implementation*. Addis Ababa.
- MoE. (2005). *A Presentation of the Ethiopian Education Sector Development Programme (ESDP)*. Addis Ababa.
- MoE. (2010). *Ethiopian Education Sector Development Program IV*. Addis Ababa.
- Montgomery, M. (2010). What is Happening in the UK. In J. Gardner (Ed.), *Developing Teacher Assessment* (pp. 53-69). New York: Open University Press.
- Moss, C. M., & Brookhart, S. M. (2009). *Advancing Formative Assessment in Every Classroom: A Guide for Instructional Leaders*. USA: ASCD.
- Murphy, R. (2006). Evaluating New Priorities for Assessment in Higher Education. In C. Bryan, & K. Clegg (Eds.), *Innovative Assessment in Higher Education* (pp. 37-47). New York: Routledge.
- Murphy, S. (2008). Some Consequence of Writing Assessment. In A. Havness, & L. McDowell (Eds.), *Balancing Dilemmas in Assessment and Learning in Contemporary Education* (pp. 33-50). New York: Routledge.
- National Research Council. (2001). *Knowing What Students Know: The Science and Design of Educational Assessment*. Washington, DC: National Academy Press.

- Nega, M. (2009). Quality of Pre-University Preparation, English Language Proficiency University Entrance Examination. *Quality of Higher Education in Ethiopian Public Institutions* (pp. 1-26). Addis Ababa: Forum for Social Studies.
- Nicol, D.J., and Macfarlane-Dick, D. (2006). Formative Assessment and Self-regulated Learning: A Model and Seven Principles of Good Feedback Practice. *Studies in Higher Education*, 199-218.
- Norton, L. (2007). Using Assessment to Promote Quality Learning in Higher Education. In A. Campbell, & L. Norton (Eds.), *Learning Teaching and Assessment in Higher Education* (pp. 92-100). British: Learning Matters Ltd.
- Nusche, D., Radinger, T., Santiago, P., & Shewbridge, C. (2013). *Synergies of Better Learning: An International Perspective on Assessment and Evaluation*. OECD.
- Odom, A. L., Marszalek, J. M., Stoddard, E. R., & Wrobel, J. M. (2011). Computers and Traditional Teaching Practices: Factors Influencing Middle Level Students' Science Achievement and Attitudes About Science. *International Journal of Science Education*, 33, 2351-2374.
- Organization for Economic Co-operation and Development. (2005). *Formative Assessment: Improving Learning in Secondary Classrooms*. Paris: OECD.
- Orr, S. (2008). Real or Imagined? The shift from Norm Referencing to Criterion Referencing in Higher Education. In A. Havnes, & L. McDowell (Eds.), *Balancing Dilemmas in Assessment and Learning in Contemporary Education* (pp. 133-143). New York: Routledge.
- Orsmond, P., and Merry, S. (2011). Feedback Alignment: Effective and Ineffective Links Between Tutors' and Students Understanding of Coursework Feedback. *Assessment and Evaluation in Higher Education*, 125-136.
- Ottevanger, W., Akker, J. v., & Feiter, L. d. (2007). *Developing Science, Mathematics and ICT Education in Sub-Saharan Africa: Patterns and Promising Practices*. USA: World Bank.
- Pallanat, J. (2001). *Spss Survival Manual: A Step by Step Guide to Data Analysis Using Spss*. Maidenhead: Open University Press.
- Panadero, E., Alonso-Tapia, J., & Reche, E. (2013). Rubrics vs. self-assessment scripts effect on self-regulation, performance and self-efficacy in pre-service teachers. *Studies in Educational Evaluation*, 1-8.
- Peel, D. (2005). The Significance of behavioural learning theory to the development of effective coaching practice. *International Journal of Evidence Based Coaching and Mentoring*, 18-28.

- Peterson, S. (2008). *The Teaching Doctor: Illuminating Evidence-Based Ideas for Effective Teaching*. Queen University : Office of health Science Education.
- Poulos, A., and Mahony, M.J. (2008). Effectivness of Feedback:The Students' Perspective. *Assessment and Evaluation in Higher Education*, 143-154.
- Race, P. (2007). *The Lecturer's Toolkit: A Practical Guide to Assessment, Learning and Teaching* (3rd ed.). New York: Routledge.
- Rggan, M., and Olah, L.N. (2011). Locating Interim Assessment Within Teachers' Assessment Practices. *Educational Assessment*, 1-14.
- Richardson, V. (2003). Constructivist Pedagogy. *Teachers College Record*, 1623-1640.
- Ridenour, C. S., & Newman, I. (2008). *Mixed Methods Research: Exploring the Interactive Continuum*. USA: Southern Illinois University Press.
- Roos, B. (2004). Learning and Assessment in the Knowledge Socitey. *New Challenges and Partnerships in an elarged European Union: Open, Distance and e-Learning in Support of Modernization*. Hungary: University of Technology and Economics.
- Sadler, D. (2010). Beyond Feedback: Developing Students Capability in Complex Appraisal. *Assessment and Evaluation in Higher Education*, 535-550.
- Scholtz, A. (2007). An analysis of the impact of an authentic assessment strategy on student performance in technology-mediated constructivist classroom. *Enternational Journal of Education and Development Using Information and Communication Technology*, 42-53.
- Sebba, J. (2006). Policy and Practice in Assessment for Learning: The Experience of Selected DECO Countries. In J. Gardner (Ed.), *Assessment and Learning* (pp. 185-196). London: SAGE Publication Ltd.
- Shafer, M. C. (2004). Expanding Classroom Practices. In T. A. Romberg, *Standards-Based Mathematics Assessment in Middle School: Rethinking Classroom Practices* (pp. 45-59). New York: Teachers College Press.
- Shepard, L. (2005). Formative assessment: Caveat emptor. *The Future of Assessment: Shaping Teaching and Learning*. New York.
- Shepard, L. (2005). Linking Formative Assessment to Scaffolding. *Educational Leadership*, 66-70.
- Shepard, L. A. (1991). Effects of High-Stakes Testing on Instruction. *American Educational Research Association* (pp. 1-37). Chicago: Office of Educational Research and Improvement, Department of Education.

