



THE IMPACT OF FORMATIVE ASSESSMENT ON STUDENTS' ASSESSMENT PREFERENCES

Kagan Buyukkarci
Suleyman Demirel University

Sehnaz Sahinkarakas
Cag University

ABSTRACT

Research in the area of Language Testing and Assessment proves that formative assessment (FA) leads to active involvement of students in assessment practices, and it is a valuable approach in promoting learning. The aim of this study was to gain information about the relationship between hands-on formative assessment practices and students' assessment preferences. From the analysis of the traditional assessment preferences of the students, it can be concluded that both the experimental and control group students still preferred traditional type of assessments such as multiple-choice tests. On the other hand, the analysis of the formative assessment preferences of the experimental group students indicated that they began to prefer the formative mode of assessment. In other words, they added new types of assessment to their preferences such as self/peer assessment. As a result, they began taking a more active role in their own assessment procedure, and took a step towards becoming autonomous learners.

INTRODUCTION

Testing or measurement is a method of evaluation of professional activities using clear criterion and frequently together with an effort at measurement either by grading on a rough scale or by assigning numerical value. On the other hand, the word assessment literally means a consideration of someone or something and a judgment about them, and it is a wider domain than testing (Brown 2004), which can sometimes interchangeably be used with the terms testing, measurement and evaluation. Lambert and Lines (2001) describe assessment as: a) "a fact of life for teachers, part of what teachers do; b) an organic part of teaching and learning; c) a part of the planning process." (2). Erwin (1991) goes in detail in his definition of assessment as the process of collecting information on student achievement and performance, and also as the process of documenting, usually in measurable terms, knowledge, skills, attitudes and beliefs. These collected documents provide the basis for decision making regarding teaching and learning (Alderson, Brunfaut, & Harding 2017).

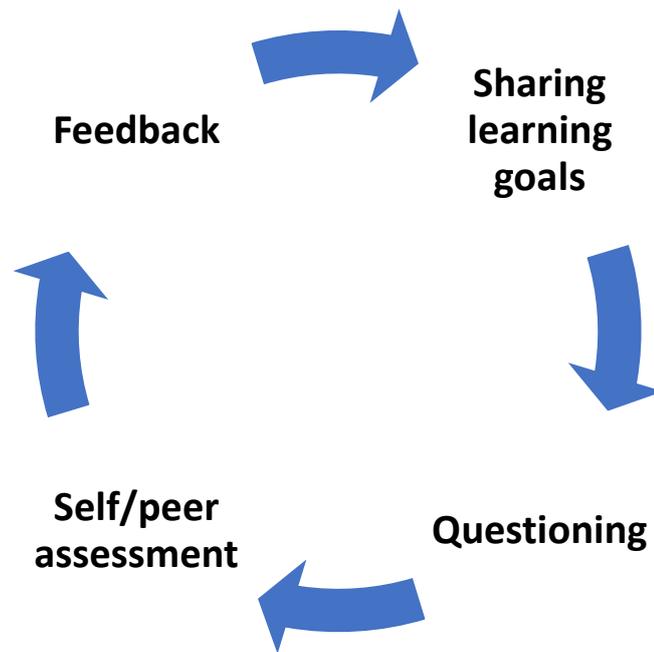
In spite of the variety in the way assessment is defined, it's commonly agreed that assessment is an essential part of teaching, by which teachers make a judgment about the level of skills or knowledge (Taras 2005), to measure improvement over time, to evaluate strengths and weaknesses of the students, to rank them for selection or exclusion, or to motivate them (Wojtczak 2002). Moreover, assessment can help individual instructors obtain useful feedback

on what, how much, and how well their students are learning (Taras 2005; Stiggins 1992). Its systematic process provides teachers evaluating an opportunity to meaningfully reflect on how learning is best delivered, gather evidence of that, and then use that information to improve.

When we go through the literature, we find that assessment can be classified in two main categories: The first one is *summative assessment* which is also called as assessment of learning (Earl 2003; Stiggins 2002; Torres, 2019). In an educational setting, summative assessments are typically used to assign students a course grade at the end of a course or project. Taras (2005) stated that summative assessment is a judgment which summarizes all the evidence up to a given point. This certain point is seen as finality at the point of the judgment. This type of assessment can have various functions, such as shaping how teachers organize their courses or what schools offer their students, which do not have an effect on the learning process.

The second category is *formative assessment* (assessment for learning) (McCallum & Milner, 2020; Stiggins 2002; Derrich and Ecclestone 2006). According to Black and Wiliam (1998b), assessment is referring to all those activities undertaken by teachers, and by their students in assessing themselves, which provide information to be used as feedback to change the teaching and learning activities in which they are engaged. Such assessment becomes formative assessment when the evidence is actually used to adapt the teaching work to meet the needs. In Threlfall's (2005) terms, "formative assessment may be defined as the use of assessment judgments about capacities or competences to promote the further learning of the person who has been assessed" (p. 54).

As stated before, assessment is a process to gain information about students' learning progress and their difficulties in learning, and make decisions about their students (Black & William 1998a; Hancock 1994; Stiggins 1992). This kind of assessment turns out to be formative when the evidence is used to adapt the teaching to meet students' needs. In general terms, formative assessment is concerned with helping pupils to improve their learning. In practice, formative assessment is a self-reflective process that intends to promote student attainment (Crooks 2001). Cowie and Bell (1999) define it as the bidirectional process between teacher and student to improve, recognize and respond to the learning. Similarly, Shepherd (2005) explains formative assessment as 'a dynamic process in which supportive teachers or classmates help students move from what they already know to what they are able to do next, using their zone of proximal development'(66). Formative assessment aims at optimizing the measurement of students' intellectual abilities. They try to provide a more complete picture of child's real and maturing cognitive structures and performance and, on this basis, advance the diagnosis of learning difficulties (Allal & Ducrey 2000). Black and Wiliam (1998a) set out four main headings for formative assessment practice: sharing learning goals, questioning, self/peer assessment, feedback (See the figure below).



