

AMERICAN UNIVERSITY OF ARMENIA

Department of English Programs

A Thesis Submitted in

Partial fulfillment of the Requirements for the Degree of Master of Arts in
Teaching English as a Foreign Language (TEFL)

Implementation of formative assessment and its effect in an Armenian setting

Thesis committee:

Irshat Madyarov, Ph.D.-Thesis Advisor

Alexan Simonyan, Ph.D.

Rubina Gasparyan-Thesis Reader

By:

Angela Balasanyan

Yerevan, Armenia

2011

AMERICAN UNIVERSITY OF ARMENIA

We hereby approve that this thesis

By

Angela Balasanyan

Implementation of formative assessment and its effect in an Armenian setting

be accepted in partial fulfillment for the requirements of the degree of M.A. in TEFL

Committee of the Thesis

Irshat Madyarov, Ph.D.-Thesis Advisor

.....

Rubina Gasparyan-Thesis Reader

.....

Catherine Buon: Associate Dean of DEP

.....

Yerevan, Armenia

2011

Dedication
To my family

Acknowledgments

I acknowledge with profound gratitude many people who helped me during my studies and research at AUA.

First, I would like to express my deepest appreciation to the supervisor of my thesis Irshat Madyarov for his willingness to help me find answers and solutions to the questions cropped up while completing my thesis, for his constructive criticism and valuable comments. His effective feedback and guidelines helped me to find the right track during the path of writing the thesis.

I owe my deepest gratitude to the reader of my thesis Mrs. Rubina Gaspayran for her valuable advice, support and assistance; who always kindly devoted her time and energy for reading my thesis and answering my questions, and who without hesitation, agreed to assist me in carrying out my research in the Experimental English Classes.

My special thanks go to Alexan Simonyan who assisted me in the most challenging part of my thesis, the statistical part. His guidelines and directed instructions helped me and facilitated my job through the statistical analysis of the data.

I would like also to thank Dr. Hossein Farhady for suggesting me the topic and giving some advice and guidelines during the initial part of the thesis. Moreover, I am thankful to all the instructors of the Department of English programs for their efforts and supports during these two years.

I'd like to thank my family, my father, mother and sister for their patience, love and understanding. And last but not least, I would like to thank my friends for encouraging and supporting me during my research and my studies at the American University of Armenia.

Table of Contents

Acknowledgments	3
Abstract.....	6
Chapter 1: Introduction.....	9
1.2 Purpose of the study	9
1.3. Significance of the study.....	12
Chapter 2: Review of the Related Literature	14
2. 1. Introduction.....	14
2.2. Formative vs. Summative Assessment	16
Table 1: The differences and goals of formative and summative assessment.....	17
Examples:	17
2.3. The purpose and benefits of Formative Assessment.....	18
2.4 Elements that formative assessment involves in practice: Formative Assessment Techniques	23
Chapter 3: Methodology	34
3.1. Research Design.....	34
3.2. Participants and Setting.....	34
3.3. Materials	35
3.3. Instrumentation	36

3.3.1. Tests	36
3.3.2. Questionnaires	37
3.3.3. Filed notes	38
3.4. Procedure and Treatment	39
3.4.1. Students' Work Samples	40
Chapter Four: Results and Discussion	43
4.1. Analysis of Quantitative Data	43
4.2. Analysis of Qualitative Data	56
4.2.2. Field notes	69
Chapter 5: Conclusion	72
5.1. Summary of the research findings	72
5.2. Limitations of the study	73
5.3. Suggestions for further research	74
References	75
Ferlazzo L. (2010) "The Best Resources for Learning about Formative Assessment" Samoset Middle School Blog	81
Appendices	87
Appendix #1	87
Appendix #2	93

List of Tables

Table 1: The differences and goals of formative and summative assessment.....	15
Table 2: Paired Sample Test for the within group comparison for the pretest for the experimental group	42
Table 3: Paired Sample Test for the within group comparison for the posttest for the comparison group.....	44
Table 4: Wilcoxon Signed Ranks Test for the within group comparison for the posttest for the experimental group.....	45
Table 5: Wilcoxon Signed Ranks Test for the within group comparison for the posttest for the comparison group	46
Table 6: Independent Samples T Test for the between group comparison for the pretest.....	47
Table 7: Independent Samples T Test for the between group comparison for the posttest	50
Table 8: Mann-Whitney U Test for the between group comparison for the pretest and the posttest.....	52
Table 9: A total number of student questionnaires' participants.....	54
Table 10: Responses of the questionnaire for the experimental group.....	55
Table 11: A total number of student questionnaires' participants.....	62
Table 12: Responses of the questionnaire for the comparison group.....	63
Table 13: Patterns emerged out of filed notes.....	67

Abstract

The purpose of the present study was to investigate what effect formative assessment can have on the achievement of the students learning English in an Armenian setting. The study was quasi experimental and was carried out in the Experimental English Classes (EEC) at the American University of Armenia (AUA). Two groups were involved in the study, the experimental and the comparison. The experimental group received treatment and the comparison group did not. Besides, the current research was aiming to find out the students' perceptions and attitude towards formative assessment.

The instruments that were used for collecting data were a pretest and a posttest, a questionnaire and field notes. The data collected from the pre and post tests between the two groups were analyzed using the parametric Independent Samples T Test and the non parametric Mann-Whitney U Test, in order to see the difference of their performance during and after the treatment. For the within group comparison, the parametric paired sample and the non parametric Wilcoxon Signed Ranks Tests were implemented. The data collected through the questionnaire were analyzed qualitatively.

The results revealed that that there was a significant difference in both groups during the performance of pretest and the posttest. However, the major limitation of the study should be taken into account. Because of some incorrectness and circumstances, the students' first exam of the course was taken as a pretest and during the test the experimental group performed better than the comparison group. In addition, during the post test the comparison group revealed better results. This can be proved through the results obtained by the Wilcoxon Signed Rank test comparing the effect sizes ($r=0.81$ (experimental) and $r=1.61$ (comparison)) of the two groups during the pretest and the posttest. So, it can be inferred that though both groups performed

better results during the posttest, the null hypothesis that there is no relationship between the implementation of formative assessment and the students' achievement is not rejected.

The qualitative data were in the form of questionnaires and field notes. The results obtained through the questionnaires revealed that the students' attitude towards formative assessment was positive. The field notes served as perceptions of the classes on the part of the teacher about the treatment in the experimental group and the core procedure of the class.

Chapter 1: Introduction

Like most terms in education, assessment evokes a wide range of conceptions and feelings. If you ask a middle school student why he is learning his particular lesson, concepts and terms, writes his assignments, etc., the response may relate to passing a test or having a success to move on to the next level. Although it is important for the students to demonstrate the knowledge they have gained from classroom instruction, this kind of response from the students leads to think that learning is for something outside. Learning outside means that students basically relate the grades that they receive to the real achievement they have had during a particular period of time. To bring the focus back to learning not for grades or for receiving a piece of sheet that will guarantee their success or achievement or not just for taking a test to pass to the next grade or level, formative assessment needs to be part of the instructional process.

1.2 Purpose of the study

English as a foreign language is extensively taught at primary, secondary and higher educational institutions of Armenia. Particularly, more than 3000 students study English as their first or second profession at many universities of Armenia, e.g. Yerevan State University, Yerevan State Linguistic University after V. Brusov, Armenian State Pedagogical University, etc. All the universities give students degree majoring in teaching English as a Foreign Language, which means that the students graduating from the universities are supposed to be future educators. The various subjects students pass during their academic years relate to those of teaching. One important aspect of teaching is assessment which has become a new subject in the Armenian educational system and that still experiences vague impressions among the students.

Though it is gradually developing in the educational structure of Armenia, people sometimes misunderstand some ideas and concepts of it which might lead to the mistaken ways of its implementation. We often think about assessment in terms of beings marked out or branded by results that stick with us after our course or program has finished. And sometimes when people think of assessment they mean summative assessment that is end-of-term exams, final projects and practical demonstrations, final presentations, and others.

As far as the summative assessment is concerned, it has been implemented in the country during many years and it is still in process. Through summative assessment educators summarize the development of learners at a particular time. After a period of work, e.g. a unit for two weeks, the learner sits for a test and then the teacher marks the test and assigns a score. The test aims to summarize learning up to that point. Besides summative assessment there is also one way of assessment which is intended to help instructors identify material which needs to be clarified or re-taught and should not be used to evaluate or grade students.

Considering the issue of formative assessment as a way of implementation in Armenian settings and taking the fact that summative assessment is widely used in most Armenian educational programs, it seems necessary to conduct research on the issue and to find out the results. Moreover, so far no research has been conducted in the field of formative assessment in an EFL Armenian setting. Therefore, it seems reasonable to conduct a research and respond to this lack. However, in order to investigate this phenomenon, a solid background with supporting data needs to be established. Hence, the first purpose of the study is to examine the relationship between the implementation of formative assessment and the students' achievement. The second purpose is to explore the students' attitude towards the implementation of formative assessment. Thus, this MA thesis will examine the following research questions:

Is there a relationship between the implementation of formative assessment and students' achievement in language learning in an Armenian setting?

What are the students' attitudes towards formative assessment?

Before stepping forward, let's at an outset, define what each of these terms mean. According to McManus, S. (2006, p. 3), "formative assessment is a process used by teachers and students during instruction that provides feedback to adjust ongoing teaching and learning to improve students' achievement of intended instructional outcomes". Formative assessment is delineated as assessment that is realized during instructional processes for the purpose of improving teaching or learning (Shepard, Hammerness, Darling-Hammond, & Rust, 2005). Black and Wiliam (1998b) confirm this fact stating that formative assessment include all activities that teachers and students carry out aiming to get information that can be used to alter teaching and learning to meet student needs.

As far as the achievement is concerned, it is the action or process of accomplishing something or the results gained by effort. Besides, it can be defined as the quality and the quantity of the students' work.

The setting for the study to be carried out was in the Experimental English Classes (EEC) at the American University of Armenia (AUA). The EEC courses are organized for students' age ranging from 5-17 of different levels. The purpose of the courses is to familiarize students with the communicative language teaching. The courses cover all the skills of language. Before taking the course, each student has to take a proficiency test in order for the authorities to decide the level of the course he or she is to take. At the end of each term, each student takes a final test that certifies his or her achievement. One term of the course lasts 10 weeks. Each week two sessions are conducted lasting one hour. Each class consists of approximately 15-20 students of different ages but the same level.

In order to be able to evaluate the students' attitudes, let us begin with the definition of the term 'attitude'. There are many definitions of the term. Among them Triandis (1971) accepts

that it is a manner of consistency toward an object. To Brown (2001), attitude is characterized by a large proportion of emotional involvement such as feelings, self, relationships in community (p 61). Brown (1994), in his work ‘Principles of Language Learning and Teaching’ adds: “Attitudes, like all aspects of the development of cognition and affect in human beings, develop early in childhood and are the result of parents’ and peers’ attitudes, contact with people who are different in any number of ways, and interacting affective factors in t human experience” (p 168). Here it seems clear that there are many stimulants lead to positive or negative attitude of an individual. Chamber (1999) asserts that learning occurs more easily, when the learner has a positive attitude towards the language and learning. To sum up, attitude refers to our feelings and shapes our behaviors towards learning.

The null hypothesis is:

There is no relationship between the implementation of formative assessment and students’ achievement in language learning in an Armenian setting.

The continuation of this thesis is the review of related literature which presents recent research related to formative assessment focusing on the strategies used by teachers during the implementation of formative assessment and also on the feedback that forms a crucial part in this type of assessment.

1.3. Significance of the study

The study can be significant for several reasons:

- It aims at finding out whether the implementation of formative assessment is effective at the Experimental English Classes (EEC) at the American University of Armenia (AUA).

- It aims at clarifying the importance of formative assessment in the classroom settings.
- It aims at finding out the effect of using this method in an Armenian setting in comparison with the traditional method.
- It aims at finding out whether feedback given in the instruction of formative assessment can be useful.
- It also aims at presenting the students' performance after the implementation of formative assessment through achievement tests.

Chapter 2: Review of the Related Literature

2. 1. Introduction

Students frequently echo frustration and disturbance when they are required to undergo regular assessment in order to demonstrate mastery of content or competency to pass to the next level of instruction (Stiggins, 2001). Assessment experts try to reveal the reasons of specific assessment practices as well as the role of assessment in society. In order to understand the main purpose of assessment, at the outset, let's find out what assessment means, where it comes from and what its purpose is. Assessment derives from the word "assess" which according to Murphy (1975), means "to sit beside" or "to assist the judge" (p.27). "To sit beside", according to James (1994), gives us information about a particular learner about his/her position among a set of people. Hence conceptually, assessment means "sitting beside a growing child with the purpose to observe, document, and describe qualitatively and quantitatively, his cognitive, affective, and psychomotor behavior, and using the results as feedback to ensure valid identification of potential as well as maximum desirable growth and development on which terminal judgment will later be made" (Nenty 1997a,) in (Nenty, 1999, p 1).

Besides its theoretical definitions, several operational definitions have been given to the word "assessment" (Wood, 1984; Izard, 1992; Title, Hecht and More, 1993; Messick, 1994). The word assessment may refer to in a variety of settings; business, health, education, psychology, etc. When it applies to education, it is the process of finding out "what knowledge, skills, habits, attitudes, or what behavior a learner does or does not have, acquire, or develop, before, during, and at the end of an instruction, a period of instructions, or a course of study" (Nenty, 1997b) (Nenty, 1999, p 1). According to McNamara (2004), assessment is an information-gathering activity which aims to gain insights into learners' level of knowledge or ability. Assessment includes observing, interviewing, professional experience/judging which involves using

questionnaires, classroom questioning, project assignment, class or seat-work, home work assignment, classroom testing, measuring; examination, etc. It is used to find out how much changes the child achieved in behavior or learning, and whether this is enough to guarantee his being passed to the next class or his being awarded a certificate. Assessment therefore provides valid evidence of teaching and learning with which to inform the learners, the teacher, the parents, and the administrators; to facilitate improvement in teaching as well as in learning and to certify that some required level of learning has been reached or has taken place (Izard, 1992; Nenty, 1997a).