- Shepard, L. A. (2000). The Role of Assessment in a Learning Culture. *Educational Research*, 29, 4-14.
- Shepard, L. A. (2000). *The Role of Classroom Assessment in Teaching and Learning*. Los Angeles: The Regents of the university of California.
- Shibeshi, A., Mekonnen, D., Semela, T., & Endawoke, Y. (2009). Assessment of Science Education Quality Indicators in Addis Ababa, Bahirdar and Hawassa Universities. *Quality of Higher Education in Ethiopian Public Institutions* (pp. 161-264). Addis Ababa: Forum for Social Studies.
- Singh, A. Yager, S.,O., Yutakom, N., Yager, R.,E., Ali, M.,M., (2012). Constructivist teaching practices used by five teacher leaders for the Iowa Chautauqua professional development program. *International Journal of Environmental and Science Education*, 197-216.
- Sjoberg, S. (2007). Constructivism and learning. In E. Baker, B. McGaw, & P. Peterson (Eds.), *International Encyclopaedia of Education* (3rd ed.). Oxford: Elsevier.
- Steadman, M., & Svinick, M. (1998). CATS: A Student's Gateway to Better Learning. *New Directions for Teaching and Learning*, 13-20.
- Stiggins, R. (2004). New Assessment Beliefs for a New School Mission. *Phi Delta Kappan*, 86, 22-27.
- Stiggins, R. (2005). From Formative Assessment to Assessment FOR Learning: A Path to Success in Standards- Based Schools. *Phi Delta Kappan*, 87, 324-328.
- Stiggins, R. (2006). Assessment for Learning: A Key to Motivation and Achievement. *Phi Delta Kappa International*.
- Stiggins, R. (2007). Assessment Through the Student's Eyes: Educating the Whole Child. *Educational Leadership*, 64, 22-26.
- Stiggins, R. (2008). *Assessment Manifesto: A Call for the Development of Balanced Assessment Systems*. Portland: Educational Testing Service.
- Stiggins, R. (2009). Assessment for Learning in Upper Elementary Grades. *Phi Delta Kappan*, 90, 419-421.
- Stiggins, R. J. (2002). Assessment Crisis: The Absence of Assessment For Learning. *Phi Delta Kappan International*.
- Stiggins, R., & DuFour, R. (2009). Maximizing the Power of Formative Assessment. *PHI DELTA KAPPAN*.
- Stobart, G. (2008). *Testing Times: The Use and Abuses of Assessment*. USA: Routledge.


- Suuramm, C., Koch, M., & Arden, A. (2010). Teachers' Assessment Practices in Mathematics: Classrooms in the Context of Reform. *Assessment in Education: Principles, Policy and Practice*, 17, 399-417.
- Swaffield, S. (2009). The Misrepresentation of Assessment for Learning and the woeful waste of a Wonderful Opportunity. *Association for Achievement and Improvement Through Assessment*. Cambridge.
- Swaffield, S. (2011). Getting to the Heart of Authentic Assessment for Learning. *Assessment in Education: Principles, Policy and Practice*, 18, 433-449.
- Swan, K. (2005). A constructivist model for thinking about learning online. In J. Bourne, & J. C. Moore (Eds.), *Elements of Quality Online Education: Engaging Communities*. Needham: MA: Sloan-C.
- Taber, K. S. (2011). Constructivism as Educational Theory: Contingency in Learning, and Optimally Guided Instruction. In J. Hassaskhah (Ed.), *Education Theory* (pp. 39-61). United Kingdom: NOVA Science Publisher, Inc.
- Tan, K. (2007). Conceptions of self-assessment: what is needed for long-term learning? In D. Boud, & N. Falchikov (Eds.), *Rethinking Assessment in Higher Education: Learning for the Longer term* (pp. 114-127). New York: Routledge.
- Tan, K. (2008). Academics' and Academic Developers' Views of Student Self-Assessment. In A. Havnes, & L. McDowell (Eds.), *Balancing Dilemmas in Assessment and Learning in Contemporary Education* (pp. 225-236). New York: Routledge.
- Tecahing Development Unit. (2009). *Assessment: Feedback to Promote Students Learning*. New Zealand: The University of Waikato.
- Torrance, H. (2007). Assessment as Learning? How the Use of Explicit Learning Objectives, Assessment Criteria and Feedback in Post-Secondary Education and Training can Come to Dominate Learning. *Assessment in Education*, 14, 281-294.
- Tucker, R., Fermelis, J., & Palmer, S. (2009). Designing, Implementing and Evaluating a Self-and-Peer Assessment Tool for E-Learning Environment. In C. Spratt, & P. Lajbcygier (Eds.), *E-Learning Technologies and Evidence-Based Assessment Approaches* (pp. 170-193). New York: Information Science Reference.
- Twing, J. S., Boyle, B., & Charles, M. (2010). Integrated Assessment for Improved Learning. *The 36th Annual Conference of the International Association of Educational Assessment (IAEA)* (pp. 22-27). Thailand: Bangkok.
- Underhill, A. F. (2006). Theories of Learning and Their Implications for on-line Assessment. *Turkish Oline Journal of Distance Education*, 165-174.