The first principle of formative assessment is sharing learning goals: learning opportunities are more likely to succeed if learners have a clear, specific understanding of what they are learning. The second one is questioning: Effective questioning is an important element in formative assessment, and it includes matching the questions with learning target, engaging the whole class and providing enough wait time for students to respond (Forbes 2007; Harlen 2007). Another principle is self/peer assessment: It provides feedback for the future adjustment of teaching and learning activities learners are engaged (Black & William 1998; Vogt, Tsgari, Csepes, Green, & Sifakis 2020) and learners take ownership of their own learning and see themselves as partners in the teaching-learning process (Harlen 2007). The last important principle is feedback: Formative feedback is more effective when it gives details about why student's answer is correct or incorrect together with commentary on good or poor strategy.

Assessment Preference

It is commonly known that assessment has a crucial role in students' learning process. The way students prepare themselves for an assessment largely depends on how they see the assessment, and this can affect their learning positively or negatively (Watering, Gijbels, Dochy, & Rijt 2008). The development and implementation of teaching practices that will promote students to obtain and apply their knowledge efficiently, think critically, analyze, synthesize, and make inferences are the important challenges for today's higher education (Gijbels, Segers, & Struyf 2008). In general, it is claimed that new learning environments have the potential to improve these educational outcomes for students in higher education by making the students' learning the core subject and defining instruction as enhancing learning (Lowe et al 2008).

Clarke, Heaney and Gatfield (2005) assume that students in all countries prefer multiple-choice testing, and they link this choice to a widespread perception of ease, a lesser risk of failure on the basis of linguistic expression, and the statistically favorable potential of blind guessing (In Bartram & Bailey 2010). However, Birenbaum (2007) argue that students preferring problem-solving tasks tend to perform better than students preferring simple, quick, and easy problems. Besides, since new assessment methods are largely available in the last

decades, students' assessment preferences are not restricted to their experiences of multiple-choice examinations and essay modes of evaluation anymore (Struyven, Dochy, and Jassens 2005).

It can be said that the students' preferences through instruction and assessment can also reflect their perceptions of learning environment, and their approaches to learning which can affect their success. Birenbaum (1997) states in her study that if the students are provided with "the assessment type they prefer from among those types of assessment considered appropriate for a given purpose, the perceived validity of the assessment will improve, thus motivating them to perform at their best" (81). That is, the assessment preferences of the students, as explained above, have some effects not only their learning preferences but also their achievements.

The Present Study

Although there are several studies on the effects of formative assessment, there are not many studies focusing on its effects on higher education students' assessment preferences. Especially the main concern of the current study is on ELT undergraduates' (possible) assessment preference changes in terms of formative assessment, self/peer assessment, and feedback preferences. Baeten, Dochy, and Struyven (2008) explain in their articles hands-on formative assessment practices have effects on students' assessment preferences. Therefore, it is aimed in this study to find out if implementing formative assessment causes any changes in students' assessment preferences.

METHODOLOGY

The study is based on a constructivist theory which argues that people produce knowledge and meaning from their experiences (Piaget 2001). The theory of constructivism is vital for formative assessment because it conceptualizes learning as an active process, which builds on students' own understanding. Meanwhile, the central principle of this study is that students should be involved in their own assessment process, and as a result they may increase their understandings and add new types of assessments to their preferences.

To provide this aim, the mixed methods were included using quantitative and qualitative strategies aiming to collect and analyze both forms of data in a single study (Creswell 2003). Being aware of that all methods have their own limitations, researcher felt that triangulating data sources- through using quantitative and qualitative methods- would lessen the danger of using only one source with its biases.

Participants

The participants of this study were the freshmen students of Cukurova University English Language Teaching (ELT) Department. The participants consisted of thirty eight (38) experimental group students and forty eight (48) control group students, and eighty six (86) students in total. The ELT department has thirteen freshmen classes. Both experimental and control groups have two classes, meanwhile, four classes in total took part in this study. The experimental group consists of nine boys and twenty nine girls; on the other hand, the control group has eight boys and forty girls. Their ages vary from eighteen to twenty.

Two classes for experimental group and two classes for control group (four classes in total) were randomly chosen out of thirteen classes because with random selection or random sampling, each student has an equal chance to be selected from the population, guaranteeing

that the sample will be representative of the population (Creswell 2003). All of the students have been learning English since 4th grade of primary school, which means for at least 8 years in total. Besides, both experimental and control group students had one year preparatory class in their first year of the university.

The formative assessment treatment was applied in “Contextual Grammar” course because this course deeply embedded in language usage, and rhetorical grammar, which requires using and producing language. Therefore, the students were not feeling so secure about this course, and had fears about the assessment of the course.

Data Collection Tools

Assessment Preference Scale was applied before and after the formative assessment implementation. It helped the researcher to find out what kind of assessment preferences the students had, and why they had those preferences; then at the end of this process, this scale showed the changes, if any, in the students’ assessment preferences.

The steps of developing this scale are as follows:

1. The relevant literature was examined, and 16 items related to assessment preferences were composed. Three assistant professors in English Language Teaching Department at Cukurova University and two experts in measurement and evaluation were asked to examine the comprehensibility and sufficiency of the items. It consisted of a 5-point scale where *1 = never*, *2 = rarely*, *3 = sometimes*, *4 = often*, and *5 = always*.
2. The scale was applied to Cukurova University English Language Teaching Department freshman students ($n = 107$) aging from 18-20 in 2008-2009 educational year.
3. Factor analyze was conducted in order to determine reliability of the scale. After these steps, this data collection tool was determined to be a 16 item scale.

The analysis showed that the scale had a high reliability (Cronbach’s $\alpha = .84$). This scale has two main categories:

a) *Traditional Assessment Preferences* including the sub-categories as:

- *selected response tasks (multiple-choice, matching)*,
- *production tasks (open-ended questions, explanatory questions)*,
- *limited-production tasks (rewrite, clause test, combining sentences)*,

b) *Formative Assessment Preferences* including the sub-categories as:

- *self/peer assessment*,
- *feedback preferences (verbal, written)*,
- *ongoing assessment preferences (weekly quizzes, classroom activities, take-home task)*.