Brown (2004), states that assessment is a popular and sometimes misunderstood term in current educational practice. He asserts that sometimes people refer to the terms "testing" and "assessment" as synonymous terms, but they are not. Assessment, according to Clapham (2000), "is used as an umbrella term to cover all methods of testing and assessment and as a term to distinguish "alternative assessment" from "testing" (p. 155). According to Kizlik (2010), assessment is an umbrella term that includes testing. A test is a special form of assessment. Tests are kinds of assessments that are made under particular circumstances for the purpose to be administered. In other words, all tests are assessments, but not all assessments are tests. "Tests are prepared administrative procedures that occur at identifiable times in a curriculum when learners muster all their faculties to offer weak performance knowing that their responses are being measured and evaluated" (Brown, 2004, p. 4). Assessment, on the other hand, is an ongoing process, that covers a much broader field, for instance, when a student answers a question, makes comments, expresses opinions to form new words or structures, the teacher subconsciously assesses the student's performance (Brown, 2004).

Sometimes, there is confusion between the term "assessment" and "evaluation" as well. A distinction between them is also made by different researchers. As it was noted above,

assessment is considered as a means of gathering information and making relevant inferences about language learners' knowledge and ability to use the language. However, evaluation mainly deals with collecting information and making judgments about the quality of the whole program (Kizlik, 2010). Assessment in its turn is divided into summative and formative assessments. Now let's apply to each of them.

2.2. Formative vs. Summative Assessment

Before starting the discussion about formative and summative assessments, let's, at the outset, talk about their history and chronology in order to find out where they come from and how they have been developed and changed through years. Formative and summative assessments have a long history. It was Michael Scriven (1967) who invented the terms formative and summative assessments. By inventing the two types of assessment, he aimed to emphasize their differences both in terms of the goals of the information they seek and how the information is used. Just a year later Benjamin Bloom (1968) made formative assessment a basis of Learning for Mastery and he stated that formative assessment might play a crucial role in the improvement of a curriculum.

In 1967, Michael Scriven identified the differences between formative and summative assessments (Marzano, 2006). The difference between the two types of assessments is that classroom formative assessments occur while content is being taught and learned and should continue throughout the period of learning and are not meant to assign grades. Its primary objective is to inform the teacher of what his or her students know or do not know. In other words through formative assessment results the teacher can be aware of what has been taught well and not so well (Nichols *et al*, 2008). The purpose of formative assessment is to enhance learning and not to allocate grades. In other words formative assessment refers to frequent,

interactive assessments of student progress and understanding to identify learning needs and change teaching instruction appropriately. The goal of formative assessment is to improve; the goal of summative assessment is to prove. Yorke (2003) defines summative assessment as an evaluation where students meet the course objectives through a midterm or final examination. Ebel and Frisbie (1991) defined summative assessment as assessment conducted at the end of instruction for the teachers to decide whether a particular student is guaranteed to pass to the next level of instruction. For more clear understanding Black and William (1998) present the differences and goals of formative and summative assessment in shape of a chart provided by Black & William, (1998) (Table 1).

Table 1: The differences and goals of formative and summative assessment

<i>Formative assessment</i>	<i>Summative assessment</i>
<p>➤ The goal of formative assessment is to gather feedback that can be used by the instructor and the students to guide improvements in the ongoing teaching and learning context. These are low stakes assessments for students and instructors.</p> <p>Examples:</p> <ul style="list-style-type: none"> • Asking students to submit one or two sentences identifying the main point of a lecture • Have students submit an outline for a paper. • Early course evaluations 	<p>➤ The goal of summative assessment is to measure the level of success or proficiency that has been obtained at the end of an instructional unit, by comparing it against some standard or benchmark.</p> <p>Examples:</p> <ul style="list-style-type: none"> • Assigning a grade to a final exam • Critique of an article • Final project

Formative assessment evaluates students' academic progress, delivers information *during* the instructional process, *before* the summative assessment which aims to summarize how much learning has occurred at a point in time attempting to measure the level of students, school, or program success by means of final exams, general proficiency exams, etc.(Brown, 2004). It means that formative assessment aims to determine whether the learning process is going on as it was planned. Its main purpose is to determine the learner difficulties and enhance teacher effectiveness, and is always administered during the teaching process.

2.3. The purpose and benefits of Formative Assessment

Formative assessment has become recognized recently as an important topic (Shepard, 2001; Stiggins, 2001). The concept of formative assessment was first used by Scriven (1967) and then taken up by Bloom *et al.* (1971). The topic of formative assessment has been widely reviewed by Black and William (1998). According to Black and William (1998) “assessment for learning is any assessment for which the first priority in its design and practice is to serve the purpose of promoting students' learning. It thus differs from summative assessment which is designed to serve the purposes of accountability, or of ranking, or of certifying competence (p 2). This was the main purpose of Scriven (1967) who tried to distinguish formative assessment from summative, stating the former occurred while learning was still occurring or *forming* and the other occurred at the end of learning. Before giving definitions of formative assessment or going deep into it, let's, at the outset, find out why it is called formative. There exist ample points of view and assumptions about it. Some authors like Airasian, 2001; Black & Wiliam, 1998, suggest that this kind of assessment is called formative as it provides a chance for both teacher and the student to *form* behavior. Here the authors highlight the fact of formation of learning and the formation of behaviors or strategies that contribute learning. Typically, “assessments

become formative when the information is used to adapt teaching and learning to meet students' needs" (Boston, 2002, p 2).

Since the beginning of the 1990s "formative assessment has enjoyed considerable attention, especially in schools. It was also called Assessment for Learning following Caroline Gipps's (1994) distinction from assessment *of* learning: summative assessment" (John Pryor and Barbara Crossouard, 2008, p 2) and the two are now considered "conceptually identical" (Threlfall, 2005). Since then, many institutions, especially in the UK, educators have attempted to make parallel formative assessment with contemporary psychological theories of learning (e.g. Boud, 1995; Gipps *et al.*, 1995; Black *et al.*, 2003; Hall & Burke, 2004; Eraut, 2006; Nichol & MacFarlane-Dick, 2006). More precisely, in 1997 the influential researchers The Assessment Reform Group (ARG) at the University of Cambridge (England) ordered Black and Wiliam (1998a, b) to deal with the issue of formative assessment. In this review Black & Wiliam (1998b, p.2) defined formative assessment as 'all those activities undertaken by teachers and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities and when the evidence is actually used to adapt the teaching to meet student needs'.

According to McManus, S. (2006, p. 3), "formative assessment is a process used by teachers and students during instruction that provides feedback to adjust ongoing teaching and learning to improve students' achievement of intended instructional outcomes". Formative assessment is delineated as assessment that is realized during instructional processes for the purpose of improving teaching or learning (Shepard, Hammerness, Darling-Hammond, & Rust, 2005). Black and Wiliam (1998b) confirm this fact stating that formative assessment include all activities that teachers and students carry out aiming to get information that can be used to alter teaching and learning to meet student needs. Under this definition, assessment includes such activities like teacher observation, classroom discussion, and analysis of student work, including homework and tests.

Kahl (2005a) argued that formative assessment is a “midstream” tool that teachers use “to measure student grasp of the specific topics and skills they are teaching” (p. 38). Many researchers and advocates for formative assessment state that its primary advantage is that it allows students to organize and improve their own learning (e.g., Stiggins, 2002).

Throughout the history of education, there have been many disagreements on the nature of formative assessment (Cizek, 2010). Some consider formative assessment as testing and giving feedback, while others claim that formative assessment exists while instruction proceeds and it can become a reason to alter the future methods and styles of teaching and learning (William, 2010). According to William and Leahy (2007), formative assessment “needs to contain an implicit or explicit recipe for future action.” (p.13). Assessments that are formative give information about the real and appropriate recognition of where the student lacks understanding of a certain concept or phenomenon, measure and indicate the progress, provide specific feedback, and include instructional correctives, different from previous instruction, that will guarantee successful achievement of learning targets (Brookhart, 2007; Guskey, 2010). In other words, assessment that is formative provides individualized feedback and information about what students need to do to improve performance.

Formative assessment is thought to have “intrinsic acceptability” (Black and William, 2003, p.634) to teachers which means that they think it is another effective way of teaching and it really has its outcomes but it faces some problems in terms of its implementations at the secondary level (Hayward & Hedge, 2005; Smith & Gorard, 2005). There is little research on the extent to which these aspects of formative assessment are used by secondary teachers. The Centre for Educational Research and Innovation (CERI, 2005) notes that “powerful bureaucratic constraints” (p.28) limit the implementation of formative assessment in secondary schools, even though its use with teenagers is particularly acceptable and justified. In their international case studies, CERI (2005) concludes that the benefits of formative assessment are more important

than the obstacles of its implementation. Black and Wiliam (1998a) conducted an extensive research review of 250 journal articles and book chapters to find out whether formative assessment raises academic standards in the classroom. They concluded that if teachers try to reinforce the implementation of formative assessment, they will produce more powerful effect on student learning and the students and the teaching process will have significant gains. This can be proved by comparing the average improvements in the test scores of the students involved in the formative assessment process with the range of scores found for the students on the same tests who did not undergo such process. Black and Wiliam report that studies of formative assessment show an effect size on standardized tests of between 0.4 and 0.7, which is larger than most known educational interventions. The authors state that formative assessment especially helps the low-achieving students. (Black and Wiliam,1998b). First of all they indicate that formative assessment produces non-threatening results. What did they mean by saying non-threatening results? First of all, formative assessments are scored but not graded. This fact is mostly accepted by the low-achieving students as sometimes seeing the low grade under their work they become frustrated which prevents their ambitions to reach their goals using their potentials. Students mark their own work and are encouraged to raise questions about the assessment and the material covered during the session. Besides, formative assessment provides direct and immediate feedback. Results of formative assessments are produced “on the spot;” teachers and students get them immediately. Teachers separate the mistakes and error made by the students, discusses them, gives feedback, and provides rules and examples of the wrong points, while students learn how well they have done. In addition to that, formative assessment designates ways to improve; it means that it provides a basis for the teacher to re-visit topics in the unit if necessary and for giving students additional experiences in areas where they performed less well.

The question is whether accordingly the students approaches change towards their learning. Research shows that students' assessment preferences and their approaches to learning are correlated with differences in learning strategies (Birenbaum, 1994, 1997; Birenbaum and Feldman, 1998). Entwistle and Tait (1990) reported that students who described themselves as 'surface learners' preferred teaching and assessment procedures which supported this learning approach, whereas students describing themselves as 'deep learners' preferred courses which were intellectually challenging and assessment procedures allowing them to demonstrate their understanding.

Sadler (1989) sums up some potential uses of formative assessment:

- ✓ To facilitate learning
- ✓ To see whether learning has taken place
- ✓ To provide feedback to teachers of how the learners are progressing, clarifying for the teacher what can be done to improve, extend or enhance learning
- ✓ To provide feedback to learners concerning their own progress, clarifying for the learner what he or she needs to do to improve, extend or enhance learning
- ✓ To diagnose learners' needs and barriers to learning and help to get any necessary changes to the course or program of the study.

2.4 Elements that formative assessment involves in practice: Formative Assessment Techniques

In the cognitive theory of classroom management, students learn best when they are provided meaningful feedback. This feedback is cognitively processed and integrated into the larger understanding of a concept. Formative assessment encourages this process by focusing intervention strategies on ensuring that students get high quality feedback in a relatively short amount of time. There exist a number of ways elements, strategies and techniques for effective implementation of formative assessment.

First of all, it should be mentioned that formative assessment differs from other assessments in a way that during its implementation teachers change the “culture” of their classrooms (OECD, 2005). What does it mean? It means that the teacher should create such an atmosphere that the students would not be stressed and would take risks to make mistakes and develop self-confidence in the classroom.

Formative assessment techniques are quick and easy ways for the teachers to understand how and what students are learning; first find out the gap and what kind of adjustments need to be made in the course to meet the learning goals.

These techniques might be applied at regular periods throughout the course: after each class or each unit, weekly, etc. Since no responses will be graded, and as anonymous feedback is welcomed and more honest towards the students, learning process will be more effective and smooth.

The following techniques are usually implemented in the courses using formative assessment.

One-minute paper:

This technique of formative assessment requires the teacher to assign at the end of each class to respond on a half-sheet of paper, to the following questions: “What was the most important thing

you learned in this class today? What was the most confusing idea presented today? What question remains unanswered (or is unclear) from class today?" (Angelo and Cross, 1993, p 2)

Then the teacher should collect these papers and review them after class to find out what the students did and did not understand about the lesson. Then he should collect the papers and use for promoting discussion, identifying misconceptions, or confusion (Ferlazzo, 2010) and consequently, respond to the problematic questions during the next class session. This gives the teacher a sense of understanding what the students learned and what they lack.

One-sentence summary:

This technique of formative assessment requires the teacher to assign the students to "summarize the "who, what, where, when and why" of something (an event, a chemical process, a story plot, etc.) in one, clear grammatical sentence. Also he/she can ask students to summarize the main ideas they've taken away from a lecture, discussion, or assigned reading (Boston, Carol, 2002). As Angelo and Cross (1993) assert this gives the teacher a sense of their ability to analyze and synthesize information in a concise format" (1993, p 2).