- Velayutham, S., Aldridge, J. M., & Fraser, B. J. (2011). Development and Validation of an Instrument to Measure Students' Motivation and Self-regulation in Science Learning. *International Journal of Science Education*, 2159-2179.
- Vitale, M. R., Romance, N. R., & Dolan, M. F. (2006). A Knowledge-Based Framework for the Classroom Assessment of Student Science Understanding . In M. McMahon, P. Simmons, R. Sommers, D. DeBaets, & F. Crawley (Eds.), *Assessment in Science: Practical Experience and Education Research* (pp. 1-14). USA: National Association for Research in Science teaching.
- Walport, M., Goodfellow, J., McLoughlin, F., Post, M., Sjøvoll, T., & Waboso, D. (2010). *Science and Mathematics Secondary Education for the 21st Century*. Crwn copyright.
- Weaver, M. R. (2006). Do Students Value Feedback? Student Perceptions of Tutors' Written Responses. *Assessment and Evaluation in Higher Education*, 379-394.
- William, D. (2006). Formative Assessment: Getting the Focus Right. *Educational Assessment*, 11, 283-289.
- William, D. (2008). Balancing Dilemmas: Traditional Theories and New Applications. In A. Havnes, & L. McDowell (Eds.), *Balancing Dilemmas in Assessment and Learning in Contemporary Education* (pp. 267-281). New York: Routledge.
- William, D. (2008). Changing Classroom Practice. *Educational Leadership*, 65, 36-42.
- William, D., Lee, C., Harrison, h., & Black, P. (2004). Teachers Developing Assessment for Learning: Impact on Students' Achievement. *Assessment in Education: Principles, Policy and Practice*, 11, 49-65.
- William, D. and Black, P. (2004). Teachers developing assessment for learning: impact on students achievement. *Assessment in education: Principles, policy and practice*, 49-65.
- Willig, C. (2008). *Introducing Qualitative Research in Psychology: Adventures in Theory and Method* (2nd ed.). New York: Open University Press.
- Willis, J. (2011). Afiliation, Autonomy and Assessment for Learning. *Assessment in Education: Principles, Policy and Practice*, 18, 399-415.
- Wragg, E. C. (2001). *Assessment and Learning in the Secondary School*. London : Routledge.
- Wyatt-Smith, C., & Gunn, S. (2009). Towards Theorising Assessment as Critical Inquiry. In C. Wyatt-Smith, & J. J. Comming (Eds.), *Educational Assessment in the 21st Century: Connecting Theory and Practice* (pp. 83-103). New York: Springer.

- Yeh, S. S. (2010). Understanding and Addressing the Achievement Gap Through Individualized Instruction and Formative Assessment. *Assessment in Education: Principles, Policy and Practice*, 17, 169-182.
- Yigzaw, A. (2013). High school English teachers' and Students' perceptions, attitudes and actual practices of continuous assessment. *Educational Research and Review*, 8, 1489-1498.
- Yung, B. H. (2006). Assessment Reform in Science. In B. H. Yung (Ed.), *Assessment Reform in Science: Fairness and Fear* (pp. 1-9). Hong Kong: Springer.

## APPENDIX A: ETHICAL CLEARANCE CERTIFICATE

**UNISA**   
University of South Africa

**COLLEGE OF EDUCATION RESEARCH ETHICS REVIEW COMMITTEE**  
15 October 2014

Ref : 2014/10/15/49059939/14/MC  
Student : Mí AA Dessie  
Student Number : 49059939

Dear Mr Dessie

**Decision: Ethics Approval**

---

**Researcher:** Mr AA Dessie  
Tel: +27913072854  
Email: askajeadamu@gmail.com  
Address: Debre Markos University  
Debre Markos,  
Ethiopia  
P O Box: 269

**Supervisor:** Prof. PJ Heeralal  
College of Education  
Department of Psychology of Education  
Tel: +2712 429 2318  
Email: [heerapi@unisa.ac.za](mailto:heerapi@unisa.ac.za)

**Proposal:** Teachers' Practices of Assessment for Learning in Science Education In East Gojjam Preparatory Schools, Amhara Regional State, Ethiopia

**Qualification:** D Ed in Didactics


---

Thank you for the application for research ethics clearance by the College of Education Research Ethics Review Committee for the above mentioned research. Final approval is granted for the duration of the research.

*The application was reviewed in compliance with the Unisa Policy on Research Ethics by the College of Education Research Ethics Review Committee on 15 October 2014.*

*The proposed research may now commence with the proviso that:*

- 1) The researcher/s will ensure that the research project adheres to the values and principles expressed in the UNISA Policy on Research Ethics.*
- 2) Any adverse circumstance arising in the undertaking of the research project that is relevant to the ethicality of the study, as well as changes in the methodology, should be communicated in writing to the College of Education Ethics Review Committee.*



University of South Africa  
Pretorius Street, Muckleneuk Ridge, City of Tshwane  
PO Box 392 UNISA, 0003 South Africa  
Telephone +27 11 429 3111 Facsimile +27 11 429 4150  
[www.unisa.ac.za](http://www.unisa.ac.za)

An amended application could be requested if there are substantial changes from the existing proposal, especially if those changes affect any of the study-related risks for the research participants.