The highest assessment preference score that can be obtained for each category is calculated by multiplying all the items of that category with 5, which is the maximum anxiety score given for *always*. For example, the first category (traditional assessment preference) has seven items and it can have thirty five as the maximum score for its assessment preference level.

Additionally, as a data collection tool, semi-structured interviews were conducted with randomly selected seven students of the experimental group before and after the formative assessment implementation. The reason for using semi-structured interview is that it is more flexible than standardized methods such as the structured interview or survey. Although the researcher in this study had some established general topics for examination, this method allowed for the exploration of rising ideas rather than just relying only on thoughts and questions defined in advance of the interview.

The questions were asked in a similar order and format to make a form of comparison between answers possible. On the other hand, there was also scope for pursuing and investigating for new and related information, through additional questions. The followings are the two examples of the interview questions:

- 1- When you were in high school, how did your teachers assess you in your courses? (Pre-test)
- 2- During this semester you had different assessment activities including your midterm exams and classroom assessments. What kinds of assessment types were useful for you? (Post-test)

The excerpts taken from students' interviews are not corrected. The interview transcripts were analyzed through content analysis that "arose out of the approach known as grounded theory" (Burnard, Gill, Stewart, Treasure, and Chadwich 2008, 429).

Lastly, field notes and teacher observations helped the researcher to get more detailed information about formative assessment practices in the classroom. During formative assessment implementation, the teacher usually observed the students while they were busy with self and/or peer assessment activities, and it yielded a description and analysis of teacher's informal assessment practices in the class. The data collected through teacher observation was recorded as the field notes of the researcher.

Main Study

The focus of the study was mainly on the formative assessment applications in the classroom. The main study application took place during fourteen weeks. The first step was to prepare and hand out checklist covering the topics of the week. The aim for using checklist was that the use of these checklists keeps students on track and allows them to take responsibility for their own learning through peer/self-evaluation. Also, the checklists clearly communicate performance expectations in terms of criteria and standards (Emery, Harvey, & Andersen 2006). In each week or for each chapter, students were given a checklist to let them see if they had any idea about that day's topic.

The second application was sharing learning goals of that week's topic or chapter. This helped them realize what they knew and what they need to know, which means the gap between pupils' current and desired level of learning or performance (Black & Wiliam 1998a). This target sharing activity made the curriculum more transparent for students, and helped them to realize what was expected from them.

The third step in formative assessment application was the feedback, both verbal and written. Black and Wiliam (1998a) consider assessment as formative when it uses feedback to adjust teaching and learning activities. Therefore, the instructor not only offered feedback to students for their learning, but he also got feedback from students to adjust his own teaching.

The next assessment application was to include students into their own assessment (self assessment), and peer assessment. Crooks (2001) states that that feedback on assessment cannot be effective if students do not accept that their work can be improved and recognize important

features of their work that they wish to develop. In classroom students were asked to evaluate their own work and also their classmates' work and correct their mistakes, mostly after a pop up quiz or an exercise.

DATA ANALYSIS AND DISCUSSIONS

This part gives an overall result between the experimental and control group students' mean scores of their traditional and formative assessment preferences.

Traditional Assessment Preferences

According to the pre-test means of the experimental and control groups ($X_{exp}= 22.70$, $X_{cont}= 22.66$), there was not a statistically significant difference between the two groups in total ($p= .971$) at the beginning of the study (Table 1). The pre-test results of the two groups show that the groups' traditional assessment preferences can be considered high. In other words, both groups highly prefer this type of assessment. This might be due to the fact that these students were almost always assessed via multiple-choice type exam or other traditional assessment types in their high schools.

Table 1. Traditional Assessment Preferences Between The Groups

PRE-TEST			POST-TEST			
	Mean	P	Mean	P		
	Experimental	Control	Experimental	Control		
TOTAL	22.70	22.66	.971	21.68	22.68	.326

Similar to pre-test results, the means of the two groups' post-test ($X_{exp}= 21.68$, $X_{cont}=22.68$) did not show a statistically significant difference in their post-test results ($p= .326$) (Table 1). When examined, the post-test means are still high. It can be concluded that both group students still prefer traditional type of assessment at a high rate. Assessment types used during their high school years have an undeniable effect on these students' traditional assessment preferences, and as they are freshman students they need more time for renovation in their assessment preferences.

What might cause this stability is that these freshmen students were selected for university by a university entrance exam that is totally based on selected response items, mostly multiple choice questions. Moreover, they probably studied for this exam more than 3-4 years. Therefore, almost all students still find it easier to find or recall the correct answer from the given choices. The students' selected response preferences are also clear in their interview conducted both before and after the study. In the whole paper, these student responses are quoted directly that includes non-native errors. Followings are two excerpts from two of the students' pre-test interview:

ST4: "In the high school, we have been tested by YDS exam. Our teachers gave us 100 questions. We solved them. They assessed us according to this exam. This exam was multiple choice. I want to be assessed by multiple choice."

ST6: “I think it can be multiple choices. For example, vocabulary tests, structures, etc. having a, b, c, d, e...”

As clear from this interview transcript, the students were assessed only by multiple choice exams in their high schools, and this type of assessment was used for preparing the student for their university entrance exam. Another student stated similar ideas when he was asked about the changes in his ideas about type of the assessment tasks at the end of the formative assessment application:

ST2: “I do not have much fears but multiple choice was easier than others.”

Both the quantitative and qualitative data revealed that students, when in high school, were assessed mostly by multiple-choice test format for getting prepared for university entrance exam. Certainly, the reason for using such tasks was that all the students had to prepare for this high-stake test. The gate-keeping role of this test gives no other choice to teachers but assessing their students in this way. This yields students’ getting used to these selected response tasks; other types of assessments, which may result in a better understanding and learning, seem hard, and/or unnecessary for the students.

Formative Assessment Preferences

This part gives an overall result between the experimental and control group students’ mean scores of their formative assessment preferences.