Frequent and short tests vs. infrequent and long tests:

Formative assessment challenges to use frequent and short tests instead of infrequent and long tests during classroom instruction. Black and William (1998) state that first of all infrequent and short tests release students of the stress they usually have during the tests, besides they prepare the student for the final test without feeling stressed or embarrassed. By writing the tests during a session a student treats it as a usual classroom activity which does not intend to decide his/her accomplishments or does not serve as a record of achievement. As a proof for this example was the research which was done by Martinez and Martinez (1992) which supports the

conclusion that formative assessment does improve student achievement. Besides they state that frequent short tests are of more effect than long and infrequent ones. Martinez and Martinez (1992) utilized an experimental design in which two groups were taught by a teacher and the remaining two were taught by another teacher. Each teacher taught one class in which the students took only one test per chapter, and the other class took three tests per chapter. The total sample size consisted of 120 college algebra students, which resulted in small numbers of participants in each of the four sub-groups (less than 30 students each). Results indicated that the only statistically significant differences in achievement were seen between the control group (one test per chapter) and the treatment group (three tests per chapter) in the novice teacher group. The authors concluded that frequent assessment is more important in the classrooms. Moreover, this study looked at the importance of frequency of assessment. However, no information was given as to the nature of the assessment or of feedback provided from the assessment.

Office Hours:

“Use your office hours not only to answering student questions but also as a way to learn about your students. Find out what they do and do not understand. Experiment with different ways to explain course material and see which ways make sense to students”. (Angelo and Cross, 1993, p 1)

Boston, Carol (2002) suggests that teachers might also assess students' understanding in the following ways:

- “Have students write their understanding of vocabulary or concepts before and after instruction”.
- Interview students individually or in groups about their thinking as they solve problems.

Portfolios:

Portfolios are also used as a tool for teachers to provide a better understanding of the materials. They serve as a means for the teachers to check student progress. “They provide an opportunity for written dialogues between teacher and student, and also provide opportunities for students to reflect on their learning process” (OECD, 2005, p 3). A portfolio is a purposeful collection of significant work, carefully selected, dated and presented to tell the story of a student’s achievement or growth in well-defined areas of performance, such as reading, writing, math, etc. A portfolio usually includes personal reflections where the student explains why each piece was chosen and what it shows about his/her growing skills and abilities.

Self and peer assessment:

Black and Wiliam (1998b) emphasize peer- and self-assessment as key techniques: “If formative assessment is to be productive, pupils should be trained in self-assessment so that they can understand the main purposes of learning and thereby grasp what is to be achieved”.

Self assessment

Students can become better language learners when they engage in deliberate thought about what they are learning and how they are learning it. In this kind of reflection, students step back from the learning process to think about their language learning strategies and their progress as language learners. Such self assessment encourages students to become independent learners and can increase their motivation. Besides, it involves children in their own learning and assessment, allows the pupils to realize their own learning needs, gives the pupils the information they require to negotiate their learning target and provides the teacher with more assessment information - the pupil’s perspective (OECD, 2005).

How should this strategy be implemented in a classroom setting to meet the needs?

The teachers should create a poster of questions:

- ✓ What did you find easy?
- ✓ What did you find difficult and what helped to move on?
- ✓ What do you need more help with?
- ✓ What are you most pleased with?
- ✓ Have you learnt anything new and what?
- ✓ How would you change this activity?
- ✓ Do you have any question?

Peer Assessment

One of the ways in which students internalize the characteristics of quality work is by evaluating the work of their peers. For peer evaluation to work effectively, the learning environment in the classroom must be supportive. Students must feel comfortable and trust one another in order to provide honest and constructive feedback. Instructors who use group work and peer assessment frequently can help students to develop trust by forming them into small groups early in the semester and having them work in the same groups throughout the term. This allows them to become more comfortable with each other and leads to better peer feedback.

Peer assessment – some practical tips

- ✓ Make sure that students understand the assessment criteria and the characteristics of work at different levels.
- ✓ Provide opportunities for useful quality feedback alongside grades.
- ✓ Use anonymous feedback to begin with (to overcome any feelings of ‘betrayal’ among friends).

- ✓ Try using peer assessment choosing random students (after collecting the works of students, muddle them and give a random student to assess).

Some examples of peer/self assessment activities are:

- ✓ commenting on final or draft essays/reports
- ✓ anonymously or publicly grading presentations/performance
- ✓ proposing a grade for their own work, after seeing/assessing others' work
- ✓ discussing and suggesting improvements to others' work
- ✓ reflecting on improvements they could implement themselves
- ✓ Discussing in groups before collectively providing a grade and feedback.

Benefits of Peer Assessment

- ✓ Students often respond more positively to a peer than a teacher
- ✓ When writing for peers they improve specifics such as handwriting
- ✓ The assessor gains as well as the assessed
- ✓ Teacher can stand back, observe and make focused interventions
- ✓ Pupils take more responsibility for their own learning

Feedback

One of the most important functions of formative assessment is providing students with 'continuous feedback', "meaning that opportunities for feedback should occur continuously, but not intrusively, as a part of instruction" (Bransford *et al.* 2000, p. 140). Bell and Cowie (2001), Bransford *et al.* (2000, pp. 140–141), Perrenoud (1998), Sadler (1989), and Williams (2004) observed that "feedback is as an essential component of successful formative assessment". And as Black and Wiliam (1998) defined formative assessment is "all those activities undertaken by

teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged” (p. 10). Sadler (1998) notes that formative assessment specifically intends to generate feedback on performance to improve and speed up learning. Bell and Cowie (2001) concluded that formative assessment is increasingly being used to refer only to assessment that provides feedback to students about learning.

Feedback in formative assessment can become a means to discover the student weaknesses that need reinforcement. Feedback is considered as an essential component of the formative assessment interaction, which aims to support learning (Sadler, 1989; Perrenoud, 1998). Teachers should make use of formative assessment to give students feedback at an appropriate point in the learning process (Brown and Knight 1994, p.2) as formative assessment will not help students much if teachers do not devote time for feedback. Wiliam and Black (1996) state that feedback of formative assessment should tell learners that there is a gap that needs to be filled in, what must be done, thus improving effective learning. This is the main point when formative feedback differs from feedback in general. In traditional classroom teaching, “teacher feedback to learners is limited” (Bransford *et al.* 2000, pp. 140–141). Limited means that teachers give just transcripts or scores to the assignments done by the students. After grades are given, students move on to the next topic and work for another grade. In other words, learning is driven by reading, writing, completing activities and making grades. “Such learning is not effective. Instead, the truly meaningful feedback that teachers give to learners should derive from formative assessment. By using meaningful feedback, students can improve weaknesses in learning and thinking, increase and transfer learning, and value opportunities to revise (Barron *et al.* 1998; Black & Wiliam 1998; Bransford *et al.* 2000, p. 141). Feedback helps learners become aware of any gaps that exist between their desired goal and their current knowledge, understanding, or skill and guides them through actions necessary to obtain the goal (Ramaprasad, 1983; Sadler, 1989). The most helpful type of feedback on tests and homework

provides specific comments about errors and specific suggestions for improvement and encourages students to focus their attention thoughtfully on the task rather than on simply getting the right answer (Bangert-Drowns, Kulick, & Morgan, 1991; Elawar & Corno, 1985). This type of feedback may be particularly helpful to lower achieving students because it emphasizes that students can improve as a result of effort rather than be doomed to low achievement due to some presumed lack of innate ability. Formative assessment helps to support the expectation that all children can learn to high levels and tries to take out the fear and stress that force students to become discouraged and unwilling to invest in further learning (Ames, 1992; Vispoel & Austin, 1995). Formative feedback informs pupils of strengths and weaknesses, provides strategies for improvement, and improves the achievement of all students. However, as Buchanan (2000) notes, formative assessment feedback is not an easy task for the teachers to provide, because they face large numbers of students, lengthy pieces of work, or practical constraints such as time and workload.

Henly (2003) and Khan *et al.* (2001) implemented the design of ‘repeat the test’ and ‘timely feedback’ in their web-based formative assessment and found that it had significant effects on the learning of the students. Learners discovered what the gap was and were motivated to make clear the concepts to master content. Buchanan (1998, 2000) noted that while implementing this strategy, the correct responses were not given to the students and they had to rethink and to be engaged more with the course materials. He suggested that these strategies will help participants know which items need reinforcement, and enable learners to concentrate on learning materials and steadily master course concepts.

One of the most important aspects of formative assessment is the students’ discussions with their teachers concerning their mistakes, problems, improvements and achievements. Besides, the reflections and the feedback on both the students and their peers’ works can be an effective means to master the content as well as to improve their learning process (APMG, 2005-

2008). This may contain a variety of insightful tasks, for example, after reporting presentations to the class, students are expected to assess the quality of the presentations of their peers, and receive the assessments of other groups in the class regarding their presentation. The process is fulfilled by a discussion between the teacher and the students and class feedback (APMG, 2005 – 2008). Thus, it can be concluded that managing feedback appropriately is a fundamental determiner of the classroom atmosphere.

One of the most important questions is how teachers should frame feedback comments, what kind of discourse should be used and how many comments are appropriate. (Laurillard) (2002) argues that one of the best ways of providing the students with a good way of feedback is to conceptualize feedback more as *dialogue* rather than as information transmission. Feedback as dialogue means that the student not only receives initial feedback information but also has the opportunity to engage the teacher in discussion about that feedback. It should include both spoken and written dialogue.

So, what should an instructor do to ease his work and at the same time provide meaningful and appropriate feedback? Let's see what are the essential features of useful feedback? It should:

- ✓ Take the form of a dialogue between the instructor and the student
- ✓ Be in the form of comments and not grades or marks
- ✓ Invite and encourage the pupils to think for themselves
- ✓ Be prompt and regular
- ✓ Include both spoken and written dialogue
- ✓ Recognize effort as well as quality

Freeman and Lewis (1998) argue discussions with the teacher help students to develop their understanding of expectations and standards, to check out and correct misunderstandings and to get an immediate response to difficulties. Lunsford (1997) also advises that the comments should always be written in a non-authoritative tone and where possible they should offer corrective advice (both about the writing process as well as about content) instead of just information about strengths and weaknesses. Besides, teachers should not only correct students' errors, mistakes, give suggestions of improvement but they should also provide positive praise and encouragement which promotes high expectations of progress. Moreover, teachers could and should invite the children to find and correct their own mistakes. Considering this issue, an example from the current study is provided when often students received feedback like this: "Armen do you remember we have discussed this issue during the class? There is a mistake here. I am sure you can find it and correct. And please bring another example that can show I really know it".

Unfortunately, with large class sizes it can be difficult for the teacher to engage in dialogue with students. Nonetheless, there are ways that teachers might increase feedback dialogue even in these situations. One approach is to structure small group break-out discussions of feedback in class after students have received written comments on their individual assignments. Another approach is to use classroom technologies. These technologies help gather student responses to in-class questions (often multiple-choice questions) using handset devices. This collated feedback has been used as a trigger for peer discussion (e.g. 'convince your neighbor that you have the right answer') and teacher-managed discussion in large classes (e.g. Nicol and Boyle, 2003; Boyle and Nicol, 2003).

Studies also found out that another source of providing feedback to students is through their peers. Boyle and Nicol (2003), note that peer dialogue enhances in students a sense of self-control over learning in a variety of ways. Firstly, students who have just learned something are

often better able than teachers to explain it to their classmates in a language and in a way that is accessible. Secondly, peer discussion exposes students to alternative perspectives on problems and to alternative devices and strategies. Alternative perspectives enable students to revise or reject their initial hypothesis and construct new knowledge and meaning through negotiation. Thirdly, peer discussion can be motivational in that it encourages students to persist. Finally, it is sometimes easier for students to accept critiques of their work from peers rather than tutors.

Research in school settings has shown that frequent high stakes assessment (where marks or grades are given) has a 'negative impact on motivation for learning that militates against preparation for lifelong learning' (Harlen & Crick, 2003). Feedback given as grades has also been shown to have especially negative effects on the self-esteem of low ability students (Craven, Marsh & Debus, 1991). Dweck (1999) argues that such assessments encourage students to focus on performance goals (passing the test, looking good) rather than learning goals (mastering the subject). For just having the aim to pass the test or to get the same grade as their peer's student intend to copy or cheat without realizing the real sense of learning. In one study, Butler (1988) demonstrated that feedback comments alone increased students' interest in learning when compared with two other controlled situations, for example one where only marks were given and the other where students were given feedback and marks. Butler argued that students paid less attention to the comments when given marks and consequently did not try to use the comments to make improvements. This phenomenon is also commonly reported by academics in higher education. Butler (1987) has also argued that grading student performance has less effect than feedback comments because it leads students to compare themselves against others (ego-involvement).

Chapter 3: Methodology

This chapter describes the process of conducting the current research attempted to explore students' attitude towards formative assessment, what feedback they can get through formative assessment, and finally to find out whether there is an effect of using this method in an Armenian setting in comparison with the traditional method. This chapter consists of three sections: *participants, setting, instrumentation, procedure and analysis*.

3.1. Research Design

The following research questions are addressed in the study:

Is there a relationship between the implementation of formative assessment and students' achievement in language learning in an Armenian setting?

What are the students' attitudes towards formative assessment?

In order to seek answers from the questions guiding this study mixed methods; quantitative and qualitative were implemented. Particularly, the quasi experimental design was used to answer the two research questions.

3.2. Participants and Setting

The participants for the study were 33 students studying at the American University of Armenia (AUA) in the Experimental English Classes (EEC). EEC classes are organized by one of the departments of AUA, the Department of English Programs (DEP). AUA is a U.S. accredited graduate institution affiliated with the University of California. The Department of English Programs at AUA offers a professional M.A. program in Teaching English as a Foreign Language (MA TEFL). The EEC courses are organized for students' age ranging from 5-17 of

different levels. The purpose of the courses is to familiarize students with the communicative language teaching. The courses cover all the skills of language which is a motivating factor for the people to attend. Besides, the classes are conducted with appropriate facilities such as equipments or materials which serve as stimuli for the students' enhancement of the learning. Before taking the course, each student has to take a proficiency test in order for the authorities to decide the level of the course he or she is to take. At the end of each term, each student takes a final test that certifies his or her achievement. One term of the course lasts 10 weeks. Each week two sessions are conducted lasting one hour. Each class consists of approximately 15-20 students of different ages but the same level.