- 3) The researcher will ensure that the research project adheres to any applicable national legislation, professional codes of conduct, institutional guidelines and scientific standards relevant to the specific field of study.

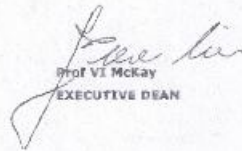
**Note:**

The reference number 2014/10/15/49059939/14/MC should be clearly indicated on all forms of communication [e.g. Webmail, E-mail messages, letters] with the intended research participants, as well as with the College of Education RERC.

Kind regards,



**Dr M Claassens**  
**CHAIRPERSON: CEDU RERC**  
mcdtc@nebactive.co.za



**Prof VI McKay**  
**EXECUTIVE DEAN**



University of South Africa  
Pretoria Campus, Muckleneuk Ridge, City of Tlokweng  
PO Box 192, UNISA, 0001 South Africa  
Telephone: +27 12 429 3111; Facsimile: +27 12 429 4190  
www.unisa.ac.za

# APPENDIX B: PERMISSION LETTER TO CONDUCT THE RESEARCH

EAST GOJJAM ZONE EDUCATION  
DEPARTMENT  
DEBREMARKOS



ref. Ed/833/1/11  
date 9/10/2006

To: Askalemariam Adamu

**RE: Permission to Conduct a Research Study in Preparatory Schools**

Thank you for your interest in conducting research in East Gojjam Second Cycle Secondary Schools' assessment for learning practice, particularly, in science education. Currently, the Ethiopia government gives high attention for science education, for the reason that science education is the heart to achieve economic development in the country. Moreover, effective continuous assessment for learning practice is the powerful means to improve students learning.

Thus, we would like to acknowledge the receipt of your application for research permit to conduct a study in second cycle secondary schools' in East Gojjam Zone. This serves to grant you permission to conduct your study in the sampled schools on the topic entitled: Teachers' assessment for learning practice in science education in East Gojjam preparatory schools, Amhara Regional State, Ethiopia.

With best regards

አዲስ አበባ ግንባታ ሚኒስቴር  
Aduzna Abajun Mehar

የምስክር ወረቀት ቁጥር 69  
2/16



የምስክር ወረቀት ቁጥር 69  
ደብዳቤ መረጃ

ስ.ቁ 058771 2297      058771 5580      fax 058771 4893      Email: - egoj.ed@ethionet.et  
058771 4093      058771 6852      ገጽ 20

## APPENDIX C: INFORMED CONSENT LETTER

Please read this document carefully before you decide to participate in this study

**Title of Research:** Teachers' Practice of Assessment for Learning in Science Education at East Gojjam Preparatory Schools, Amhara Regional State, Ethiopia.

**Researcher Name:** Askalemariam Adamu, DEd candidate, Didactics, College of Education, University of South Africa

**Supervisor Name:** Prof. PJH Heeralal, Department of Science and Technology, College of Education, University of South Africa

### A. Information to Participants

1. **Purpose of the research:** The purpose of the research is to examine teachers' practice of integrating assessment with daily instruction to improve students' science learning in preparatory schools and to identify the barriers of such practices.
2. **Research participation:** Participation in this study will involve: completing a questionnaire regarding your background information, your assessment for learning perception and practice, professional trainings you gained, school supports provided for you to implement assessment for learning and major factors that hinder your assessment for learning practice. Moreover, there is semi-structured interview and unstructured observation regarding the implementation of assessment for learning in science classrooms with selected schools and teachers. The interview will be held in a place of your choice.
3. **Selection of participants:** All of the participants in the selected school will be included in the study for completing the questionnaire. First, schools were clustered in to two based on their year of establishment. Second, five schools were selected randomly from each cluster to be included in this study. In the selected ten schools there are 186 science teachers. 127 teachers in cluster one and 59 teachers in cluster two. Thus, the sample size for this study will be 186. While, in the interview and observation session purposive sampling techniques will be employed to select more experienced teachers from the nearby four schools. Totally, eight teachers will be included in the interview and

observation sessions. The observation will be conducted two times before and after the interview session.

4. **Confidentiality:** The information gathered from this research will be kept confidential. Your identity will be coded, and all data will be kept in a secured, limited access location. Your identity will not be revealed in any publication or presentation of the results of this research. The data will be used to examine science teachers' practice of assessment for learning in second cycle secondary schools. It is applicable only for educational purpose.
5. **Risks and benefits:** There is no risk associated with participating in this research. This research is in no way connected to your efficiency. By participating, you are contributing to research to improve the practice of assessment for learning in science subjects to provide quality education.
6. **Right to withdraw:** participation in this study is completely voluntary. You can withdraw at any time with simple oral notice. You can ask any question regarding this study.
7. **Payment:** There will be no compensation.
8. **Time Limit:** This study will take approximately 5 months for data collection including the pilot study. Specifically, the questionnaire, interview, and observation session may take approximately 30 minutes, 40 minutes and 2 lessons respectively to complete.
9. **How will results be used:** The results of this research are only for educational purposes. The results will be shared with other researchers and participants through workshops and research symposiums to assess science teachers' practice of assessment for learning in second cycle secondary schools in East Gojjam Zone.
10. **Institution that gave Ethical approval:** East Gojjam Zone Education Department

## **B. Consent**

I have read and discussed the above information with the researcher. I have had the opportunity to ask questions about the purposes and procedures regarding this study. The purpose, nature, risks and benefits of the study has been fully explained to me by Askalemariam Adamu. I clearly understand what is being asked of me. My participation in this research is voluntary. I may refuse to participate or withdraw from participation at any time without prejudice, on simple oral notice and without having to justify my decisions.