Table 2. Formative Assessment Preferences between the Groups

PRE-TEST		POST-TEST					
Mean		P		Mean		P	
Experimental		Control		Experimental		Control	
TOTAL	26.87	27.12	.857	30.60	25.95	.002	

According to the pre-test means of the two groups ($X_{\text{exp}}= 26.87$, $X_{\text{cont}}= 27.12$), at the beginning of the study, there was not a statistically significant difference between the two groups in total ($p= .857$) in their formative assessment preferences. Contrary to the pre-test results presented in Table 2, there is a statistically significant difference in experimental and control group means in the post test results ($p= .002$). While the experimental group shows a certain increase in their post-test means ($X_{\text{pre}}= 26.87$, $X_{\text{post}}= 30.60$), the control group shows a decrease ($X_{\text{pre}}= 27.12$, $X_{\text{post}}= 25.95$).

The decrease in the means of control group in Table 2 may be explained with measurement error. When the analysis was done between the experimental and control groups, the SPSS program accepted the number of the both groups as 16 for pre-test and 20 for post-test. However, the numbers were accepted as 17 for experimental group and 21 for control group when the analysis was conducted within the groups as can be seen in Table 4.26.

Table 3. Formative Assessment Preferences within the Groups

		N	X	sd	p
Experimental Group	Pre-test	17	26.52	1.07	.004
	Post-test	17	30.11		
		N	X	sd	p
Control Group	Pre-test	21	25.76	1.12	1.000
	Post-test	21	25.76		

When the formative assessment preferences are analyzed within the groups in Table 3, the mean experimental group ($X_{pre}= 26.52$, $X_{post}= 30.11$) shows a sharp increase whereas the control group ($X_{pre}= 25.76$, $X_{post}= 25.76$) remains the same at the end of the study, and this increase in the mean of the experimental group is statistically significant ($p= .004$).

Although both groups had similar ideas at the beginning of the study, the experimental group students' preferences about formative assessment tasks moved to a higher level. In other words, their attitudes to this category developed in a better way after the formative assessment application. This high increase in experimental group students' formative assessment preferences may be due to several reasons. Firstly, one of the main activities of formative assessment was in the perception by the learner of a gap between the desired goal and his/her present state. This gap filling activity was provided by the instructor through communicating the targets of the topics of the day.

Another reason was the action by the learner to close that gap to attain the goal that the teacher shared at the beginning of the lesson. For this purpose, the key action was carried out by the students, which were marking their own work and raising questions about the assessment and the material covered by the assessment. This self assessment process created awareness in students' learning.

The third reason for this increase in experimental group students' formative assessment preferences was formative feedback. By direct and immediate feedback, not only the teacher got a view of both individual and class performances but the students also learned how well they did. The teacher could judge the students' success and plan adjustments based on the formative feedback obtained from the students. Moreover, summarized formative feedback provided a basis for the teacher to reexamine topics in the unit and adjust his teaching if necessary. In addition to benefits of formative feedback for the teacher, students could also see progress in their learning. Namely, both teachers and students learned from the feedback results.

Since students became active in their own assessment, they were motivated to learn, which, in turn, resulted in a less threatening learning and assessment environment. This anxiety-free environment is also clear in students' interviews. Following excerpt taken from a student's first interview is an example related to this issue:

ST3: "The first time I entered your course, I was not sure about how can I do that. I felt the first time this one. But the other classes like speaking I can do, but this one requires more energy than other ones. That's why I have fears. I want to be assessed by multiple choice. You know, our behaving in the class, attendance in the class, could be very nice. "

The same student expressed how he got rid of this fear at the end of the study:

ST3: "We used...like weekly quiz, we matched the words. It was very good because we learned how to match the words, close sentences each other. And I learned a lot of things from

it because in multiple choices we only choose the answer. We do not give anything, we do not focus on. But in these, we learn more than multiple choice. Actually, I do not have any fears anymore, because, at the beginning, I did not know anything about the lesson. That's why; if I do not know anything about the lesson, I was afraid. Now I learned how to match, how to make sentences. I learned all of them; that's why, I do not afraid now."

As can be seen from the example above, the student had fears about both the course itself and its assessment at the beginning of the study, and preferred to be assessed by only multiple choice tests. After the application, he began to understand that there were much better assessment activities than multiple choices tests for their own learning. Similarly, another student stated in his second interview that all those formative assessment activities were more enjoyable and useful for them because they could understand the lesson better:

ST6: "Actually, with peer assessment, I can understand lesson better. It will be better if I continue like that. I like evaluating their situation and my situation. It improves me to know structures in language. To write good sentences, to combine sentences with conjunctions or coordinating conjunctions... It helps me do less mistakes...I get many feedbacks from my teacher, from my friends even. These feedbacks were useful for me because I knew that I won't have any problems about lesson, about the assessments you have given us. And so they were all useful for me...Actually there were many changes in my education, especially in this lesson. At the beginning of this semester, I had difficulty in understanding the lesson, I could not concentrate on the structures... but now I have less difficulty and I can write sentences better."

As we can understand from the statistically significant data given in Table 4.26 and the interviews, the students' preferences about the formative assessment seem to be very useful for them. Their assessment preferences moved from a multiple choice test format to a more learning-based assessment method, formative assessment. They clearly preferred this type of assessment tasks because they felt that they learned more and felt less frightened about the assessments. More specific analysis and discussions about three categories (*self/peer assessment, feedback, and ongoing assessment preferences*) related to formative assessment task preferences are done in the following sections.

Self/Peer Assessment Preferences

This part includes two assessment tasks: *self/peer assessment*. Therefore, the maximum total score of the two groups in this section may be maximum 10 at most.

Table 4. Self/Peer Assessment Preferences within the Groups

		N	X	sd	p
Experimental Group	Pre-test	24	6.41	2.10	.000
	Post-test	24	8.33		
		N	X	sd	p
Control Group	Pre-test	40	6.25	2.04	.645
	Post-test	40	6.10		

The analysis of self/peer assessment preferences of the experimental group given in Table 4 reveals that there is a significant difference ($p = .000$) between the pre and post-test means ($X_{pre} = 6.41$, $X_{post} = 8.33$). The increase surely means that the students of experimental group found self and peer assessment useful and began to prefer more in the end. The means of the control group, on the other hand, ($X_{pre} = 6.25$, $X_{post} = 6.10$) do not show a statistically significant difference ($p = .645$).