The age of participants for my study ranges from 10-15. The level of participants is intermediate. The classes consist of both males and females. Most of the students have been exposed to English for at least 4 years. All of them are studying at Armenian schools. The classes are held mostly in the afternoon because most of the students are in the second shift in their schools.

3.3. Materials

The textbook used for the class for this research was 'New Parade 5' by Herrera, M and Zanetta, T. (2000). New Parade is a set of seven - level, communicative language program that features rhymes, songs, pair work, cooperative learning and hands on projects. (Herrera and Zanetta, 2000). Each set includes a textbook accompanied by a workbook. The textbook covers all four skills (listening, reading, writing and speaking), but special emphasis is put on maximizing opportunities for discussion and promoting the development of both linguistic and communication skills (Herrera and Zanetta, 2000).

The textbook consists of nine units, each unit focusing on one selected topic. The workbook consists of different tasks and activities. The workbook enables students to practice

materials for writing, grammar and language practice covered in the textbook. After each three units students take a test. During the ten weeks of this research, 3 units of this book were covered in class: units 1, 2 and 3.

It should be noted that the textbook served as the main source of the treatment and all the extra activities including the essays, short tests, and the power point presentation presented at the end of the term were based on the textbook. It should be noted that both the experimental and the comparison groups were taught by the same teacher.

3.3. Instrumentation

The data for the study were collected through the following instruments:

- The quantitative data were collected through:
 - ✓ a pre test
 - ✓ a post test
- The qualitative data were collected through:
 - ✓ a questionnaire
 - ✓ field notes

All the instruments were administered in English except the questionnaires. There was a need to develop them in both Armenian and English taking into account the fact that students might experience some difficulties in understanding of some concepts or specific terminology.

3.3.1. Tests

At the beginning of the study, the decision was to take the results of the achievement test for passing to the next level as the students' pre-test. Out of the 33 students participating in the study, some were newcomers meaning it was their first time studying at EEC and the others were

former students who were continuing their course. It happened so that some of the students had already taken the test and some were still to take. Later, in the process of checking the tests, it turned out that there were two different raters and this fact was not taken into account. Besides, the details and the circumstances of the test of the previous level were not available. This can be considered as a limitation of the study.

For the results of the study to be valid, it was decided to take the results of the first test and the final test of the same level and check the students' performance based on the two unit process. Both tests were measuring the same sets of skills; they involved listening, reading and writing exercises.

3.3.2. Questionnaires

The purpose of developing a questionnaire was to reveal the attitude of students towards the treatment process. In each group two different questionnaires were administered as each one experienced a different way of teaching.

As noted above, the questionnaires were developed both in the students' native language and in English. They were translated into Armenian to eliminate the possibility of misunderstanding. In order to make sure that the items were appropriate, before administering the questionnaires, each questionnaire was validated by the supervisor and the reader and piloted among the current MA students of DEP (Department of English Programs). When the necessary corrections were made the questionnaires were ready for distribution among the participants of the study.

The questionnaire developed for the experimental group consisted of 10 questions. It involved 8 closed-ended items in Likert four point Scale format (from Strongly Disagree to Strongly Agree) each question consisting of four items (Strongly disagree, Disagree, Agree,

Strongly agree) (Appendix 1). The other two items of the questionnaire were open ended, which aimed to elicit students' suggestions and comments about the quality of the course.

The questionnaire developed for the comparison group consisted of 7 questions. It involved 5 closed-ended items in Likert Scale format (from Strongly Disagree to Strongly Agree). Each question consisted of four items (Strongly disagree, Disagree, Agree, Strongly agree). (The questionnaire is included in the appendix 2). Again, the other two items of the questionnaire were open ended, which aimed to elicit students' suggestions and comments about the quality of the course.

Two questionnaires were distributed to explore the students; attitude toward the teaching process. The attitudinal questionnaires for students consisted of open-ended and closed-ended items in Likert scale format (from Strongly Disagree to Strongly Agree). As the treatments were different in both groups, each group had a separate questionnaire.

3.3.3. Filed notes

As the main purpose of the formative assessment is to elicit the problems and difficulties during the instruction to alter teaching to a better was, there was a need to keep field notes which could provide more valid information about the treatment and teaching process: each step designed by the teacher, her perceptions and reactions towards each step. Emerson (1995) defines field notes as accounts describing experiences and observations by the researcher during the experiment. The main objectives of field notes are to understand the true perspectives of the subject being studied. For this reason, after each class, some 10-15 minutes were devoted to the writing of field notes that described the essential points of the session. Further, at the end of each week they were summarized into a descriptive summary of the past two sessions of the experimental group. This treatment was applied only in the experimental group as the purpose was to clarify the outcomes of formative assessment only.

3.4. Procedure and Treatment

The two groups involved in the research used the same textbooks and had English classes for the same amount of time (1 hour per session, for 10 weeks). The research lasted for 10 weeks, starting from April 1 to June 16, 2010. Classes were conducted twice a week for one hour each (a total of 14 hours).

The textbook used in the study was “New Parade 5” by Herrera, M. and Zanetta, T. (2000). New Parade is a set of seven - level, communicative language program that features rhymes, songs, pair work, cooperative learning and hands on projects.

During the term the students were to examine three units of the textbook. After each unit they took a test; the first exam, the midterm and the final. All tests consisted of three sections: Listening, Reading and Writing.

Throughout the term the students wrote many essays based on the topics involved in their textbook. For instance, the first unit if the textbook was about family. After completing some parts of the unit, getting introduced to the main vocabulary and structures, they were asked to prepare a family tree and to write a detailed description of their family history. Both experimental and the comparison groups wrote the essay. For the experimental group, the essays were collected, checked, provided feedback and returned to the students. On the part of the teacher all the errors were taken out and brought to class. During the class, all the errors were written on the board and explained in detail. Saying in detail, means that if the error was concerning the verb tense, first it was corrected either by the teacher of by the students, then the rules on this tense were explained, some examples were brought, etc. An important issue need to be noticed here that it was announced whose error was taken out for the student not be offended or demotivated. Those students, who had too many mistakes, were asked to stay after the class for further consultation and explanation after which they had to rewrite the essay. For the

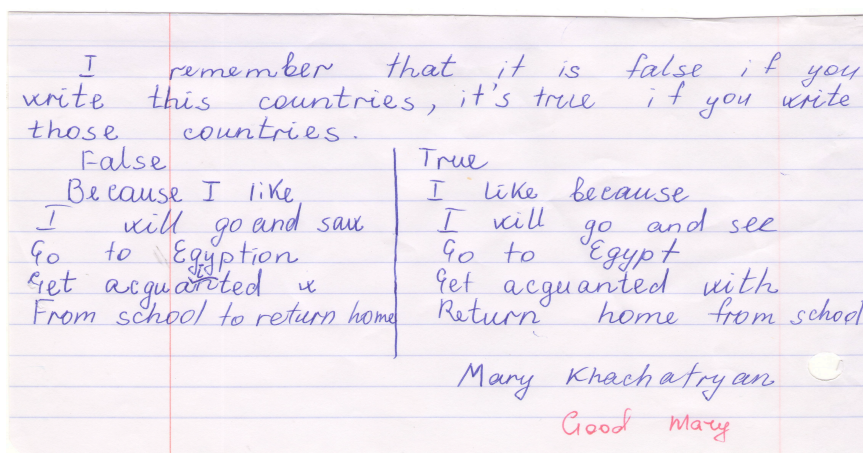
comparison group, all the essays were checked and graded. Some explanations regarding the errors were provided but not in detail.

After having passed a certain unit or session, at the end of a class the students were asked to write down on a sheet of paper what things they remembered from the class; what they understood the best and where they had challenges. The difficulties or issues that were not grasped easily by the students were explained later after collecting the sheets.

3.4.1. Students' Work Samples

During each session, one technique of the formative assessment was applied. It depended on the main purpose of the class, the activities that a particular session involved, etc. In the field notes which serve as a proof of the real procedure of the class are presented. Not all the classes are introduced in that section but a few of them as I thought there was no need to record all the classes. During this time the records are presented that reflect what students remember from a particular of previous class, what the difficulties and concerns are and finally what the things that are remembered well are. This technique of formative assessment was implemented quite often as it worked well for the students to recall issues and focus on the issues that are not understood properly yet.

One-minute paper:



Peer and self assessment

As peer and self assessments are essential parts of formative assessment, during the treatment students were asked to write down and assess their peers and themselves. They were to assess based on some aspects:

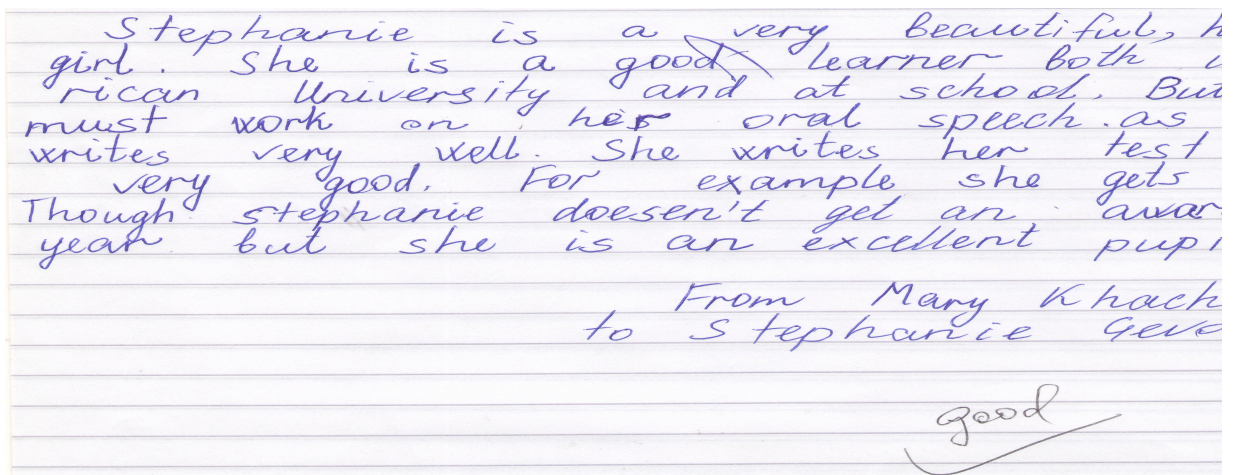
For self assessment:

- ✓ *What have I learnt during this period?*
- ✓ *What are the aspects that I remember very well?*
- ✓ *Where do I need some improvement?*

For peer-assessment:

- ✓ *What are the strengths and weaknesses of my peer?*
- ✓ *What he should do to improve?*
- ✓ *How much has he improved since the beginning of the course?*

Some examples of peer and self assessment are presented below:



Stephanie is a very beautiful, h
girl. She is a good learner both u
rican University and at school. But
must work on her oral speech. as
writes very well. She writes her test
very good. For example she gets
Though Stephanie doesn't get an award
year but she is an excellent pupil

From Mary Khack
to Stephanie Geve

good

Formative feedback

Feedback plays a crucial role during the implementation of formative assessment. Here feedback differs greatly from the traditional class. It means when a student made an error, it was written out, explained during the class period, after which some examples were brought by the students for better understanding and retention. And those students that had made too many mistakes or needed extra explanation and activities, were asked to stay longer after classes to discuss issues. To be honest, this idea was not impressive for the students as after the class they desired leave out.

Chapter Four: Results and Discussion

The present study was carried out to investigate to find out whether there is any relationship between the implementation of formative assessment and students' achievement in language learning. It was also aimed at determining what attitude students have towards this method. For the current study both quantitative and qualitative data were collected. The quantitative data were collected through pre and post tests. The qualitative data were collected through a questionnaire to both experimental and comparison groups. This chapter presents the results of the data analyzed both quantitatively and qualitatively.

Both quantitative and qualitative results seek to answer the following research question guiding the study:

Is there a relationship between the implementation of formative assessment and students' achievement in language learning in an Armenian setting?

What are students' attitudes towards the implementation of formative assessment?

4.1. Analysis of Quantitative Data

The quantitative data included one pretest and one posttest scores. The pretest and the post test results were analyzed quantitatively through the statistical package for social sciences (SPSS program). The quantitative analysis was obtained through two sets of scores from each group. As mentioned above because of the circumstances, the first exam of course was taken as the pre test and the final test was taken as a post test. The purpose was to see which group showed the highest achievement.

As it is stated in Hatch and Farhady (1995), all the variables of the experiment should be described and identified according to the type of relationship which is investigated. It should be noted that there was one independent and one dependent variable. The dependent variable was the scores of the first and the final test of both groups. The independent variable was the treatment.

4.1.1. Within group comparison

For comparing whether there has been a progress in students' performance, for each group first test results are compared with the posttest because, as mentioned above, there has been some incorrectness in the pre-test scores. It means we want to find out if there have been significant differences in the performance of each group with the respect to itself during two unit part of the three unit course. In this case samples are dependant and we have applied the parametric paired-samples t-tests and non-parametric Wilcoxon Signed Ranks tests. Both parametric and non-parametric versions were used for the findings to be more supportive. Both tests are bringing to the same results which mean that these results do not depend on type of probability distribution to be normal or not. Both tests are demonstrating significant differences between the pretest and the post test of the both groups in favor of the final tests.