By considering all of the above information, I agree to participate in the study.

Name of the teacher: \_\_\_\_\_

Signature: \_\_\_\_\_

Date: \_\_\_\_\_

#### Researcher's Verification of Explanation

I declare that I have carefully explained the purpose and nature of this research to preparatory science teachers and provided answers with the best of my knowledge to all of the questions asked.

Name of the researcher: \_\_\_\_\_

Signature: \_\_\_\_\_

Date: \_\_\_\_\_

For any questions regarding to this research, you can contact Askalemariam Adamu by telephone number: 0913072854; Email: [askleadamu@gmail.com](mailto:askleadamu@gmail.com)



## **APPENDIX D: QUESTIONNAIRE TO TEACHER**

Dear Teachers,

I am a doctoral student at the University of South Africa (UNISA). Now, I am doing my dissertation entitled “Teachers’ Practice of Assessment for Learning in Science Education in East Gojjam Preparatory Schools”. This is the reason why I am inviting you to participate in this research. The purpose of this study is to examine teachers’ current practice of using assessment to improve students’ learning and identify hindering factors of teachers’ practice of assessment for learning. Your responses to the items in this questionnaire are invaluable to conduct the research.

There is no incentive given to you, but this research is helpful to deepen your assessment understanding and to improve educational practices in Ethiopia. The questionnaire must be completed anonymously, and it may take approximately 30 minutes to complete the questionnaire. I would be very grateful if you could find the time to complete this questionnaire to assist me and play your own professional role in my research. Please feel free to write down your opinions and experiences.

Thank you in advance.

Askalemariam Adamu

Debre Markos University

Department of Psychology

Tel: 0913072854

Email: [askaleadamu@gmail.com](mailto:askaleadamu@gmail.com)

## Section I: Background Information

**Instruction:** Please indicate your response by marking "X" in the box for optional questions and by writing your response on the space provided

1. Gender: Female  Male
2. Educational background:  
 Certificate  Diploma  Bachelor Degree  Master's Degree
3. Teaching experience: \_\_\_\_\_years
4. Name of the school you are teaching in now \_\_\_\_\_
5. Which subject are you teaching now:  
 Mathematics  Biology  Chemistry  Physics
6. Teaching workload per week, including laboratory works: \_\_\_\_\_periods
7. The average number of students in each class in which you are teaching this year:  
 Below 35  35-50  51- 65  More than 65

## Section II: Science Teachers' Practice of Assessment for Learning

**Instruction:** Please indicate your appropriate response by marking an "X", how regularly you practice the assessment activities listed below, using the following 1-5 scale.

(1=Never; 2= Rarely; 3= Sometimes; 4=Often; 5= Always)

No	Items	Scale				
		1	2	3	4	5
1.	<b>How often do you practice the following formative assessment activity in your lesson?</b>	1	2	3	4	5
	• Identify learning objectives and assessment criteria					
	• Examine students' prior knowledge in the subject					
	• Plan how to share learning objectives and assessment criteria					
	• Design student-centered assessment methods and tasks					
	• Plan how and when to provide feedback					
	• Design better questions and questioning strategies					
2.	<b>How often do you employ the following assessment methods described below?</b>	1	2	3	4	5
	• Quizzes					
	• Tests					
	• Observations (using checklists, rating scales, anecdotal records)					
	• Class participation					
	• Project work					
	• Practical work					
	• Self-assessment					
	• Peer-assessment					