This high increase in experimental group's self/peer assessment preferences may be due to in-class self/peer assessment activities. For example, after each topic was taught in the classroom, students were required to assess their own learning through some ways like checklists given at the beginning of the lesson or by evaluating their own learning through exercises of the topic. The checklists covered the topics and goals of that lesson were given to the students at the beginning of the lesson.

The students were expected to evaluate what they knew about that day's topic. Then, at the end of the lesson they were asked to look at the same checklist to control what they learned and what could not understand. This self assessment process also helped the teacher to observe students' learning process because while the students were assessing their own learning, he had the chance to review their understanding individually.

Also, the students tried to assess their peers or classmates. This gave them a lot of opportunities to correct their own mistakes while trying to correct their peers'. At the beginning, the students were uninterested and unwilling for this practice because they were not sure about getting the ideas of their classmates or peers about their own work or learning, but during this period the students felt less threatened, and they became more involved and began to enjoy assessing their peer's learning. The following interview excerpts taken from two of the students in the second interview reveal this involvement process in students' self and peer assessment:

ST3: "I believe peer assessment is very useful for me because I get knowledge from it. I let my friends assess me because I can learn my wrong usage of verb from them and he gives me how to correct them. And it was very good for me, and self assessment we made in the classroom was very good. Because we let ourselves control us and what we know. We see our knowledge, and we can use them in the writing. If we learn, we should use them in the writing. It will stay life-long in our memory. "

ST4: "I think self-assessment is useful for me because I can look my error, or peer assessment is useful too because I can look my friend's error or my friend can look my error. He says me and I correct my error."

As it is clear from the interview data, students find self and peer assessment useful because they helped them to raise their awareness on their own strengths and weaknesses. More to the point, the process of self/peer assessment yielded an increased engagement with learning and assessment process. They helped the students especially set goals, clarify objects, take responsibility for their own learning, and thus increased students' self confidence. Through self and peer assessment, students were able to see the mistakes in their thinking and could correct them for future assignments. By assessing their own papers, students were better in understanding the assessment process and could recognize their own strengths and weakness. Students learned how to think while completing assignments.

Students' Feedback Preferences

Feedback is one of the key elements of formative assessment. Feedback is the information which can be written or oral and should focus on helping pinpoint areas of strengths and weaknesses. The feedback preferences of the study are as follows: *feedback, verbal feedback, written feedback*. The maximum total score of the experimental and control groups related to feedback assessment preferences may be maximum 15 because this part includes two items, each of which may be 5 at most.

Table 5. Feedback Preferences within the Groups

		N	X	sd	p
Experimental Group	Pre-test	21	10.71	2.57	.033
	Post-test	21	12.00		
		N	X	sd	p
Control Group	Pre-test	25	10.96	2.48	.873
	Post-test	25	11.04		

Like self/peer assessment preferences, the experimental group's feedback preference means ($X_{pre}=10.71$, $X_{post}= 12.00$) indicate a significant difference ($p= .033$) at the end of the study (Table 5). The means of the control group pre and post-test are almost same ($X_{pre}= 10.96$, $X_{post}= 11.04$) in both tests, and do not show a significant difference ($p= .873$). The sharp increase in the means of experimental group shows that they began to prefer getting different types of feedback from their teachers and classmates.

One of the possible reasons for this increase of the experimental group might be that students regularly got written and mostly verbal feedback from the instructor during classroom activities. This feedback about students' learning was always constructive and encouraging. After the students' benefits of such feedback, they began to prefer getting both written and verbal feedback for their work.

The feedback offered during classroom sessions was in two ways: external and internal feedback. The external feedback is the information about students' work provided both by their peers and instructor. This feedback provided additional information which helped students to reexamine their knowledge about the topics. Teacher feedback served as a reliable external reference point which evaluated their progress and their own internal goals. The information provided by the teacher was in a dialogue form rather than information transmission. This dialogue with the teacher helped students to develop their understanding of expectations and standards, and to correct their mistakes.

Another source of external feedback was the classmates of the students. Peer dialogue was beneficial for students in some ways. First of all, students who just learned something were generally better than the teacher in explaining it to their classmates in a language which was accessible. In other words, the students were sometimes more capable of explaining what needed to be corrected than the teacher. Second, peer discussion showed the students alternative perspectives and tactics on problems. Those alternative views enabled students to revise or reject their initial premise and construct new knowledge and meaning through negotiation. Thirdly, by commenting on their peer's work, students developed objectivity of judgment (about their work in relation to goals or standards) that could be transferred to the assessment of their own work. Lastly, peer discussion was motivating and easier for students to accept

critiques of their work from peers than the teacher. In the second interview with one of the participants, for example, it can be seen that the students felt more comfortable during peer assessment that provided them feedback:

ST7: ...”Peer assessment is good because when I show my writing to my friends, when they show my mistakes I feel more comfortable because I can improve in a good way. That’s why, peer assessment is really important...”

Another student stated that he liked peer assessment because it helped them increase the communication among their classmates:

ST5: “I liked peer assessment most because it is some kind of classroom activity. At lesson students do not talk much each other, and I think; it was useful for our class. And peer assessment improved our communication with our friends, also, help our teacher improved our learning.”

Internal feedback, on the other hand, is the information provided by the students themselves. During classroom sessions, the students did not only get feedback from their instructor and their classmates, but they also got feedback from self assessment activities. This self-generated feedback led to a reinterpretation of the task and to adjust their internal goals or strategies. Following excerpt taken from a student in the second interview supports this issue:

ST3: ... “and self assessment we made in the classroom was very good. Because we let ourselves control us, what we know. We see our knowledge, and we can use them in the writing. If we learn we should use them in the writing. It will stay life-long in our memory.”

As it is clear from students’ excerpts, the feedback offered from the teacher and the peers were useful for the students, and helped them to internalize the objectives identified by the teacher at the beginning of the lesson. Besides, the formative feedback was also useful for the teacher to do necessary adjustments in his teaching.