Table 2: Paired Sample Test for the within group comparison for the experimental group

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair 1 Posttest (40 points total)	-2.5000 0	1.94454	.47162	-3.49979	-1.50021	-5.301	16	.000

a. group name = experimental

According to the table 2, for the experimental group the Paired-sample t-test statistic $t = -5.301$ and the probability $p = 0.000$. This value is substantially smaller than the specified alpha value of 0.05. Therefore we can conclude that there is significant difference between the pretest and the posttest in the experimental group.

For the results to be more valid and taking into account the major limitation of the study, effect size for the experimental group should be calculated. According to Pallant (2007), one of the ways to assess the importance of the finding is to calculate the effect size. "This is a set of statistics that indicates the relative magnitude of the differences between means" (Tabachnick & Fidell in Pallant, 2007). And although the results presented above tell us that the difference we obtained in the two sets of scores was unlikely to occur by chance, it does not tell us much about the magnitude of the intervention's effect. And one way to do so is to calculate the effect size. The effect size for Paired-samples t-test is calculated through Eta squared. Here

$$\text{Eta squared} = \frac{t^2}{(t^2) + N - 1}$$

So, from our example $t = -5.3$ and $N = 16$.

Hence, eta squared value = **0.63**

Taking into account the guidelines proposed by Cohen (1998) (0.01=small effect, 0.06=moderate effect, 0.14=large effect), it can be concluded that there was a great effect between the pretest and the post test of the experimental group.

Table 3: Paired Sample Test for the within group comparison for the comparison group

	Paired Differences					T	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair 1 Posttest(40 points total)	-5.05000	2.97630	.94119	-7.17911	-2.92089	-5.366	9	.000

a. groupname=comparison

As the table 3 reveals, for the comparison group the Paired-sample t-test statistic $t=-5.366$ and the probability $p=0.000$. This value is substantially small that the specified alpha value of 0.05. Therefore we can conclude that there is significant difference between the pretest and the posttest in the comparison group.

The same procedure for the comparison group was implemented. And for the results to be more valid and taking into account the major limitation of the study, effect size for the experimental group was also calculated. Following the formula of eta squared $\text{Eta squared} = \frac{t^2}{(t^2) + N-1}$, in this case $t=-5.37$ and $N=10$.

Hence, eta squared value= **0.76**

Taking into account the guidelines proposed by Cohen (1998) (0.01=small effect, 0.06=moderate effect, 0.14=large effect), it can be concluded that there was a great effect between the pretest and the post test of the comparison group.

To conclude with, it can be seen that the statistical analysis' results show that as a result of the treatment both the experimental and the comparison groups had achievement; it means that during the post test the groups performed better than during the pretest. However, considering the fact that during the first exam the experimental group performed better that the comparison

group and taking into account the effect size numbers, we can conclude that both groups did have achievement but the progress of the comparison groups was much better. Maybe if the circumstances of the present situation were not as follows, the statistics would reveal other results, but now we should state that according to the interpreted results our null hypothesis that there is no relationship between the implementation of formative assessment and the students' achievement is not rejected.

Wilcoxon Signed Ranks Tests

For the purpose of determining whether the treatment of formative assessment has had positive impact on students' achievement, a Wilcoxon Signed Ranks test was performed. The results can be seen in table 6 for the experimental group.

Table 4: Wilcoxon Signed Ranks Test for the within group comparison for the experimental group

Test Statistics^{b,c}	
	Posttest(40 points total)
Z	-3.370 ^a
Asymp. Sig. (2-tailed)	.001

As the table 4 shows, the non-parametric Wilcoxon test for the experimental group has test-statistic $Z=-3.370$ and $P=0.001$ which is less than 0.05. Therefore we can conclude that the two sets of scores are significantly different.

For the results to be more valid and again, taking into account the major limitation of the study, effect size for the experimental group should be calculated. According to Pallant (2007), one of the ways to assess the importance of the finding is to calculate the effect size. "This is a

set of statistics that indicates the relative magnitude of the differences between means” (Tabachnick & Fidell in Pallan, 2007). The effect size for this test to be calculated is as follows: we should divide the z value by the square root of N. In this situation, $Z=3.37$ and $N =17$. Therefore, $r=0.81$. Taking into account the guidelines proposed by Cohen (1998) (0.1=small effect, 0.3=moderate effect, 0.5=large effect), it can be concluded that there was a great effect between the pretest and the post test of the experimental group.

Table 5: Wilcoxon Signed Ranks Test for the within group comparison for the comparison group

	final exam(40 points total) – Newpre
Z	-2.805 ^a
Asymp. Sig. (2-tailed)	.005

In addition, for the comparison group, for the Wilcoxon tets, $Z=-2.805$ and $P=0.005$ which is less than 0.005. (See table 4.8.) Again, we can conclude that the two sets of scores are significantly different.

The same procedure should be implemented for the comparison group. Here again, for the results to be more valid and again, taking into account the major limitation of the study, effect size for the experimental group should be calculated. In this situation, $Z=2.8$ and $N =10$. Therefore, $r=1.61$. Taking into account the guidelines proposed by Cohen (1998) (0.1=small effect, 0.3=moderate effect, 0.5=large effect), it can be concluded that there was a great effect between the pretest and the post test of the comparison group as well.

As a conclusion, it can be stated that the statistical analysis’ results show that as a result of the treatment both the experimental and the comparison groups had achievement; it means that

during the post test the groups performed better than during the pretest. However, considering the fact that during the first exam the experimental group performed better than the comparison group and taking into account the effect size numbers, we can conclude that both groups did have achievement but the progress of the comparison groups was much better. So, the null hypothesis that there is no relationship between the implementation of formative assessment and the students' achievement is not rejected.

4.1.2. *Between group comparisons*

For comparing the average performances of both groups during pre-test and post test Independent-sample t-test was applied. An independent-samples t-test is used when you want to compare the mean scores, on some continuous variable for two different groups of subjects (Pallant, 2007, p. 232).

Table 6: Independent Samples T Test for the between group comparison for the pretest

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	Df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
First exam	Equal variances assumed	20.156	.000	4.221	28	.000	7.27778	1.72437	3.74558	10.80998
	Equal variances not assumed			3.589	12.915	.003	7.27778	2.02766	2.89433	11.66122

Table 6 presents the results of the comparison of the pretests for the experimental and comparison groups. According to table 1, Levene's test for Equality of Variances is rejecting the null hypothesis about the equality of variances because the significant value $P = 0.000$ and it is less than the selected significance level of 0.05. If the significance level of Levene's test is $p = 0.05$ it means that we should apply the information in the second line of the t-test table, which refers to Equal variances not assumed. Here T-test statistic value $t = 3.589$ and $P = 0.003$ which is smaller than 0.05. It follows that the average score for the first group is greater than the average score for the second group by value which falls between 2.9 and 11.67 with 95% of confidence.

Although the results presented above indicate difference between the two groups, they do not reveal much about the magnitude of the treatment's effect. According to Pallant (2007), one of the ways to assess the importance of the finding is to calculate the effect size. "This is a set of statistics that indicates the relative magnitude of the differences between means" (Tabachnick & Fidell in Pallant, 2007). There are a number of different effect size statistics and the most commonly used is eta squared. Eta squared can range from 0 to 1 and represents the proportion of variance in the dependent variable that is explained by the independent variable. SPSS does not provide eta squared values for t-tests; however, it can be calculated using the information in the output. The formula for eta squared is as follows:

$$\text{Eta squared} = t^2 / (t^2 + (N1 + N2 - 2))$$

According to our Table 1:

$$\text{Eta squared} = 3.5^2 / (3.5^2 + (15 + 18 - 2)) = 0.28$$

The guidelines (proposed by Pallant, 2007, p 236) for interpreting this value are as follows:

0.01=small effect

0.06=moderate effect

0.14=large effect

So from the calculated formula, we can see that the effect size of 0.28 which means that there was a large effect with a considerable difference in the pretest of the experimental and the comparison group.

As mentioned above, at the beginning of the study, the decision was to take the results of the achievement test for passing to the next level as the students' pre-test. Out of the 33 students participating in the study, some were newcomers meaning it was their first time studying at EEC and the others were former students who were continuing their course. It happened so that some of the students had already taken the test and some were still to take. Later, in the process of checking the tests, it turned out that there were two different raters and this fact was not taken into account. Besides, the details and the circumstances of the test of the previous level were not available. This can be considered as a limitation of the study.

For the results of the study to be valid, it was decided to take the results of the first test and the final test of the same level and check the students' performance based on the two unit process. This is the major limitation of the study and it somehow invalidated the results. The results of the Independent Sample T Test show during the pretest the experimental group performed better than the comparison group. But the fact that the tests were implemented after some treatment received should be taken into account. It means that maybe after the treatment the experimental group performed better which could not have been the same when that had taken the test before getting any treatment.

Table 7: Independent Samples T Test for the between group comparison for the posttest

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
final exam(40 points total)	Equal variances assumed	24.367	.000	3.284	29	.003	4.61765	1.40621	1.74162	7.49367
	Equal variances not assumed			3.000	14.151	.009	4.61765	1.53931	1.31947	7.91582

For the post test for the both groups the same procedure was applied. As the Table 2 reveals, Levene’s test for Equality of Variances is rejecting the null hypothesis about the equality of variances because $P = 0.000$ and it is less than the selected significance level of 0.05. We are considering t-test for Equality of Mean scores for the post exam with Equal variances not assumed. For the t-test $P = 0.009$ is less than 0.05, the null hypothesis about Equality of mean scores is becoming rejected. It means there is a significant difference in performances of both groups in the post test. T-test statistic value $t = 3.000$ and $P = 0.009$ which is smaller than 0.05. It follows that the average score for the first group for the posttest is significantly higher the average score for the second group by value which falls between 1.4 and 7.9 with 95% of confidence.

In order to find out how great the differences are between the test scores of the posttest of the groups, the effect size was calculated as well.

Considering the formula of calculating the effect size for the independent samples t-test and according to Table 2:

$$\text{Eta squared} = \frac{3 \times 3}{3 \times 3 + (15 + 18 - 2)} = 0.22$$

So from the calculated formula, we can see that the effect size of 0.22 which means that there was a large effect with a considerable difference in the posttest of the experimental and the comparison group.

The same implications for the post test between the groups can be made. It means that though the results indicate that during the post test the students in the experimental group performed better, the effect sizes for the posttest (experimental group Eta squared=0.28, comparison Eta squared=0.22) show that the pretest was better performed. It means we cannot come to a valid conclusion stating that the hypothesis is rejected.

Mann-Whitney U Tests for between group comparison

Because the sample sizes are small, Independent samples t-test is assuming the normality for populations of scores which is not possible to test on the basis of small samples. That's why the non-parametric Mann-Whitney U Test has to be applied instead of Independent-sample t-test. Non-parametric tests are used for comparing the average ranks. Non-parametric two independent samples Mann Whitney's U test converts the scores to ranks across the two groups. It also shows whether the ranks for the two groups differ significantly (Pallant, 2007, p. 220). Mann Whitney U tests are giving the same results as Independent-sample t-tests.

Table 8: Mann-Whitney U Test for the between group comparison for the pretest and the posttest

	pretest (40 points total)	posttest(40 points total)
Mann-Whitney U	32.000	49.500
Wilcoxon W	110.000	154.500
Z	-3.245	-2.783
Asymp. Sig. (2-tailed)	.001	.005
Exact Sig. [2*(1-tailed Sig.)]	.001 ^a	.005 ^a

In contrast with Independent t-tests, Mann-Whitney U test is comparing mean ranks of scores instead of mean scores. So, from the point of view of average ranks we have a significant difference between groups in favor of the performance of the experimental group.

According to Table 3, for the pretest, *Mann-Whitney U* test statistic value is equal to 32 and probability $p= 0.001$ which is less than the significance level of 0.05. It means that the null hypothesis (no significant difference between two groups' performances during pretest) is rejected.

Again referring back to table 1 for the posttest, *Mann-Whitney U* test statistic value is equal to 45.9 and probability $p= 0.005$ which is less than the significance level of 0.05. It means that the null hypothesis (no significant difference between two groups' performances during pretest) is rejected.

Although the results indicate difference between the two groups, they do not reveal much about the magnitude of the treatment's effect. Here again as for the Independent Samples T Test, the effect size for the pretest and the post test should be calculated in order to find out the amount of difference among the tests.

SPSS does not provide an effect size statistic, but the value of z that is reported in the output can be used to calculate an approximate value of r . So, the formula goes as follows:

$$r = z / \sqrt{N}$$

N = total number of cases

From our example, $z = -3.245$ and $N = 33$, therefore the r value = 0.56

This would be a large effect on the pretest of the groups based on Cohen (1988) criteria of 0.01 = small effect, 0.03 = medium effect, 0.05 = large effect.

The same procedure was applied for the post test. Here $z = -2.783$ and $N = 33$, therefore r value = 0.48. This again proves about the large effect on the posttest of the groups.

The same interpretation can be made for Mann-Whitney U Test as in the Independent Sample T Test. Taking into account the results of the tables provided, we can come to the conclusion that during the pretest and the posttest there was a progress in their performance. But also regarding the major limitation of the study, we can conclude that the groups progressed each on its own measure. And the effect sizes reveal (pretest $r = 0.56$, posttest $r = 0.48$) that the pretest was better performed. It means we cannot come to a valid conclusion stating that the hypothesis is rejected.

4.2. Analysis of Qualitative Data

4.2.1. Students' Questionnaires

4.2.1.1. The questionnaire for the experimental group

The qualitative data was gathered through a questionnaire which was administered to both groups after the post-tests. The aim of this questionnaire was to determine the attitude of the students towards the formative assessment.

As noted earlier the questionnaires were in the mother tongue of the students (Armenian). The reason for this was to ensure that the students clearly understood the questions they were asked and could give clear answers to the questions.