	<ul style="list-style-type: none"> <li>Laboratory work</li> </ul>					
	<ul style="list-style-type: none"> <li>Presentation</li> </ul>					
	<ul style="list-style-type: none"> <li>Homework</li> </ul>					
	<ul style="list-style-type: none"> <li>Class work</li> </ul>					
	<ul style="list-style-type: none"> <li>Group assignment</li> </ul>					
	<ul style="list-style-type: none"> <li>Criteria and objective setting with students</li> </ul>					
	<ul style="list-style-type: none"> <li>Written feedback</li> </ul>					
	<ul style="list-style-type: none"> <li>Oral feedback</li> </ul>					
	<ul style="list-style-type: none"> <li>Peer feedback opportunities</li> </ul>					
	<ul style="list-style-type: none"> <li>Students' feedback or reflection of ideas on the lesson learnt</li> </ul>					
	<ul style="list-style-type: none"> <li>Self- reflection through drawing, artifacts, concept mapping, etc.</li> </ul>					
	<ul style="list-style-type: none"> <li>Oral questions</li> </ul>					
	<ul style="list-style-type: none"> <li>Written questions in group</li> </ul>					
	<ul style="list-style-type: none"> <li>Open-ended questions</li> </ul>					
	<ul style="list-style-type: none"> <li>Student's questions</li> </ul>					
	<ul style="list-style-type: none"> <li>Peer-to-peer questions</li> </ul>					
	<ul style="list-style-type: none"> <li>Teacher –to- student dialogue</li> </ul>					
	<ul style="list-style-type: none"> <li>Student- to- student dialogue</li> </ul>					
	<ul style="list-style-type: none"> <li>Self-evaluation questions at the end of lesson</li> </ul>					
<b>3.</b>	<b><i>How often do you assess students...?</i></b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
	<ul style="list-style-type: none"> <li>Scientific reasoning ability in new situations</li> </ul>					
	<ul style="list-style-type: none"> <li>Application of problem solving skills in new issues</li> </ul>					
	<ul style="list-style-type: none"> <li>Knowledge of scientific facts</li> </ul>					
	<ul style="list-style-type: none"> <li>Higher order thinking (i.e. analysis, synthesis, ...)</li> </ul>					
	<ul style="list-style-type: none"> <li>Difficulties during the teaching learning process</li> </ul>					
<b>4.</b>	<b><i>How often do you carry out the following activities during a lesson as a part of assessment for learning strategies?</i></b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
	<ul style="list-style-type: none"> <li>Encourage every students to ask questions</li> </ul>					
	<ul style="list-style-type: none"> <li>Advise students to assess their own work against learning objectives</li> </ul>					
	<ul style="list-style-type: none"> <li>Advise students to assess others' work against learning objectives</li> </ul>					
	<ul style="list-style-type: none"> <li>Encourage students to take risks and listen to others ideas carefully</li> </ul>					
	<ul style="list-style-type: none"> <li>Encourage students to share ideas</li> </ul>					
	<ul style="list-style-type: none"> <li>Encourage students to answer questions quickly</li> </ul>					
	<ul style="list-style-type: none"> <li>Create opportunities for students to act on feedback provided</li> </ul>					
	<ul style="list-style-type: none"> <li>Engage every student to answer questions</li> </ul>					
	<ul style="list-style-type: none"> <li>Repeat the learning objectives and criteria during the lesson</li> </ul>					
	<ul style="list-style-type: none"> <li>Provide examples of quality work that shows the standards required</li> </ul>					
<b>5.</b>	<b><i>How often do you use the following criteria to interpret the assessment evidences that have been collected?</i></b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
	<ul style="list-style-type: none"> <li>Against learning objectives</li> </ul>					
	<ul style="list-style-type: none"> <li>Against their own previous performance</li> </ul>					

	<ul style="list-style-type: none"> <li>Compared to other students' result within the class</li> </ul>					
	<ul style="list-style-type: none"> <li>Compared to other class of students with the same grade level</li> </ul>					
	<ul style="list-style-type: none"> <li>Against assessment criteria or standards</li> </ul>					
<b>6.</b>	<b><i>How often do you provide the type of feedback described below to your students?</i></b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
	<ul style="list-style-type: none"> <li>Provide the detailed answers for each assessment task along with marks</li> </ul>					
	<ul style="list-style-type: none"> <li>Orally suggest on how to improve their work against learning objectives</li> </ul>					
	<ul style="list-style-type: none"> <li>Provide written comments on how to improve their work</li> </ul>					
	<ul style="list-style-type: none"> <li>Tell their achievement on a task against other students result</li> </ul>					
<b>7.</b>	<b><i>For what purpose do you regularly use assessment evidences or results? To:</i></b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
	<ul style="list-style-type: none"> <li>Categorize students into different groups</li> </ul>					
	<ul style="list-style-type: none"> <li>Plan what to teach next</li> </ul>					
	<ul style="list-style-type: none"> <li>Identify the gaps in students' understanding</li> </ul>					
	<ul style="list-style-type: none"> <li>Advise students how to fill the gap in their learning</li> </ul>					
	<ul style="list-style-type: none"> <li>Approve students who score high</li> </ul>					
	<ul style="list-style-type: none"> <li>Modify my teaching strategies accordingly</li> </ul>					
	<ul style="list-style-type: none"> <li>Suggest means for students to plan their future learning</li> </ul>					
	<ul style="list-style-type: none"> <li>Allow peer discussion on how to improve their work</li> </ul>					
	<ul style="list-style-type: none"> <li>Permit students to resubmit their work once they improved it</li> </ul>					
	<ul style="list-style-type: none"> <li>Record assessment results</li> </ul>					

### Section III: Factors Affecting the Practice of Assessment for Learning

**Instruction:** Please mark "X" how likely the following barriers hinder your effective formative assessment practices, based on the five options given at the right side of each item. (1=Very Unlikely; 2=Slightly Unlikely; 3=Neither; 4=Slightly Likely; 5=Very Likely)

No	Items	Scales				
		1	2	3	4	5
1.	Lack of professional development activities (such as in-service training, pre-service courses, workshops, etc.)					
2.	Students' negative perception on formative assessment					
3.	Teachers' negative perception on formative assessment					
4.	Shortage of instructional time					
5.	Impact of summative assessments (mid exams, final exams...)					
6.	Pressure of national examinations					
7.	Government mandates on assessment issues					
8.	Lack of support from colleagues					
9.	Lack of support from school principals					
10.	Lack of support from supervisors					

11.	Lack of available science resources (textbooks, lab rooms, lab equipment, demonstration sites...)					
12.	Lack of instructional materials (e.g. teacher's assessment guideline...)					
13.	Large number of students in one class					
14.	Many number of periods per week					

#### Section IV: Science Teachers' Perception on Assessment for Learning

*Instruction: Please indicate your appropriate response by marking an "X", the degree to which you agree with the statement listed below, using the following 1-5 scale.*