CONCLUSION

The current study aimed to find out if implementing formative assessment would yield to any changes in students’ assessment preferences. From the analysis of the traditional assessment preferences of the students, it can be concluded that both the experimental and control group students still prefer traditional type of assessment at a high rate. This is because of strong effect of the assessment tasks used during their high school years, and as they are freshman students they need more time for renovation in their assessment preferences. Another reason for preferring this type of assessment was that the students had to take a university entrance exam, the high-stake test they had to take to be a university student, which is mostly based on traditional (summative) assessment and does not directly measure the ability to produce language. As Stoneman (2005) states that these kinds of test results are still the main criteria for students’ promotion or being accepted to educational institutions. For this reason, the gate-keeping role of this test gives no other choice to teachers for assessing their students, and this type of assessment yields students getting used to this selected response tasks. Then again, other types of assessments, which may result in a better understanding and learning, seem hard and unnecessary for the students.

Additionally, the students studied for this high-stake test for three or four years, which means that all they were directed to choose from the given options during this period. Although students find some other assessment types useful for their personal achievement, they still showed a tendency for traditional type of assessment. This situation is also stated by Black and Wiliam (1998a), “they spend time and energy looking for clues to the *right answer*” (143). Namely, as a result of this powerful exam, instead of producing any kind of language, the students prefer to be assessed through multiple-choice tests highly even after this formative assessment practice period.

The analysis of the *formative assessment preferences* of the experimental group students indicated that students' preferences changed towards formative assessment. In other words, they added new types of assessment to their preferences. This change had some causes: firstly, one of the main activities of formative assessment was sharing the goals with the learners, which aimed to closing the gap between the desired goal and the students' present state. This gap filling activity was provided by the instructor through communicating the targets of the topics of the day. Similar to the findings Lynch and Maclean's in their study (2003), the present study found that the goal setting and target sharing activity created awareness in students' learning and assessments.

Another reason for this increase in experimental group students' formative assessment preferences was the feedback. According to Ramaprasad (1983), feedback is ‘information about the gap between the actual level and reference level of a system parameter which is used to alter the gap in some way’ (p. 4.; in Sadler, 1989, p. 120). Similar to findings of Çakır and et al. (2016), the way for the students to close the gap during this study was the feedback from their peers and instructor, which enabled them to revise or reject their first idea and construct new knowledge and meaning through negotiation.

The present study's findings echo in Narciss and Huth's (2004) study where they put forward that systematic feedback had positive effects on students' motivation. Moreover, summarized formative feedback provided a basis for the teacher to reexamine topics in the unit and adjust his teaching if necessary. As in the study of Sivaci (2020), who concluded that students have benefitted from peer feedback, students learned from the feedback results in this our study.

From the analysis of assessment preference scale, another significant change was found to be in students' *self/peer assessment* preferences. As Black and et al. (2003) pointed, self/peer assessment is exclusively important in students' learning; these types of assessment were found to improve student motivation to work carefully. In line with their study, the self/peer assessment activities during formative assessment implementation in this study had a positive effect on students' assessment preferences because they found this type of assessment useful:

STI: “I think self-assessment is useful for me because I can look my error; or peer assessment is useful too because I can look my friend's error or my friend can look my error. He says me, and I correct my error.”

To sum up, the process of self/peer assessment resulted in an increased engagement with learning, and it also increased students' confidence. During self and peer assessment, students were able to get a view of their thinking and mistakes; then, they could correct them for future assignments. By assessing their friends and own papers, students were better in understanding the assessment process and could recognize their own strengths and weakness.

SUGGESTIONS FOR CURRICULUM AND TEACHING

As formative assessment has been accepted and proved to be a very powerful way of enhancing student learning through various assessment tasks (Black & Wiliam 1998a; Black & Wiliam 1998b; Brookhart 2007; Butler & McNunn 2006), in an effort to distill the quality of classroom assessment, students' learning, teachers' professional improvements, the following statements are the suggestions derived from the current research for a better student learning:

- There should be a link between curriculum, instruction and assessment, which can be clearly understood by the students. Squires (2004, p. 4) stated in his book that “the curriculum, the curriculum-embedded assessment, and the instruction are aligned with each other. Assessments answer the question of how much knowledge and skill are good enough to meet the standards aligned in the unit. Teachers use assessments to determine how good is good enough. Classroom assessment is inexorably linked to the curriculum.” Students should sufficiently be informed about this alignment of curriculum with assessment.
- Teachers should have a consistent and ongoing plan for their professional development. Assessment of learning does not only aim to enhance learning of students in the classroom, but it also aims to redesign the instruction and the curriculum according to assessment results of the classroom. Thus, the results obtained from classroom assessment should also act as a mirror for teachers to make necessary changes in their teaching methods and the materials they use.
- As the students are one of the main actors of teaching-learning process, teachers should share the learning goals with students. This sharing will yield to increase student learning in basic knowledge and higher order cognitive processes such as application and transfer (Fulmer, 2017). She also states that when students know that they are expected to learn in that lesson, they will focus more on those areas. With clear goals aligned with the assessment, students will spend less time for the search of what to learn in that lesson. Moreover, discussing the assessment questions critically with the learners will yield to more reflection on what is being assessed. As assessment can be used a tool for a better learning, this discussion of assessment tasks and the answers may certainly be great way of teaching.
- Instead of over-stressing the grading function, teachers should focus more on the learning functions of assessment. As the main aim of grading is to evaluate individual students' learning gains in the classroom, it does not really add to a better learning. However, assessment generally goes far beyond grading because it systematically examines student teaching/learning process during the course, and uses this information for improving the educational process (both teaching and learning process) in the classroom.
- Formative assessment aims to identify and eliminate the learning deficiencies and difficulties of the students and to increase their learning while the education continues (Baird, Andrich, Hopfenbeck, & Stobart 2017). For an assessment to be formative, the assessment should be used as feedback for closing the gap between what the students know and what else they need to learn (Vogt, Tsgari, Csepes, Green, & Sifakis, 2020). “For example, a formative assessment technique could be as simple as a teacher asking students to raise their hands if they feel they have understood a newly introduced concept, or it could be as sophisticated as having students complete a self-assessment of their own writing (typically using a rubric outlining the criteria) that the teacher then reviews and comments on” (the Glossary Of Education Reform).