The questionnaire was analyzed quantitatively using the SPSS program through frequency analysis. As is seen in Table 1, a total of 18 students 17 were present during the answering of the questionnaire. Both questionnaires from both groups were analyzed quantitatively using the SPSS program through frequency analysis

Table 9: A total number of student questionnaires' participants

LEVEL	AGE	COURSE OF STUDY		FREQUENCY
Intermediate	12-15	Experimental	English	17 (Total 18)
		Classes (Com. 4 A)		

In this section the most evident attitudinal changes of the students towards the issues under consideration are presented. The table of descriptive statistics for the experimental group is presented to show more detailed information about the attitudes of the students. The letter x in

the table indicates the absence of students. It means that 5.6% of the 18 students of the experimental group did not answer the questionnaire.

Table 10: Responses of the questionnaire for the experimental group

		<i>Column Total N %</i>
Item 1	a	27.8%
	b	66.7%
	x	5.6%
Item 2	a	77.8%
	b	11.1%
	c	5.6%
	x	5.6%
Item 3	a	16.7%
	b	5.6%
	c	27.8%
	d	44.4%
	x	5.6%
Item 4	a	16.7%
	b	50.0%
	c	27.8%
	x	5.6%
Item 5	a	11.1%
	b	44.4%
	c	22.2%
	d	16.7%
	x	5.6%
Item 6	a	16.7%
	b	11.1%
	c	38.9%
	d	27.8%
	x	5.6%
Item 7	a	27.8%
	b	27.8%
	c	38.9%
	x	5.6%
Item 8	a	44.4%
	b	16.7%
	c	27.8%
	d	5.6%

Item 1: I liked the idea of identifying the most significant (useful, meaningful, disturbing, etc.) things I have learnt during a particular session.

- a. Strongly agree*
- b. Agree*
- c. Disagree*
- d. Strongly disagree*

As can be seen from the chart, the answers of the students are basically between a and b options. 66.7% of the students choose the answer b which means that they liked the ideas when the teacher was identifying the most significant things they had learnt during a particular session. Though the answers for the option a were not so many, 27.8%, it also proves that the above mentioned. There were no choices for the options c and d which means that the students did accept this method and technique of formative assessment.

Item 2: I remembered the vocabulary better when during the next lesson the teacher asked to write them and give examples.

- a. Strongly agree*
- b. Agree*
- c. Disagree*
- d. Strongly disagree*

The interpretation of the data shows that students basically answered the options a, b and c. The option a was dominating as 77.8% of the students preferred this option. 11.1% of the students preferred the answer b which shows that this method of vocabulary teaching was effective for the students as they confirmed that that remembered the vocabulary better when during the lesson

they were asked to write them and give examples. 5.6% of the students disagreed on the fact that this method was useful.

Item 3: I find useless when the instructor indicated incorrect responses, asked me to correct them and redo the work.

- a. *Strongly agree*
- b. *Agree*
- c. *Disagree*
- d. *Strongly disagree*

Sometimes students learn better when they think over a problem and they try to find the solution on their own. It helps them to remember the rule better. To this item whether they find useless when the instructor indicated the incorrect responses , asked them it corrects then and redo the work, 44.4% of the class strongly disagreed and 27.8% disagreed which proves that they approved this idea of implementation. However, almost 16.7% of the students did not like the idea and find it useless.

Item 4: I'd prefer infrequent long quizzes rather than what we did in class.

- a. *Strongly agree*
- b. *Agree*
- c. *Disagree*
- d. *Strongly disagree*

Formative assessment challenges to use frequent and short tests instead of infrequent and long tests during classroom instruction. Black and William (1998) state that first of all infrequent and short tests release students of the stress they usually have during the tests, besides they prepare

the student for the final test without feeling stressed or embarrassed. To find out what students think about this statement, this item tries to give answer. But as we see, the answers are quite opposite to what Black and Wiliam state. So, according to the chart, about 77% of the students preferred infrequent long quizzes rather than frequent short quizzes. However, 27.8% disagreed which means that they liked the quizzes.

Item 5: Self-assessment helped me to understand where I was and what I needed to do for improvement.

- a. Strongly agree*
- b. Agree*
- c. Disagree*
- d. Strongly disagree*

One of the implications of formative assessment is that students can become better language learners when they engage in deliberate thought about what they are learning and how they are learning it. In this kind of reflection, students step back from the learning process to think about their language learning strategies and their progress as language learners. Such self assessment encourages students to become independent learners and can increase their motivation. In order to find out the students' attitude towards this statement the item 5 was developed. According to the output of the item, it turned out that about 55 % of the students agreed on the fact that self-assessment helped them to turn back and think what they have missed about a particular issue, go back, study again and remember better. About 40 % percent of the students did not like the idea of self-assessment technique and found it useless.

Item 6: I don't like keeping portfolios. I would not keep it if my teacher would not ask me.

- a. *Strongly agree*
- b. *Agree*
- c. *Disagree*
- d. *Strongly disagree*

One of the techniques of formative assessment is keeping portfolios. A portfolio is a purposeful collection of significant work, carefully selected, dated and presented to tell the story of a student's achievement or growth in well-defined areas of performance, such as reading, writing, math, etc. "They provide an opportunity for written dialogues between teacher and student, and also provide opportunities for students to reflect on their learning process" (OECD, 2005, p 3). During the treatment of formative assessment, the students were asked to keep portfolios but whether they were interested in the idea or not, it was intended to be found out through this item. By the way, the statement of the item was expressed negatively which I think would give students more freedom to answer more honestly. So based on the results it turned out that 65% of the students found useful the implementation of portfolios. And only about 28% answered negatively.

Item 7: I liked informal meetings with the instructor, when we discussed my assignments in detail.

- a. Strongly agree
- b. Agree
- c. Disagree
- d. Strongly disagree

Formative assessment challenges to use office hours for answering student questions and find out what they do and do not understand. During the experiment there happened many cases when students were to stay after or before class to discuss the problematic issues they had about their writings. Basically this happened with the low-achieving students as they had more errors in their writings than other and needed more explanation and assistance. But it happened so that most of the students had office hours. According to the chart, almost 56% of the students approved the idea of staying long and discussing problems and only about 39% disagreed.

Item 8: I preferred to receive my instructor's assessment on my performance in the form of comments rather than grades.

- a. Strongly agree*
- b. Agree*
- c. Disagree*
- d. Strongly disagree*

Formative assessment asserts that the feedback students get should be in form of comments rather than grades. And the comments should act as a dialogue between the teacher and the student. As during the experiments the student got feedback only in form of comments and not grades, it was interesting to find their opinion about this method. According to interpretation of the data, it turned out that almost 66% of the participants preferred comments rather than grade and about 28% preferred getting grades instead of comments.

From the results of the questionnaire in administered in the experimental group, it can be inferred that most of the students approved the techniques of formative assessment and agreed on the fact that they were useful.

As mentioned before the questionnaire contained open ended items as well for the purpose to let the students express their attitude about the course, mainly mentioning the issues they liked most about the course and what they would like to change. As the open ended items were not possible to analyze with the use of SPSS package, they needed written explanation and interpretation.

Item 9: Please mention a few things that you liked best about the course.

Most of the students liked the fact that the teacher explained everything in detail and gave more feedback. Besides, they mentioned that she helped them a lot during the quizzes. Many students mentioned that they liked the formative assessment technique when at the end of a class or at the beginning of the next session they had to write down what was the most important thing they learned during the session, what was the most confusing idea, what questions remained unanswered (or were unclear), etc. Besides, they also liked the process of making presentations, the topics of essays and group works.

Item 10: Please mention a few things that you would like to change in the course.

Almost 98 % of the students desired nothing to be changed about the course. Only one person suggested that the classes lasted longer than one hour. One student suggested not losing points from the test for not having done the homework. In reality this was a trick on the part of the teacher to make students do homework and study more.

The main purpose of the questionnaire was to find out the students' attitude towards formative assessment. The items in the questionnaire were created on the basis of the formative assessment techniques implemented in during the course. The results of the questionnaire

revealed that most of the students were pleased with strategies and the techniques implemented. So, looking from the perspective of the research question “What is students’ attitude towards formative assessment”, it can be answered that students’ attitude towards formative assessment was quite positive.

4.2.1.2. The questionnaire for the comparison group

In order to find out the students’ attitude about the treatment in the comparison group, a questionnaire was developed. The questionnaire consisted of 7 items: 5 closed ones in Likert scale format (from Strongly Disagree to Strongly Agree) and two open ended. The open-ended items of the questionnaire aimed to find out whether students had any suggestions about the course and what changes they can suggest for the improvement of the course. The questionnaire was distributed among the students at the end of the study to elicit information on their attitude towards the treatment in the classroom; activities used in the classroom and the perception of their own role in the classroom (Appendix 2).

The questionnaire was analyzed quantitatively using the SPSS program through frequency analysis. As is seen in table 1, a total of 15 students 14 were present during the answering of the questionnaire and the letter x in the table indicates the absence of students. It means that 6.7% of the 15 students of the comparison group did not answer the questionnaire.

Table 11: A total number of student questionnaires’ participants

LEVEL	AGE	COURSE OF STUDY		FREQUENCY
Intermediate	12-15	Experimental	English	15 (Total 14)
		Classes (Com. 4 B)		

In this section the most evident attitudinal changes of the students towards the issues under consideration are presented. The table of descriptive statistics is presented to show more detailed information about the students' attitude towards the issues under consideration.

Table 12: Responses of the questionnaire for the comparison group

		Column Total N %
Item 1	b	26.7%
	c	46.7%
	d	20.0%
	x	6.7%
Item 2	a	20.0%
	b	60.0%
	c	13.3%
	x	6.7%
Item 3	a	26.7%
	b	20.0%
	c	26.7%
	d	20.0%
	x	6.7%
Item 4	a	13.3%
	b	60.0%
	c	20.0%
	x	6.7%
Item 5	b	13.3%
	c	73.3%
	d	6.7%
	x	6.7%

A brief explanation of the responses follows.

Item 1: When I wrote my essays to read in class, I would prefer to get more feedback and explanation from my instructor on my mistakes.

- a. *Strongly agree*
- b. *Agree*
- c. *Strongly disagree*
- d. *Disagree*

As during the treatment in the comparison group the students got feedback in forms of grades rather than comments, it was interesting to find out whether they needed more explanation from instructor on their mistakes. According to the interpretation in the chart, it turned out that about 57% of the students were pleased with receiving grades and feedback; however, 26.7% think they needed more explanation and feedback.

Item 2: When I had too many mistakes in my work, I would prefer my instructor to ask me to redo the work. It would help me to learn better.

- a. *Strongly agree*
- b. *Agree*
- c. *Strongly disagree*
- d. *Disagree*

During the treatment students used to write some short essays, come to class, read them in front of the peers and get immediate feedback and the grade. This item aimed to elicit the information about whether the students would like to get more explanation and had a chance to redo the work. According to the chart, 80 % of the students preferred to redo the job which means they needed more explanation and more thorough feedback. However, 14% were please with the situation.

Item 3: I would prefer shorter and more frequent quizzes less frequent quizzes that we did in this course.

- a. *Strongly agree*
- b. *Agree*
- c. *Strongly disagree*
- d. *Disagree*

During the term this group had only one chance to have a quiz. So it was interesting to find out whether they would like to have shorter and more frequent quizzes. The chart provided interesting results. 50% of the students agreed and 50% disagreed.

Item 4: I would like to keep a portfolio (a collection of my assignments) for my assignments.

- a. *Strongly agree*
- b. *Agree*
- c. *Strongly disagree*
- d. *Disagree*

As this group did not have a chance to keep a portfolio, it was purposeful to find out whether they would like to. And as the term “portfolio” might not be similar to them, it needed an explanation in the item. 73% of the students were attracted by the idea and answered positively. However, 20% disagreed.

Item 5: I would like to have individual meetings with my instructor to discuss my assignments.

- a. *Strongly agree*
- b. *Agree*

c. *Strongly disagree*

d. *Disagree*

According to the chart results, it turned out that 86% of the students refused to have individual meeting with the instructor. And 6.7% agreed.

From the results of the questionnaire in administered in the comparison group, it can be inferred that most of the students approved and liked their way of teaching, they did not need any further explanations or clarifications of their mistakes, however, they approved the idea of redoing their assignments in case of failure or low grade.

The questionnaire for the comparison groups also contained open ended items for the purpose to let the students express their attitude about the course, mainly mentioning the issues they liked most about the course and what they would like to change. As the open ended items were not possible to analyze with the use of SPSS package, they needed written explanation and interpretation.

Item 6: Comment on what you like best about the course.

For the comparison group, the answers for the open-ended items varied greatly among the students. One of the students liked the fact that during the session, only English was used. She explained that thus helped her to get more used to the language. Most of the students liked the process of the classes, the types of activities, essays and mentioned that they helped them to speak more fluently.

Item 7: Comment on what you would like to change in the course.

99% of the students answered that there was nothing necessary to be changed and one of them suggested the classes lasting longer.

Though the comparison group did not receive treatment, it was interesting to find out their attitude towards the class procedure as well. The items in the questionnaire were created on the basis of the basic strategies implemented in during the course. Besides, there were some items concerning formative assessment as well, with the aim to find out whether the students would like this type of procedure. The results of the questionnaire revealed that most of the students were pleased with the strategies and would not like to be treated differently.

4.2.2. Field notes

The reason for taking some field notes was to consider the teacher's perceptions of the treatment implemented in the classroom. As the treatment involved basically the use of the techniques of the formative assessment suggested, after each implementation of a technique some notes were recorded on the part of the teacher to have a better picture and express her perceptions about each technique. The field notes were categorized according to some patterns which basically describe the positive and negative aspects of each technique. Besides, field notes included some student works which served as a proof of the procedure throughout the course.