**(1 = Strongly Disagree; 2 = Disagree; 3= Neutral; 4 = Agree; 5 = Strongly Agree)**

No	Items	Scale				
		1	2	3	4	5
1.	Formative assessment is a series of tests administered to evaluate student's learning.					
2.	Formative assessment uses a variety of assessment tools to gather evidence of learning.					
3.	Using a variety of assessment methods provides useful insights about students' progress.					
4.	Using one crucial assessment method allows students to perfect their performance.					
5.	Formative assessment inspires students' deep thinking in science subjects.					
6.	Formative assessment is more vital to assess the effectiveness of teaching than final exams.					
7.	Formative assessment helps students to know more about their own learning problems.					
8.	Formative assessment allows teachers to make their students active.					
9.	Formative assessment helps teachers to easily identify students' problems in learning.					
10.	Formative assessment fosters students' internal motivation to learn science subjects.					
11.	Formative assessment is an integral part of the teaching learning process.					
12.	Formative assessment follows the teaching learning process.					
13.	Mixing formative assessment with daily lesson is useful to improve science education.					
14.	Formative assessment is vital to assess students' higher order learning (e.g. application, analysis, synthesis etc.) more than summative assessments.					
15.	Formative assessment improves every student achievement in science subjects.					
16.	Formative assessment is vital to ensure all students to have an experience of success and makes all of them competent learners.					
17.	Formative assessment helps students to develop positive self-esteem and confidence.					
18.	Formative assessment helps students to be independent learners.					
19.	Formative assessment is an engaging and enjoyable process for students.					
20.	Formative assessment is important to capture students' attention and effort.					
21.	Formative assessment reduces the rate of repetition and dropout more than summative assessment.					
22.	Formative assessment empowers students to study from day one class.					
23.	Sharing learning objective and assessment criteria to students is vital to improve learning.					

24.	Transparency in learning objectives and assessment criteria enable learners to assess their own progress.					
25.	Sharing learning objectives and assessment criteria motivates students to learn.					
26.	Feedback is a key component of formative assessment to improve learning.					
27.	Creating competitive classroom environment among students improves learning.					
28.	Feedback is helpful to identify and fill the gaps in students' science understanding.					
29.	Teacher-student dialogue during feedback is effective in fostering science learning (i.e. dialogue refers to exchange of ideas).					
30.	Feedback is a powerful technique to support or inhibit students' motivation to learn.					
31.	Detailed correction of students work is effective way of feedback to improve learning.					
32.	Scores provide direction for students about their progress against learning objectives.					
33.	Suggesting students on how to improve their work is effective feedback in learning.					
34.	Active involvement of students in the assessment process is vital to improve learning.					
35.	Using self-assessment in science subjects improves students' learning.					
36.	Using peer assessment in science subjects fosters students' learning.					
37.	Long waiting time during questioning is vital to engage every student in answering.					
38.	Questioning is effective to engage every student in the lesson through thinking.					

### Section V: Science Teachers' Professional Learning

**Instruction:** Please indicate your appropriate response by marking an "X" based on the two options on the right of each item. The statements indicate the professional training you got about formative or continuous assessment.

No	Items	Options	
		Yes	No
1.	I took classroom assessment course during pre-service training.		
2.	I took more than one course dedicated to classroom assessment.		
3.	The assessment method I learning during pre-service training had little value in the classroom		
4.	I took in-service training which focuses on formative assessment.		
5.	I received adequate training on how to design formative assessment methods.		
6.	I learned assessment method during in-services training that has little value in the classroom.		
7.	I received no training on formative assessment implementation.		
8.	I need more training about formative assessment to implement it effectively.		
9.	I understand how to implement formative assessment in my daily instruction.		

## Section VI: School Support Provided for Science Teachers

*Instruction: Please mark "X" to indicate the degree to which you agree with the statements listed below about the supports provided by your school based on the five options given at the right side of each item.(1=Strongly Disagree; 2= Disagree; 3= Neutral; 4= Agree; 5= Strongly Agree)*

No	Items	Scale				
		1	2	3	4	5
1.	School principals and supervisors support is very essential to implement formative assessment in science education.					
2.	Our school promotes collaborative work among science teachers.					
3.	Our school encourages mutual observation and sharing of experience within our school.					
4.	Our school encourages mutual observation and sharing of experience with other schools.					
5.	Our school gives time for teachers to meet on regular basis and provides opportunity for them to report their work.					
6.	Our school gives opportunity for teachers to tryout and evaluates new ideas.					
7.	Our school provides models of quality practice to further teachers' professional development.					
8.	Our school tries to find resource funds to fulfill science laboratory equipment.					
9.	Our school provides opportunity for short-term trainings on classroom assessment.					
10.	Our school motivates teachers to carry out research with one or more colleagues to improve classroom practice.					
11.	Our school acknowledges teachers who implement formative assessment effectively.					

**I thank you for your kind cooperation.**

## **APPENDIX E: INTERVIEW GUIDE FOR TEACHERS**

1. Please tell me who you are, including your title, academic rank, what courses you teach, and how long have you been teaching?
2. Why do you carry out assessment in the classroom?
3. What is formative assessment means to you?
4. Do you use formative assessment in your classroom? If so, what kind?
5. How do you describe your experience in sharing learning objectives and assessment criteria?
6. How do you describe your experiences with implementing peer and self-assessment?
7. How do you describe your experiences in providing feedback to your students?
8. How do you describe your experience in questioning technique?
9. For what purpose do you use assessment evidence?
10. Have you received classroom assessment trainings? If so? What kind? How do you describe their relevance to implement formative assessment?
11. What kind of support school principals, supervisors, department heads and colleagues provided to you to implement formative assessment effectively?
12. Are there challenges to implement formative assessment strategies? If so, what are?
13. What you suggest to improve the implementation of formative assessment in science subjects?
14. Is there anything we missed in our discussion? Would you like to add anything before we close?



## APPENDIX F: SAMPLE INTERVIEW TRANSCRIPTS

**Teacher E (Female, BSc degree in chemistry and 9 years of teaching experience):**

1. Why do you carry out assessment in the classroom?

She stated: *I used assessment to evaluate students understanding in the lesson, to know whether students get enough knowledge or not. After knowing their understanding I classify them into higher, medium and lower achiever, then as much as possible I help medium achievers to be high achiever and low achievers to be medium achiever. Basically, I used assessment to select students, secondly, to evaluate their understanding, because my intension was to make all students competent to pass into the next grade and to make them good citizens.*

2. What is formative assessment means to you? (What is the difference between formative and summative assessment for you?)

*Formative assessment means an assessment that I used to evaluate or assess students' day to day progress in a lesson. If we did not assess students' understanding continuously, we did not know students' daily progress in a subject. But, using formative assessment is both advantageous and disadvantageous. When we see its' advantage: It helps to know students' level of understanding in the subject, to identify and classify students based on their knowledge and to give tutorial for those students who achieve low.*

*However, if we implement formative assessment continuously, we are not teaching students to make them competent and to achieve the required competencies (skills, knowledge and attitude) in the subject. Our work becomes continuous assessment, which does not help students to be competent in a subject. Implementing formative assessment continuously does not make students to be knowledgeable and skillful in the subject.*

*Formative assessment by itself is not an active teaching method, but it helps us to modify our teaching methods. When students results becomes low, the teacher asks him/herself to*

*change his/her methodology to make all students high achiever or to modify their results. But we are not implementing it in our schools, because of large number of students' and our skill gap.*

3. Do you use formative assessment in your classroom? If so, what kind?

*Yes, I used various continuous assessment tools such as: Students' participation during group discussion, home works, classwork, test, quizzes, and assignment. I always give mark for any assessment task. For example, when I give group assignment from text books or from other sources, I randomly call students to reflect what they do in the group. It helps me to identify those students who actively participate in the group and who did not, to give marks accordingly. After that I give mark based on their report on the topic. Moreover, I give marks for quizzes, test, assignments, class attendances, home works and class works. After I do questions in the class and then I allow them to see it.*

4. How do you describe your experience in sharing learning objectives and assessment criteria?

*I always write it on the blackboard, because every text book starts from its objective and assessment tasks are designed based on it, but there are students who did not see it and they choice their own ways. Moreover, at the beginning of the lesson I told them the weight of quizzes, assignment etc.*

5. How do you describe your experiences with implementing peer and self-assessment?

*What do you mean self-and peer assessment? Is it assessing of myself? Or what it is? I did not use it. However, when I give group assignment ...I randomly call students to reflect what they do in the group to identify those students who actively participate in the group and who did not. After this I give different marks according to their reflections. If we allow them to assess themselves and their peers they are not volunteer, they see it carelessly.*

6. How do you describe your experiences in providing feedback to your students?

*What is feedback? Before scoring quizzes or test, assignments, class works... I do the questions together with students in the class then score it and allow them to see it. However, in the case of mid exam and final exam, first I score it and after we do the question together. In both cases I provide only their marks. Moreover, in the case of assignments I provide answers to be added and wrong answers to be removed.*

7. How do you describe your experience in questioning technique?

*During the lesson I mostly ask questions which are simple, but I did not write questions before the lesson.*

8. For what purpose do you use assessment evidence?

*Do you mean the result? I used assessment results to compare students' performance within the same class and with other class of students for the same subject. I used assessment results to classify students into different groups (high achiever, medium achiever and low achiever) and as a result to give tutorial accordingly. Moreover, I used it to create awareness among students about their level of understanding or rank against other students, because I believed that creating competitive environment between students is good to improve students' learning. Finally, I record the result on the mark sheet for final result to decide whether a student passes or fails in the subject. Moreover, the result is important to inform the school and student's parents.*

9. Have you received classroom assessment trainings? If so? What kind? How do you describe their relevance to implement formative assessment?

*I took one evaluation course during pre-service training and it is more valuable to prepare exams, but still I did not use the principles and rules to prepare tests. I did not take any in-service training related to formative assessment.*

10. What kind of support school principals, supervisors, department heads and friends provided to you to implement formative assessment effectively?

*I didn't remember any support related to assessment, but school principals and supervisor encourage and order us to use continuous assessment and to give mark for each task.*

11. Are there challenges to implement formative assessment strategies? If so, what are?

Yes;

- *Formative assessment takes time and the number of students should be small.*
- *Teachers' background and experience in using self-and peer assessment*
- *Students' background and experience in using self-and peer assessment and their level of understanding were the main factors that hinder the implementation of formative assessment. When I allocate tasks and permit them to assess themselves and their peers, they are not volunteer, rather they see it carelessly.*
- *Students' want the teachers to do everything for them.*
- *Students level of understanding (top, medium and low...). So to make effective teaching students should be classified based on their understanding. Because top students understand the lesson quickly, whereas other left.*

12. What you suggest to improve the implementation of formative assessment in science subjects?

*If we give mark for each assessment task formative assessment is more advantageous for medium achievers than lower achievers and high achievers, because high achievers do tasks quickly, whereas low achievers are not volunteer to do it. However, if we did not give mark for assessment tasks it may be advantageous for low achievers. But in our school we are not doing it. We used each assessment task for the collection of marks.*

13. Is there anything we missed in our discussion? Would you like to add anything before we close?

*No*