As Black and William (1998a) characterize assessment broadly to include all activities that teachers and students carry out to get information used diagnostically to alter teaching and learning, teachers can create many opportunities to assess how students are learning, and then use this information to make constructive changes in instruction. Meanwhile, if teachers do not use tests and assessment tasks to give feedback about learning, they are just the indicators of a final summative test.

LIMITATIONS OF THE STUDY

There are some limitations of this study. The first one is that our entire study group was the freshmen students of English Language Teaching Department, and they took the same university entrance exam and were all almost the same age group. However, those individual variables such as age, sex, and socio-economic and cultural factors were not taken into consideration. Due to the fact that they took the same exam to be ELT department students, it was assumed that the students would perform similarly.

Another limitation is that the study focused on only four freshman classes at Cukurova University whereas there were thirteen classes during the study, which means that it would give a clearer picture if all the freshman class students could have been used as the participants in this study. Thirdly, the study was conducted with only one teacher, the researcher himself. However, the students had eight different courses and teachers. It would be more useful if this study could have been conducted in students' other courses to see the effects of formative assessment on students' assessment preferences. The last limitation is the period of the implementation. Since it is the process that is significant in this study, longer period of formative assessment application would have been appropriate to find results that would reflect in greater accuracy.

REFERENCES

- Alderson, J. C, Brunfaut, T., & Harding, L. (2017). Bridging assessment and learning: a view from second and foreign language assessment. *Assessment in Education: Principles, Policy & Practice*, 24(3), 379-387, DOI: 10.1080/0969594X.2017.1331201
- Allal, L. & G. P. Ducrey. (2000). Assessment of-or in- The Zone of Proximal Development. *Learning and Instruction*, 10, 137-152.
- Bartram, B. & C. Bailey. 2010. Assessment preferences: a comparison of UK/international students at an English university. *Research in Post-Compulsory Education*, 15(2), 177-187.
- Baeten, M., F. Dochy, and K. Struyven. (2008). Students' Approaches to Learning An Assessment Preferences In A Portfolio-Based Learning Environment. *Instr. Sci.* 36, 359-374.
- Baird, J., Andrich, D., Hopfenbeck, T. N., & Stobart, G. (2017). Assessment and learning: fields apart?, *Assessment in Education: Principles, Policy & Practice*, 24(3), 317-350, DOI: 10.1080/0969594X.2017.1319337
- Black, P. & D. Wiliam. (1998a). Inside the Black Box-Raising the Standards through Classroom Assessment. *Phi Delta Kappan*. 80(2), 139-148.
- Black, P. & D. Wiliam. (1998b). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74.

- Black, P., C. Harrison, Lee, C., B. Marshall & D. Wiliam 2003. *Assessment for Learning: putting it into practice*, New York, McGraw-Hill.
- Birenbaum, M. (1997). Assessment preferences and their relationship to learning strategies and orientations. *Higher Education*, 33, 71–84.
- Birenbaum, M. (2007). Assessment and Instruction Preferences and Their relationship with Test Anxiety and Learning Strategies. *Higher Education*, 53, 749-768.
- Brookhart, S. M. (2007). Expanding Views about Formative Classroom Assessment: A Review of the Literature. In J. H. McMillan, *Formative Classroom Assessment: From Theory to Practice* (p. 116-136). New York: Teachers College Press.
- Brown, H. D. (2004). *Language Assessment-Principles and Classroom Practices*. New York: Pearson Education, Inc.
- Burnard, P., P. Gill, K. Stewart, E. Treasure, and B. Chadwick. (2008). Analyzing and presenting qualitative data. *British Dental Journal*, 204(8), 429-432.
- Butler, M. S. & N. D. McMunn. (2006). *A Teacher's Guide to Classroom Assessment: Understanding and Using Assessment to Improve Student Learning*, San Francisco: Jossey-Bass.
- Çakır, R., Korkmaz, Ö., Bacanak, A., & Arslan, Ö. (2016). An Exploration of the Relationship Between Students' Preferences For Formative Feedback and Self-Regulated Learning Skills. *Malaysian Online Journal of Educational Sciences*, 4(4), 14-30.
- Cresswell, J. W. (2003). *Research Design: Qualitative, quantitative, and mixed method approaches (2nd edition)*. USA: Sage Publications.
- Cowie, B., & B. Bell. (1999). A model of formative assessment in science education. *Assessment in Education*, 6. 101-116.
- Crooks, T. (2001). The Validity of Formative Assessment. *British Educational Research Association Annual Conference*, University of Leeds. Retrieved November 19, 2009 from <http://www.leeds.ac.uk/educol/documents/00001862.htm>
- Derrich, J. & K. Ecclestone. (2006). Formative Assessment in Adult Literacy, Language and Numeracy Programmes. *A Literature Review for the OECD*, Draft.
- Earl, L. (2003). *Assessment as Learning: Using Classroom Assessment to Maximize Student Learning*, Thousand Oaks, CA: Corwin Press.
- Emery, L. J., Harvey, H. & Andersen, M. C. (2006). Formative Evaluation Using Checklists to Improve Research Papers. *Perspectives in Health Informative Management*, 3(2), Retrieved April 13, 2010 from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2047304/>
- Erwin, T.D. (1991). *Assessing Student Learning and Development*. Jossey-Bass, Retrieved October 17, 2008 from <http://www.qualityresearchinternational.com/glossary/assessmentoflearning.htm>
- Forbes, E. W. (2007). *Improving the Knowledge and Use of Formative Assessment: A Case Study of A Model of Formative Assessment in A K-3 Science Curriculum*. Unpublished Doctoral Dissertation. (UMI No: 3267194).
- Fulmer, S. M. (2017). *Should we share Learning Outcomes / Objectives with students at the start of a lesson?* Online document, retrieved from <https://khsbpp.wordpress.com/2017/10/12/should-we-share-learning-outcomes-objectives-with-students-at-the-start-of-a-lesson/>
- Giebelhaus, C. & C. Bowman. (2002). Teaching Mentors: Is It Worth the Effort? *Journal of Educational Research*, 95(4), 246-254.
- Gijbels, D., M. Segers, E. Struyf. (2008). Constructivist Learning Environments And The (Im)Possibility To Change Students' Perceptions of Assessment Demands And Approaches To Learning. *Instructional Science*, 36, 431-443.

- Glossary of Education Reform. (2021). *Formative Assessment*. Online document, retrieved from <https://www.edglossary.org/formative-assessment/#:~:text=Formative%20assessment%20refers%20to%20a,lesson%2C%20unit%2C%20or%20course.>
- Hancock, C. R. (1994). Alternative Assessment and Second Language Study: What and Why? *Eric Digest*, Retrieved August 18, 2008 from http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/13/6f/18.pdf
- Hangstrom, F. (2005). Formative Learning. *Communication Disorders Quarterly*, 28(1), 24-36.
- Harlen, W. (2007). Formative Classroom Assessment in Science and Mathematics. In J. H. McMillan. *Formative Classroom Assessment: From Theory to Practice*, 116-136, New York: Teachers College Press.
- Lambert, D. & Lines, D. (2001). *Understanding Assessment: Purposes, Perceptions, Practices*. London: Routledge Falmer.
- Lowe, P. A., S. W. Lee, K. M. Witteborg, K. W. Prichard, M. E. Luhr, C. M. Cullinan, et al. (2008). The Test Anxiety Inventory for Children and Adolescents (TAICA): Examination of the Psychometric Properties of a New Multidimensional Measure of Test Anxiety among Elementary and Secondary School Students. *Journal of Psychoeducational Assessment*, 26(3), 215-230.
- Lynch, T. & J. Maclean (2003). *Effects of Feedback on Performance: A Study of Advanced Learners on an ESP Speaking Course*. Edinburg Working Papers in Applied Linguistics.
- McCallum, S. & Milner, M M. (2021). The effectiveness of formative assessment: student views and staff reflections. *Assessment & Evaluation in Higher Education*, 46(1), 1-16, DOI: 10.1080/02602938.2020.1754761
- Narciss, S. & K. Huth (2004). How to design informative tutoring feedback for multimedia learning. In H. M. Niegemann, D. Leutner, & R. Brunken (Ed.), *Instructional design for multimedia learning* (pp. 181–195). Munster, New York: Waxmann.
- Piaget, J. (2001). *The psychology of intelligence* (2nd Ed.), London: Routledge. Retrieved October 12, 2009 from <http://www.google.com/books?hl=tr&lr=&id=jJZ7aWY4tIcC&oi=fnd&pg=PR5&dq=The+psychology+of+intelligence&ots=3ZFbpwKbbv&sig=wT5iPnoGw8CeuziQ17vG1PIQBmA#v=onepage&q=&f=false>
- Sadler, D.R. (1989). Formative Assessment and the Design of Instructional Systems. *Instructional Science*, 18(2), 119-144.
- Shepard, L. A. (2000). *The role of classroom assessment in teaching and learning*. Center for the Study of Evaluation (CSE Technical Report # 517). University of California, Los Angeles, CA.
- Shepard, L. A. (2005). Linking Formative Assessment to Scaffolding. *Educational Leadership*, 63(3), 66-70.
- Sivaci, S. (2020). The Effects of Peer Feedback on Writing Anxiety Levels of Pre-Service English Teachers. *The Reading Matrix: An International Online Journal*, 20(2), 131-139.
- Squires, D. A. (2004). *Aligning and Balancing the Standards-Based Curriculum*. Southern Connecticut State University, USA.
- Stiggins, R. (1992). *High quality classroom assessment: what does it really mean?* NCME *Instructional Topics in Educational Measurement Series*. Module 12, Summer 1992. (Online) at <http://www.ncme.org/pubs/items/19.pdf>

- Stiggins, R. J. (2002). Assessment Crisis: The Absence of Assessment FOR Learning. *Online article, Kappan Professional Journal*, Retrieved August 12, 2009 from <http://electronicportfolios.org/afl/Stiggins-AssessmentCrisis.pdf>.
- Stoneman, B. W. H. (2005). *The Impact of an Exit English Test On Hong Kong Undergraduates: A Study Investigating The Effect Of Test Status On Students' Test Preparation Behaviors*. Unpublished PhD Dissertation, the Department of English, the Hong Kong Polytechnic University.
- Struyven, K., F. Dochy, & Janssen, S. (2005). Students' Perceptions About Evaluation And Assessment In Higher Education: A Review. *Assessment and Evaluation in Higher Education*, 30(4), 331–347.
- Taras, M. (2005). Assessment- Summative and Formative- Some Theoretical Reflections. *British Journal of Educational Studies*, 53(4), 466-478.
- 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.
- Torres, J. O. (2019). Positive Impact of Utilizing More Formative Assessment over Summative Assessment in the EFL/ESL Classroom. *Open Journal of Modern Linguistics*, 9(1), 1-11.
- Vogt, K., Tsagari, D., Csépes, I, Green, A., & Sifakis, N. (2020). Linking Learners' Perspectives on Language Assessment Practices to Teachers' Assessment Literacy Enhancement (TALE): Insights from Four European Countries, *Language Assessment Quarterly*, 17(4), 410-433, DOI: 10.1080/15434303.2020.1776714
- Watering, G., Gijbels, D., Dochy, F., Rijt, J. (2008). Students' Assessment Preferences, Perceptions of Assessment and Their Relationships to Study Results. *High Education*, 56, 645-658.
- Wojtczak, A. (2002). *Glossary of Medical Education Terms*. Online document, retrieved May 13, 2009 from <http://www.iime.org/glossary.htm>

Kagan Buyukkarci is an Associate Professor and Head of School of Foreign Languages at Suleyman Demirel University, where he has been teaching and researching since 2011. His main research area is foreign language assessment.

Email: kaganbuyukkarci@sdu.edu.tr

Sehnaz Sahinkarakas is a Professor and Vice-Chancellor at Cag University, Turkey. Her main research area is language assessment.

Email: sehnazkarakas@cag.edu.tr