Table 13: Patterns emerged out of filed notes:

Positive attitudes	Negative attitudes
✓ The students remember things better when they are explained each pattern corrected and explained during the next class.	✓ Students don't like the idea of keeping portfolios
✓ Students like when at the end of a class they summarize the main parts of the class.	✓ They don't like the frequent short tests
✓ They like staying late after class	✓ They were subjective and superficial during the peer and self assessments

Students like when at the end of a class they summarize the main parts of the class.

As one of the most important issues of formative assessment was to provide a detailed and thorough feedback, this technique was implemented quite often, almost every class and the focus was not only the explanation of errors, but other activities as well, for instance when students asked a vocabulary question or something not concerning the lesson, more explanations and clarifications were provided on the part of the teacher. And the perception was that the students liked this idea and were eager to ask more questions and to elicit more answers.

Students like when at the end of a class they summarize the main parts of the class.

At the end of some classes the students were asked to summarize in a few sentences the important patterns of the class, basically what they remember and where they had difficulties. In my opinion, after implementation this techniques once or twice, during next classes students were more attentive as they thought the techniques might be implemented during that particular session.

They like staying late after class

Those students who happened to make many errors in their assignments were asked to stay 10-15 minutes to discuss and clarify the points. In my opinion, they were enthusiastic about the idea they could be explained in a more relaxed atmosphere and could ask more questions and seek answers than they could do them in the presence of their peers.

Students don't like the idea of keeping portfolios

The students were asked to keep a portfolio of their works as it could serve as a guide for them for further actions as well as for revising issues. In my opinion, this did not motivate

students as some of them especially boys found it uncomfortable to carry out the large folder of portfolio.

They don't like the frequent short tests

Sometimes the students were stressed during these short and frequent tests as they compared them with the midterm and the final tests where they did their best to get good results. Sometimes they misunderstood the real meaning of these tests.

They were subjective and superficial during the peer and self assessments

When expressing their ideas about their peers and themselves, sometimes students were subjective and sometimes they did not get the real point of peer and self assessments. For instance, some of the students wrote about the character of their peers, what taste he/she has, etc. Maybe this was a result of not proper instruction of the part of the teacher.

Chapter 5: Conclusion

This chapter presents a summary of the research findings, answers the research questions, points out the limitations and makes suggestions for the further research.

5.1. Summary of the research findings

The research findings can be summarized in accordance with the research questions:

- 1. Is there a relationship between the implementation of formative assessment and students' achievement in language learning in an Armenian setting?*

According to the findings, both the comparison and the experimental groups had a progress during the course. The qualitative results revealed that both groups performed better in the posttest. However, taking into account the major limitation of the study, that is the fact that based on the circumstances, the pretests were administered after some treatment being implemented and the fact that during the pretest the experimental group performed much better it can be concluded that the implementation of the formative assessment did not have its effect on the students' achievement in the process of English language learning.

- 2. What is the students' attitude towards formative assessment?*

The results of the qualitative data (questionnaire and field notes) revealed that students in both classes were motivated to learn English. They were willing to complete any activity assigned by their teacher. The questionnaires developed revealed an interesting point that the students on both groups liked the style of teaching though they were quite different. For instance, in the experimental group when students were writing essays, after the checking all the errors and mistakes were thoroughly discussed and analyzed. However, in the comparison group the students were only corrected and graded. The answers to the questionnaires revealed that both

groups liked each way of teaching, meaning the comparison group did not think over a better explanation of their mistakes and more clarifications. But as the questionnaire results show they were inclined to redo their work in case of bad results which the experimental group experienced during the treatment.

The qualitative results also revealed that the students in the experimental groups liked almost strategies emerged through formative assessment and thought that it helped them to learn in an easier way and remember things better.

5.2. Limitations of the study

Several limitations of the present study should be observed. First and the most important limitation of the study concerned the pretest which in this case appeared to be the first exam of the course. The reason for doing so was that some of the students had already taken the test and some were still to take. Later, in the process of checking the tests, it turned out that there were two different raters and this fact was not taken into account. Besides, the details and the circumstances of the test of the previous level were not available. The second limitation of the study was limited to low-intermediate level students, it was not possible to see the effect formative assessment on advanced level students because formative assessment of basically implemented in more academic level. The third limitation was that the study was not in a large-scale: it included only 33 students and besides, the samples were not randomly selected. Thus, the findings of this research are limited to AUA and its results cannot be generalized to other educational institutions. And the last limitation was that though the treatment lasted 10 weeks, only the two weeks results of the tests were taken into account.

5.3. Suggestions for further research

Taking into account the first limitation of the study, it would be better to carry out similar research but having a separate pretest for both groups and what is more important it should be administered under the same circumstances. Besides, the test should be administered at the beginning of the course and not after the treatment. Further, it would be interesting to carry out a similar study with advanced level students in order to determine the effect formative assessment on advanced level students, basically on an academic level, as well as their attitude towards the process. Besides, for further investigation it is recommended to conduct more research and find out the relationship between formative assessment and students' motivation. And finally, it would be more effective to conduct longitudinal studies, using more groups as it may yield more information about the issues under consideration and may affect the results of the study.

References

Airasian, P. W. (2001). *Classroom assessment: Concepts and applications*. New York: McGraw-Hill.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Ames, C. (1992). *Classrooms: Goals, structures, and student motivation*. *Journal of Educational Psychology*, 84 (3): 261-271.

Angelo, Thomas A. and Cross, K. Patricia. (1993). *Classroom Assessment Techniques: A Handbook for College Teachers (2nd ed.)*. San Francisco: Jossey-Bass.

Assessment Action Group (AAG)/ AiFL Programme Management Group (APMG). (2002 – 2008). *AiFL - Assessment is for learning*. Retrieved on February 8, 2008 from:
<http://www.ltscotland.org.uk/assess>

Assessment Reform Group. (1999). *Assessment for learning: Beyond the black box*. Cambridge, University: Cambridge School of Education.

Bangert-Drowns, R.L., Kulick, J.A., and Morgan, M.T. (1991). *The instructional effect of feedback in test-like events*. *Review of Educational Research*, 61 (2): 213-238.

Barron B.J.S., Schwartz D.L., Vye N.J., Moore A., Petrosino A., Zech L., Bransford J. & The Cognition and Technology Group at Vanderbilt (1998). *Doing with understanding: lessons from research on problem- and project-based learning*. *Journal of the Learning Sciences* 7, 271–311.

Bell B. & Cowie B. (2001) *The characteristics of formative assessment in science education*. *Science Education* 85, 536–553.

Birenbaum, M. (1994) *Toward adaptive assessment—the students' angle*, *Studies in Educational Evaluation*, 20, 239–255.

Birenbaum, M. (1997). *Assessment preferences and their relationship to learning strategies and orientations*. *Higher Education*, 33, 71–84.

Birenbaum, M. & Feldman, R. (1998) *Relationships between learning patterns and attitudes towards two assessments formats*, *Educational Research*, 40(1), 90–98.

Black, P. (2003, April). *The nature and value of formative assessment for learning*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Black, P., Harrison, C., Lee, C., Marshall, B. & Wiliam, D. (2003) *Assessment for learning: putting it into practice* (Buckingham, Open University Press).

Black, P. J., & Wiliam, D. (1998). *Inside the black box: Raising standards through classroom assessment*. *Phi Delta Kappan*, 80, 139–148.

- Black, P. & William, D. (1998) *Assessment and classroom learning*, *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74.
- Black, P. & Wiliam, D. (1998a). *Assessment and classroom learning*. *Assessment in Education: Principles, Policy, and Practice*. 5(1): 7-74.
- Black, P. & Wiliam, D. (1998b). *Inside the black box: Raising standards through classroom assessment*. London: School of Education King's College.
- Black, P. & Wiliam, D. (2003). 'In praise of educational research': *Formative assessment*. *British Educational Research Journal*, 29 (5), 623-637.
- Bloom, B. S. (1968) *Learning for mastery*. *Evaluation Comment*. Los Angeles: University of California.
- Bloom, B. S., Hastings, J.T. & Madaus, G.F. (Eds.) (1971). *Handbook on the formative and summative evaluation of student learning*. New York: McGraw-Hill Book Company.
- Boud, D. (1995) *Assessment and learning: contradictory or complimentary?* in: P. Knight (Ed.) *Assessment for learning in higher education*. London: Kogan Page.
- Boston, Carol (2002). *The concept of formative assessment*. *Practical Assessment, Research & Evaluation*, 8(9).

Bransford J.D., Brown A. & Cocking R. (2000) *How People Learn: Mind, Brain, Experience and School, Expanded Edition*. Washington, DC, National Academy Press.

Boyle, J.T. and Nicol, D. J. (2003) *Using classroom communication systems to support interaction and discussion in large class settings*, *Association for Learning Technology Journal*, 11(3), 43-57.

Brookhart, S. M. (2007). Expanding views about formative assessment: A review of the literature. In J. H. McMillan (Ed.), *Formative classroom assessment: Theory into Practice* (pp. 43-62). New York: Teachers College Press.

Brown S. & Knight P. (1994) *Assessing Learners in Higher Education*. Kogan Page, London.

Brown H. Douglas (1994). *Principles of Language Learning and Teaching*. Prentice Hall. Inc. Englewood Cliffs. New Jersey 07632.

Brown H. Douglas (2001). *Teaching by Principles: An Interactive Approach To language Pedagogy*, Second Edition, San Francisco Public University.

Brown H.D. (2004) *Classroom Assessment: Principles and Classroom Practice*. Pearson Education.

Buchanan T. (2000) *The efficacy of aWorld-WideWeb mediated formative assessment*. *Journal of Computer Assisted Learning* 16, 193–200.

Butler, R. (1987) *Task-involving and ego-involving properties of evaluation: effects of different feedback conditions on motivational perceptions, interest and performance*. Journal of Educational Psychology, 78(4), 210-216.

Butler, R. (1988) *Enhancing and undermining intrinsic motivation: the effects of task involving and ego-involving evaluation on interest and involvement*. British Journal of Educational Psychology, 58, 1-14.

Centre for Educational Research and Innovation. (2005). *Formative assessment: Improving learning in secondary classrooms*. Retrieved March 2005 from the Organization for Economic Co-operation and Development Web site:

<http://www.oecdbookshop.org/oecd/display.asp?sf1=identifiers&st1=962005021P1>

Chamber G.N. (1999). *Motivating Language Learners*. Clevedon. Multilingual Matters. Ltd.

Cizek, G. J. (2010). *An introduction to formative assessment: History, characteristics, and challenges*. In G. J. Cizek and H. L. Andrade (Eds.), *Handbook of formative assessment* (pp. 3-17). New York, Routledge.

Cizek and H. L. Andrade (Eds.), *Handbook of formative assessment* (pp. 106-124). New York, Routledge.

Clapham, C. (2000). *“Assessment and Testing”*. Annual Review of Applied Linguistics. Cambridge: Cambridge University Press.

Craven, R.G., Marsh, H. W., & Debus, R.L. (1991) *Effects of internally focused feedback on the enhancement of academic self-concept*, *Journal of Educational Psychology*, 83(1), 17-27.

Dweck, C. (1999) *Self-theories: Their Role in Motivation, Personality and Development*> Philadelphia. Psychology Press.

Earl, L. M. (2003). *Assessment as learning: Using classroom assessment to maximize student learning*. Thousand Oaks, CA: Corwin Press, Inc.

Ebel R.L. & Frisbie D.A. (1991) *Essentials of Educational Measurement*, 5th edn. Prentice Hall, Inc, Englewood Cliffs, NJ.

Elawar, M.C., and Corno, L. (1985). *A factorial experiment in teachers' written feedback on student homework: Changing teacher behavior a little rather than a lot*. *Journal of Educational Psychology*, 77 (2): 162-173.

Emerson, R., Fretz, R., & Shaw, L. (1995). *Writing ethnographic fieldnotes*.

Entwistle, N. J. & Tait, H. (1990) *Approaches to learning, evaluations of teaching, and preferences for contrasting environments*, *Higher Education*, 19, 169–194.

Eraut, M. (2006) *Formative assessment in the workplace*, paper given to the *Annual Meeting of the American Educational Research Association*, San Francisco, April 2006.

Ferlazzo L. (2010) "The Best Resources for Learning about Formative Assessment" *Samoset Middle School Blog*

Freeman, R. & Lewis, R. (1998) *Planning and Implementing Assessment* (London, Kogan Page).

Lunsford, R. (1997) When less is more: principles for responding in the disciplines, in: M.

Sorcinelli and P. Elbow (Eds) *Writing to learn: strategies for assigning and responding to writing across the disciplines* (San Francisco, Jossey-Bass).

Gipps, C. (1994) *Beyond testing: towards a theory of educational assessment* London, Falmer Press.

Gipps C., Brown, M., McCallum, B. & McAlister, S. (1995) *Intuition or evidence? Teachers and national assessment of seven year olds*. Buckingham, Open University Press.

Guskey, T. R. (2010). *Formative assessment: The contributions of Benjamin S. Bloom*. In G. J.

Hall K. & Burke, W. (2004) *Making formative assessment work: effective practice in the primary classroom*. Maidenhead, Open University Press.

Harlen, W. & Crick, R.D. (2003) Testing and motivation for learning, *Assessment in Education*, 10(2), 169-207.

Hastings, T. and Madaus, G. (1971) *Handbook of formative and summative evaluation of student learning*. New York: McGraw-Hill Book Company Kifer,

Hayward, L. & Hedge, N. (2005). Travelling towards change in assessment: Policy, practice and research in education. *Assessment in Education*, 12 (1), 55-75.

Izard, J. (1992). Assessing learning achievement. Paris: UNESCO.

James, M. (1994). *Different countries, common problems?* *Assessment in Education*, 1(2), 261-268.

John Pryor and Barbara Crossouard (2008) *A socio-cultural theorization of formative assessment* *Oxford Review of Education University of Sussex*, UK Vol. 34, No. 1, February 2008, pp. 1–20

Kahl, S. (2005a, October 26). *Where in the world are formative tests? Right under your nose!* *Education Week*, 25(9), p. 38.

Kizlik B. (2010), “*Measurement, Assessment, and Evaluation in Education*”, *Education Information for New and Future Teachers*.

Laurillard, D. (2002) *Rethinking University Teaching: a conversational framework for the effective use of learning technologies*, 2nd edition London, RoutledgeFalmer.

Martinez, J. G. R., & Martinez, N. C. (1992). Re-examining repeated testing and teacher effects in a remedial mathematics course. *British Journal of Educational Psychology*, 62, 356-363.

Marzano, R. J. (2006). *Classroom assessments and grading that work*. Alexandria, VA: Association for Supervision and Curriculum Development.

McManus, S. (2006). "*Attributes of Effective Formative Assessment*." Accessed March 20, 2009 from The Council of Chief State School Officers website. Website: <http://www.ncpublicschools.org/docs/accountability/educators/fastattributes0408>

- Messick, S. (1994) *The interplay of evidence and consequences in the validation of performance assessment*. *Educational Researcher*, 23(2), 13-23.
- Murphy, R.T. (1975). *Assessment*. In S.B. Anderson, S. Ball, & R.T. Murphy (Eds.), *Encyclopedia of educational evaluation*. San Fransisco : Jossey-Bass Publishers.
- Nenty, H.J. (1997 a). *Basic assesement skills for classroom teachers*. Unpublished manuscript. Institute of Education, National University of Lesotho, Lesotho.
- Nenty, H.J. (1997b). *Assessing learning and managing examinations*. A paper presented in a Management Training Workshop of Primary School Principal at National University of Lesotho, July 13-23.
- Nenty, H.J. (1999). *Assessment a means of enhancing improved quality of life through education* Institute of Education, National University of Lesotho
- Nicol, D.J. & Boyle, J.T. (2003) *Peer Instruction versus Class-wide Discussion in large classes: a comparison of two interaction methods in the wired classroom*, *Studies in Higher Education*, 28(4), 457-473.
- Nicol, D. & Macfarlane-Dick, D. (2006) *Formative assessment and self-regulated learning: a model and seven principles of good feedback practice*, *Studies in Higher Education*, 31(2), 199–218.
- Nichols, P. D., Meyers, J., & Kelly B. (2008). *What is formative assessment?* *Educational Measurement Research Bulletin*, Issue 5. Retrieved from www.PearsonEdMeasurement.com

OECD (2005), *Formative Assessment: Improving Learning in Secondary Classrooms*. Paris, ISBN: 92-64-00739-3

Perrenoud P. (1998) *From formative evaluation to a controlled regulation of learning processes. Towards a wider conceptual field*. *Assessment in Education* 5, 85–102.

Popham, W. J. (October 2006). *Defining and enhancing formative assessment*. Paper presented at the Annual Large-Scale Assessment Conference, Council of Chief State School Officers, San Francisco, CA.

Ramaprasad, A. (1983). *On the definition of feedback*. *Behavioral Science*, 28 (1): 4-13.

Sadler D.R. (1989) *Formative assessment and the design of instructional systems*. *Instructional Science* 18, 119–144.

Scriven, M. (1967). *The Methodology of Evaluation*. In R. w. Tyler (Ed.), R. M. Gange (Ed.) & M. Scriven (Ed.), *Perspectives of curriculum evaluation* (39-83). Chicago IL: RandMaNally.

Shepard, L., Hammerness, K., Darling-Hammond, L., Rust, F. (2005). *Assessment*. In L. Darling-Hammond & J. Bransford (Eds.), *Preparing teachers for a changing world: What teachers should learn and be able to do* (pp. 275-326). San Francisco: Jossey-Bass.

Smith, E. & Gorard, S. (2005). *They don't give us our marks: The role of formative feedback in student progress*. *Assessment in Education*, 12 (1), 21-38.

Stiggins, R. J. (2001). *Student-involved classroom assessment* (3rd ed.). Upper Saddle River, NJ: Prentice-Hall

Stiggins, R. J. (2002). *Assessment crisis: The absence of assessment for learning*. Phi Delta Kappan, 83, 758–765.

Tittle, C. K., Hecht, D., & More, P. (1993) *Assessment theory and research for classrooms: From taxonomies to constructing meaning in context*. Educational Measurement: Issues and Practice, 12(4), 13-18

Threlfall, J. (2005). *The Formative Use of Assessment Information in Planning – the Notion Of Contingent Planning*. British Journal of Educational Studies, 53 (1), 54–65.

Triandis, H. C. (1971). *Attitudes and attitude change*. New York: Wiley.

Vispoel, W.P., and Austin, J.R. (1995). *Success and failure in junior high school: A critical incident approach to understanding students' attribution beliefs*. American Educational Research Journal, 32 (2): 377-412.

Williams C.E. (2004) *School district of Philadelphia uses Web-based system to increase student achievement*. *T H E Journal* 31, 51–52.

William, D., & Leahy, S. (2007). *A theoretical foundation for formative assessment*. In J. H. McMillan (Ed.), *Formative Classroom assessment: Theory into practice* (pp. 29-42). New York: Teachers College Press.

William, D. (2010). *An integrative summary of the research literature and implications for a new theory of formative assessment*. In G. J. Cizek and H. L. Andrade (Eds.), *Handbook of formative assessment* (pp. 18-40). New York, Routledge.

Wood, R. (1984). *Assessment has two many meanings and the one I think we want isn't clear enough yet*. *Educational Measurement: Issues and Practice*, 3(4), 5-7.

Yorke, M. (2003). *Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice*. *Higher Education*, 45(4), 477-501.

g. Disagree

h. Strongly disagree

4. I'd prefer infrequent long quizzes rather than what we did in class.

e. Strongly agree

g. Disagree

f. Agree

h. Strongly disagree

5. Self-assessment helped me to understand where I was and what I needed to do for improvement.

e. Strongly agree

g. Disagree

f. Agree

h. Strongly disagree

6. I don't like keeping portfolios. I would not keep it if my teacher would not ask me.

e. Strongly agree

g. Disagree

f. Agree

h. Strongly disagree

7. I liked informal meetings with the instructor, when we discussed my assignments in detail.

e. Strongly agree

g. Disagree

f. Agree

h. Strongly disagree

8. I preferred to receive my instructor's assessment on my performance in the form of comments rather than grades.

e. Strongly agree

g. Disagree

f. Agree

h. Strongly disagree

9. Please mention a few things that you liked best about the course.

10. Please mention a few things that you would like to change in the course.

□□□□□□□ □□□□□□□□□□ □□□□□□□□□□

□□□□□ _____

□□□□□□□ □□□□□□□□□□□□□□□□ □ □□□□□ □□□□□□□□□□□□□□

□□□□:

1. □□□ □□□□ □□ □□□□□, □□□ □□□□ □□□□□□□□□□ □□
□□□□□□□□□□□□ □□ □□□□□□□□□□□, □□□□□□□, □□□□□□□□□□,
□□□□□□□□□□□□□□ □□□□□□□:
□. □□□□□□□ □□□□□□□□□□ □□ □. □□□□ □□
□. □□□□□□□□□□ □□ □. □□□□□□□□□□□□ □□□ □□

2. □□ □□□□□□ □□ □□□□□□ □□□□□ □□□, □□□ □□□□□□□□□□ □□□□□□
□□□□□ □□□□□□□□□□□□ □ □□□□ □□□□□ □ □□□□□ □□□□□□□□□□□:
□. □□□□□□□ □□□□□□□□□□ □□ □. □□□□ □□
□. □□□□□□□□□□ □□ □. □□□□□□□□□□□□ □□□ □□

3. □□□ □□□□□□□□ □□□□□, □□□ □□□□□□□□□□ □□□□□ □□ □□□□□ □□
□□□□□□□□□, □□□□□□□□□□□□□□ □□ □□□□□ □ □□□□□□□ □□ □□□□□□
□□□□□□□□ □□□□□□□□□□□:
□. □□□□□□□ □□□□□□□□□□ □□ □. □□□□ □□
□. □□□□□□□□□□ □□ □. □□□□□□□□□□□□ □□□ □□

4. □□ □□□□□□□□□□ □□ □□□□ □□ □□□□□□□□ □□□□□ □□□□□□□□□□
(quizzes), □□□□ □□□□, □□□□ □□□□□□□□, □□□□□□ □□□□ □□□□□ □□□□
□□□□□□ □□□□□□□□□□:

- . □□□□□□ □□□□□□□□ □□
- . □□□□□□□□ □□

- . □□□□ □□
- . □□□□□□□□□□ □□□□ □□

5. □□□□□□□□□□□□□□ □□□□□□ □□ □□□ □□□□□□□□, □□ □□□□
□□□□□□ □□ □ □□□ □ □□□ □□□□ □□□□□□□□□□□□ □□□□□□□□□□□□
□□□□□□:

- . □□□□□□ □□□□□□□□ □□
- . □□□□□□□□ □□

- . □□□□ □□
- . □□□□□□□□□□ □□□□ □□

6. □□□□ □□□□ □□ □□□□□ □□□□□□□□□□ □□□□□□□□:

- . □□□□□□ □□□□□□□□ □□
- . □□□□□□□□ □□

- . □□□□ □□
- . □□□□□□□□□□ □□□□ □□

7. □□□□ □□□□ □□□□ □□□□□ □□□□□□□□□□ □□□□□□□□□□□□
□□□□□□□□ □□□, □□□□ □□□□ □□□□□□□□ □□□□ □□ □□□□□□□□
□□□□□□ □□□□□□□□□□:

- . □□□□□□ □□□□□□□□ □□
- . □□□□□□□□ □□

- . □□□□ □□
- . □□□□□□□□□□ □□□□ □□

8. □□ □□□□□□□□□□ □□ □□□□□□ □□ □□□□□□□□□□□□□□□□
□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□□□, □□□□ □□□□□□□□□□□□
□□□□□□□□:

□. □□□□□□ □□□□□□□□ □□

□. □□□ □□

□. □□□□□□□□ □□

□. □□□□□□□□□□ □□□ □□

9. □□□□, □□ □□□□ □ □□□ □□□□□□□□ □□□□ □□□□:

10. □□□□, □□ □□□ □□□□□□□□□□ □□□□□:

Appendix #2

Questionnaire for further improvement

Name _____

Read the statements and choose one of the options.

1. When I wrote my essays to read in class, I would prefer to get more feedback and explanation from my instructor on my mistakes.

i. Strongly agree

k. Strongly disagree

j. Agree

l. Disagree

2. When I had too many mistakes in my work, I would prefer my instructor to ask me to redo the work. It would help me to learn better.

e. Strongly agree

g. Strongly disagree

f. Agree

h. Disagree

3. I would prefer shorter and more frequent quizzes less frequent quizzes that we did in this course.

e. Strongly agree

g. Strongly disagree

f. Agree

h. Disagree

4. I would like to keep a portfolio (a collection of my assignments) for my assignments.

e. Strongly agree

g. Strongly disagree

f. Agree

h. Disagree

5. I would like to have individual meetings with my instructor to discuss my assignments.

e. Strongly agree

g. Strongly disagree

f. Agree

h. Disagree

6. Comments on what you like best about the course.

7. Comments on what you would like to change in the course.



□□□□□□□□ □□□□□□□□□□□□□□□□ □ □□□□□□ □□□□□□□□□□□□□□

□□□□:

1. □□□ □□ □□□□□ □□ □□ □□□□□□□□□□□□□□□□□ □□□□, □□□□□□□□
□□□□□□□□ □□□□ □□□□ □□□□□□□□, □□ □□□□□□□□□□ □□□□□ □□□□
□□□□□□□□□□□□□□□□ □□□□□□ □□ □□□□□□□□ □□□□□□□□:
□. □□□□□□ □□□□□□□□ □□ □. □□□ □□
□. □□□□□□□□ □□ □. □□□□□□□□□□ □□□ □□

2. □□□ □□ □□□□□□□□ □□□ □□□□□□□□ □□ □□□□□□□□ □□
□□□□□□□□□ □□□, □□ □□□□□□□□□□, □□ □□□□□□□□□□
□□□□□□□□□□□□ □□□□□ □□□□□□□□ □□□□□□□□□□: □□□□□□ □□
□□□□□ □□□ □□□□□□□□□□:
□. □□□□□□ □□□□□□□□ □□ □. □□□ □□
□. □□□□□□□□ □□ □. □□□□□□□□□□ □□□ □□

3. □□ □□□□□□□□□□ □□□□ □□□□ □□□□□□□□ □□□□□□□□□□, □□□□
□□□□ □□□□□ □□□□□□□□□□:
□. □□□□□□ □□□□□□□□ □□ □. □□□ □□
□. □□□□□□□□ □□ □. □□□□□□□□□□ □□□ □□

4. □□ □□□□□□□□□□ □□ □□□□□□□□□□□□□□□□ □□□□□ □□□□□□□□□□
(□□□□□□□□□□□□□□□□ □□□□□□□□□□) □□□□□□□□□□:
□. □□□□□□□□ □□
□. □□□ □□
□. □□□□□□ □□□□□□□□ □□ □. □□□□□□□□□□ □□□ □□

5. □□ □□□□□□□□ □□□□□□□□ □□□□□□□□□□ □□□□□□ □□
□□□□□□□□ □□□, □□□□□□ □□□□□□□□ □□ □□□□□□□□□□□□□□:

□. □□□□□□ □□□□□□□□ □□

□. □□□□□□□□ □□

□. □□□ □□

□. □□□□□□□□□□ □□□ □□

6. □□□□, □□ □□□□ □ □□□ □□□□□□□□□□ □□□□ □□□□:

7. □□□□, □□ □□□ □□□□□□□□□□ □□□□□:
