



DOI: 10.18039/ajesi.1219007

Evaluation of the Effect of an In-Service Training Program on Assessment for Learning for English Language Teachers¹

Manolya TUNÇER², Ali MERÇ³

Date submitted: 14.12.2022

Date accepted: 17.07.2023

Type⁴: Research Article

Abstract

The aim of this paper is to present the effect of an In-Service Training (INSET) Program designed to develop *Assessment for Learning* (AfL) practices on English language teachers. For this purpose, English language teachers working in the state high schools attended the five-session-long INSET program in 2020-2021 Spring semester. Whether there were any changes between teachers' responses to the tasks assigned before and after the sessions and what their opinions were related to the training were examined. Data collection tools included a task prepared for each session and a questionnaire including four-point Likert type questions and open-ended items. Descriptive statistics were used in the analysis of the quantitative data while deductive coding was employed to analyze the qualitative data. The results demonstrated that most of the teachers thought that AfL was 'quite important'. The findings also yielded that the first three items about what changes were detected most were identified as 'providing more opportunities to express their opinions', 'clarifying learning intentions in a detailed way', and 'using clear comments'. The outcomes shed light on the participant teachers' ideas about the differentiations between their first and second responses to the same tasks, and most teachers reported variations between pre- and post-task responses as 'partially changed' or 'changed a lot'. The techniques mentioned during the training were rated as 'practical' or 'partial practical' by a great number of teachers. The findings suggested that English language teachers should be given more chances for taking part in in-service training programs, especially on AfL; in this way, it can be provided that they can use the new strategies in English language classes efficiently.

Keywords: Assessment for Learning, EFL Teachers, Formative Assessment, In-Service Training Program, Language Assessment

Cite: Tunçer, M., & Merç, A. (2023). Evaluation of the effect of an in-service training program on assessment for learning for english language teachers. *Anadolu Journal of Educational Sciences International*, 13(2), 435-458. <https://doi.org/10.18039/ajesi.1219007>



¹ This study was produced from the doctoral dissertation titled "Developing an In-Service Training Program on 'Assessment for Learning' for English Language Teachers" that was carried out by the first author supervised by the second author.

² (Corresponding author) Dr., Afyon Kocatepe University, Faculty of Science and Literature, Department of Western Languages and Literatures, mtuncer@aku.edu.tr, <https://orcid.org/0000-0002-8406-9595>.

³ Assoc. Prof. Dr., Anadolu University, Faculty of Education, Department of Foreign Language Education, amerc@anadolu.edu.tr, <https://orcid.org/0000-0003-3061-7899>.

⁴ This research was conducted with Research Ethics Committee approval of Anadolu University, dated 24.06.2020 and issue number 34139.

Introduction

Assessment, in language teaching context, can be defined as to measure and evaluate language learners' knowledge and skills in an organized way for the purpose of enhancing their learning (Coombe, 2018). Language assessment is a general term that includes formal tests and other techniques such as detecting learners' performance (Purpura, 2016). Here, it is important to put an emphasis on the differentiation between two types of assessment which are formative assessment and summative assessment. Formative assessment is about everyday classroom implementations by gathering data and working on them to provide suggestions for better learning outcomes. Summative assessment provides a general idea about the learners' attainment, and this can be a requirement for completing a grade level or for having an official document such as certificate at the end of a period (Black, 1999). *Assessment for Learning* has been an outstanding topic in recent years (Wu et al., 2021).

Assessment for Learning offers the potential to reinforce students' learning (Wu et al., 2021). *Assessment for Learning* can be defined as a process that both teachers and students look for and understand the signs in determining students' current level, next steps they require to achieve, and the way they reach this point (Assessment Reform Group, 2002). This term is also explained as "part of everyday practice" (Klenowski, 2009, p. 264) with a special emphasis on the improvement of learning (Klenowski, 2009). The following one is a comprehensive definition suggested by Black and Wiliam (2009):

Practice in a classroom is formative to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited (p. 9).

Assessment for Learning has appeared as the enrichment of Formative Assessment (Assessment Reform Group, 2002; DeLuca et al., 2019; Wiliam, 2011a), and it aims to indicate students' language levels in learning, to address their strong and weak points, and to predict the points which they possibly reach (James et al., 2006). The key strategies of Assessment for Learning are as follows:

1. *Clarifying and sharing learning intentions and criteria for success*: This strategy is related to providing information about learning intentions and criteria for success and making learners comprehend these intentions and criteria (Wiliam, 2010). The duty of teacher is to establish learning intentions which indicate the points students should learn, comprehend, and succeed at the end of educational tasks, and the learning intentions can be framed for only a lesson or a sequence of lessons. Success criteria, another related term with learning intentions, elaborate the details of success and give both teacher and students a chance to assess level of learning (PDST, *Leaving Certificate Applied, Teacher Handbook*, 2022). To enable students to share their opinions on learning intentions and success criteria, teachers are advised to decide on intentions with students (Wiliam, 2011b).
2. *Engineering effective classroom discussions and other learning tasks that elicit evidence of student understanding*: It is about obtaining indications of success which is often seen as questioning (Wiliam, 2010). Assessment seeks to build a connection between teaching and learning; however, students' learning may not match with the aims of teacher (Wiliam, 2005). Waiting time, related to questioning, is also crucial for AfL, and this can be achieved

through allowing students to prepare their responses for successful classroom interaction (Stobart & Gipps, 2010).

3. *Providing feedback that moves learners forward*: Effective formative assessment has two important features: (1) it is prospective, and (2) modifications may be needed in teachers' instruction (William, 2010). According to Hattie and Timperley (2007), three important questions should be focused, and these are "Where am I going?", "How am I going?", and "Where to next?" (p.102). Teachers' responsibility is to maintain feedback according to their assessment of students' performances, and they should decide on the most appropriate ways and suggest feedback relying on their beliefs (Popham, 2011). Black et al. (2003) also call attention to the difference of providing feedback from giving mark for comparing students.
4. *Activating students as the owner of their own learning*: Self-assessment is considered as a crucial point in learning, and it requires students to become aware of the learning objectives and conditions necessary for gaining success (Black et al., 2003). Popham (2011) suggests some methods enabling the implementation of self-assessment which are "teaching metacognition" (p. 73) and displaying how they assess their performance. 'Self-regulation' is another related term offered to complete the processes with self-assessment in succeeding learning efficiently (Andrade, 2010).
5. *Activating students as instructional resources for one another*: This strategy is aimed to improve students' performances through their own assessment (William, 2018). Peer assessment is implemented on the students' studies sharing the same classroom (Topping & Ehly, 1998), and it is fulfilled in the way that quality of the studies of students is evaluated by other students in the class (Topping, 2009). By means of peer assessment, students can check out their studies once more and make some changes on them (Harrison, 2010).

Assessment for Learning has received considerable scholarly attention around the world in recent years. In 2015, the journal of 'Assessment in Education: Principles, Policies & Practices' has published a special issue on AfL implementations, and a number of researchers have examined AfL in their own contexts: Flórez Petour in Chile; Jonsson et al. in Sweden; DeLuca et al. in Canada; Wylie and Lyon in the USA; Hayward in Scotland; Hopfenbeck et al. in Norway; Ratnam-Lim and Tan in Singapore; and De Lisle in Trinidad and Tobago. The overall picture obtained were summarized by Black (2015) as that these studies described a variety of strategies to support the use of AfL in different cultural settings and the result indicated the restriction in changes. There have been other studies examining the factors that have an influence on AfL implementations (Fulmer et al., 2015; Heitink et al., 2016; Yan et al., 2021). In a similar vein, several researchers (Birenbaum et al., 2015; Swaffield, 2011) have emphasized the position that Assessment for Learning holds around the world. In these studies, one of the prominent issues has been 'teacher'. In Singapore, where there is an exam-dominated education system, Deneen et al. (2019) investigated the relations among teachers' values, practices, and proficiencies. Teachers taking part in this study were found to approve the value of formative assessment; however, at the same time, their incompetency and lack of chance to use formative assessment were also revealed. Thus, in-service training was suggested for the teachers in the study of Deneen et al. (2019). In a systematic review study, the aim of Heitink et al. (2016) was to identify the elements facilitating or impeding AfL implementations, and one category of the necessary conditions was 'teacher' in addition to the other categories: students, context, and assessment. Fulmer et al. (2015) also carried out a systematic inquiry on the contextual elements regarding three levels: micro-referring classroom, meso-out of classroom, and macro-distal one. The findings of this study shed light

on several studies at the micro level indicating teachers. In a more recent study, Yan et al. (2021) drew attention to the need for considering teachers' intention to implement AfL. As a consequence of this study, "education and training" was regarded as the primary focus to improve teachers' Formative Assessment (FA). In the same study, it was suggested to integrate FA into pre-service education and in-service training.

Over the last decades, several changes have also been observed in various educational contexts around the world. A good summary of the educational developments of AfL around the world has been provided in the study of Birenbaum et al. (2015). The case of Asia Pacific was one example where impacts of AfL could be seen, and the case of Australia was given as an example. In Australia, some regulations were enacted to enhance knowledge and skills of teachers as well as students' learning and AfL stepped forward in the way of arranging achievement standards. In Canada, AfL was identified as a key feature in assessment policies, and some arrangements were conducted for teachers to understand and practice assessment policies. In Israel, it was intended to cover AfL in teacher training programs for fostering learning. In New Zealand, to achieve AfL implementations in the classes, long-term programs were organized for teacher training by the Ministry of Education. In Norway, 'Inside the Black Box: Raising Standards through Classroom Assessment' published by Black and Wiliam in 1998 initiated the development of AfL which began to be an important issue with the PISA results. In the USA, there was a growing recognition of the importance of AfL, for instance, several meetings were arranged in the American Educational Research Association (Birenbaum et al., 2015).

A number of innovations have been made in Turkey and arrangements in 2013 curriculum can illustrate these changes (Karaman & Şahin, 2017). Learners' individual development has been at the center of the 2023 Vision prepared by the Ministry of National Education. In the field of assessment, an approach grounded on "development of the learner" has been adopted to decide on the methods to be used and how assessment should be conducted. As part of the vision, electronic portfolio system, which is grounded on individualized learning, is essential, and the purpose here is to enhance learners' competency individually (MoNE, 2018a). The present paper has aimed to investigate the effect of an In-Service Training Program designed on AfL in the Turkish EFL context. The main concern of the study is to reveal the impact of the in-service training by focusing on differences detected in teachers' answers for the same tasks before and after the training sessions and these teachers' ideas about the program. The primary focus, here, is to determine how the in-service training program influenced these teachers' practices and opinions regarding the implemented program. Thus, it is aimed to develop a better understanding of AfL implementations of English language teachers working in the Turkish high school contexts.

Problem Situation

There has been a growing interest in *Assessment for Learning*, which is considered as an enhanced way of Formative Assessment (Assessment Reform Group, 2002; DeLuca et al., 2019; Wiliam, 2011a). Although AfL has been accepted as crucial in providing powerful teaching (Black & Wiliam, 1998a), there have been some suspects identified related to the implementation of these techniques (DeLuca et al., 2019), issues about several AfL definitions

and its effectiveness (Bennett, 2011), misinterpretations of AfL (Swaffield, 2011), and variations in teachers' AfL understanding and practices (DeLuca et al., 2015). However, the crucial point here is that "the assumption that AfL can lead to learning gains for students is not disputed" (Willis, 2011, p. 5). Teachers' perception and classroom practices are important in successful implementation of FA (Yan et al., 2021). 'Personal factors' has been revealed as a category which consists of teachers' ideas, attitudes, abilities, and knowledge, and these factors are thought to influence achievement of FA practices (Heitink, et al., 2016; Yan, 2014; Yan et al., 2021).

Teachers' daily classroom assessment has also been underlined in English language teaching context (Tsgari & Banerjee, 2016). In the field of teaching English as a second or additional language, taking the advantage of assessment has been discussed for improving learning for nearly 10 years (Davison, 2019). The role of *Assessment for Learning* has received increased attention across a number of disciplines and an interest has been growing in studies of English Language Teaching (Cindrić & Pavić, 2017; Ghaffar et al., 2020; Lu & Mustapha, 2020; Nasr et al., 2018; Nasr et al., 2019; Nasr et al., 2020; Umar, 2018; Vattøy, 2020; Xu & Harfitt, 2019). For achieving successful AfL implementation, two requirements have been reported for teachers: "(1) to develop their AfL skills and (2) to ensure that students will respond positively to the activation to develop their skills required in learning" (Alonzo, 2016, p. 46). Thus, teacher training can be seen as an essential element in the successful implementation of AfL.

In Turkey, Yücel et al. (2017) carried out a study on the investigation of the English language curricula between 2002 and 2017, reflecting a period of regulations, and these researchers concluded that the curriculum objectives were not congruous with the assessment and evaluation techniques. Similarly, Çelik and Filiz (2018) examined the 2014 secondary school English language curriculum and found that there was not an agreement between theory and practice. Thus, it has been understood that further studies are required for achieving formative assessment implementation despite the crucial changes adopted in the curriculum. The reform in Turkish Education System has also been included in OECD report in 2019. This report focused on the discrepancy seen between theoretical and practical results. According to the report, questions giving rise to memorization was identified as a way of assessment held in the classroom rather than leading to critical thinking or problem solving. Consequently, this report suggested that various assessment tools be elaborated, and thus teachers should be able to utilize several assessment techniques. Besides, teachers should be given chances to improve their practical implementations with teacher education programs (Kitchen et al., 2019).

In the Turkish EFL context, Öz (2014) called attention to the restricted number of AfL studies in the language teaching settings where English occupies the position of a second or a foreign language. To the best of the researcher's knowledge, there has not been sufficient research in the Turkish EFL context in the present day as well.

In the current English Language Teaching Programs in Turkey, there are two courses in the curriculum of universities ELT programs: "Assessment and Evaluation" and "English Language Testing and Evaluation" for the undergraduate students. However, in a recent study, Ölmezer-Öztürk and Aydın (2019) indicated a significant point that undergraduate training could be inadequate for enabling pre-service teachers to be knowledgeable. Additionally, a system, which deals with the professional development of English language teachers, is not

available in Turkey. From the EFL context in Turkey, several researchers (Hatipoğlu, 2015; Mede & Atay, 2017; Ölmezer-Öztürk & Aydın, 2018) pointed out the lack of comprehensive in-service training on assessment and evaluation at the same time they emphasized their low levels of assessment literacy skills concluding with a recommendation for offering in-service training on assessment. The focus of the current study was to evaluate the In-Service Training Program on *Assessment for Learning* prepared for the Turkish EFL teachers working in the high schools.

Purpose and Significance of the Study

As the consequence of the previous studies abovementioned, *Assessment for Learning* has been accepted as one of the essential factors that develop learning. This study set out to explore the influence of an In-Service Training Program on AfL. In this scope, an INSET program was prepared and presented for EFL teachers working in the state high schools. The objectives of the INSET program were to increase the levels of teachers' awareness about *Assessment for Learning* and to enable them to use the latest implementations in their classes. In this study, it was also aimed to examine the opinions of teachers related to this training at the end of the program. The study sought to answer the following specific research questions:

1. What are the differences between the English language teachers' responses for the tasks before and after the in-service training courses?
2. What are the opinions of the English language teachers for the in-service training program?

Method

Research Design

Language Curriculum Design Model of Nation and Macalister (2010) was used to design this study, and the reason for choosing this model for the current study was that it is specifically prepared for the use of EFL or ESL teachers. The model includes 'Needs Analysis' and 'Environment Analysis', 'Principles', 'Goals, Content and Sequencing', 'Format and Presentation', 'Monitoring and Assessment' and 'Evaluation'. 'Environment Analysis' consists of the elements probably influencing the judgments about the course objectives and contents while 'Needs Analysis' is related to learners' necessities, wants, and lacks in the learning process (Nation & Macalister, 2010). Although all these steps were followed in the development of the INSET program, this paper focused on investigating the effects of the In-Service Training Program prepared for English language teachers working in the Turkish high school contexts. The details of the other steps are described in the following parts.

Participants

The participants were the English language teachers working in the state high schools in Afyonkarahisar, a city of the central Anatolia region of Turkey. They took part in the INSET program conducted in the spring term of 2020-2021 education year. The teachers were selected as the participants from this city due to easy access for the researchers. Thus, convenience sampling that refers to the way that researchers could access easily (Creswell,

2012) was chosen as the method of sampling for the current study. The In-Service Training Program was open to all these English language teachers, and they took part on a voluntary basis. The courses were designed as five sessions depending on the five strategies of *Assessment for Learning* (Table 1).

Table 1

Number of the Participants in the Sessions of the Program

Courses	Types of Participation		
	Audiences of the Training Program N	Respondents to the Tasks N	Respondents to the Questionnaire N
1. Providing feedback that moves learners forward	59	10	47
2. Engineering effective classroom discussions and other learning tasks that elicit evidence of student understanding	36	16	26
3. Clarifying and sharing learning intentions and criteria for success	28	11	17
4. Activating students as the owners of their own learning	20	10	20
5. Activating students as instructional resources for one another	16	10	13

Table 1 shows the numbers of participants taking part in all these sessions. Accordingly, the highest audience numbers in the training program and teachers responding the questionnaire are detected in the session of 'Providing feedback that moves learners forward' while the lowest numbers are seen in 'Activating students as instructional resources for one another'. The numbers of teachers who responded to the tasks both at the beginning and end of the sessions are close to each other which change between 10 and 16.

Data Collection Tools and Process

In the present study, an In-Service Training Program was designed and implemented. In relation to the category of 'Principles', learning objectives were determined under five AfL strategies. The main objectives were as follows: (1) to make the meanings of AfL strategies clear, (2) to clarify the importance of AfL strategies in enhancing students' learning, and (3) to improve teachers' capabilities to use AfL strategies with the help of some sample demonstrations. The contents and sequence of these sessions can be seen in Table 2.

Table 2

Contents and Sequencing of the In-Service Training Program

Courses	Contents
1. Providing feedback that moves learners forward	Meaning and importance of feedback Features of effective feedback Ways to improve feedback quality Difference between 'feedback' and 'marking'

Table 2*(Continued)*

Courses	Contents
2. Engineering effective classroom discussions and other learning tasks that elicit evidence of student understanding	Role of 'questions' in effective feedback Types of questions & classroom discussion Waiting time Dialogue & refining process
3. Clarifying and sharing learning intentions and criteria for success	Meaning of learning intention & success criteria Features of strong learning intention & success criteria Strong & weak samples
4. Activating students as the owners of their own learning	Key concepts of self-assessment Ways to improve students' assessing their own performances Online tools for self-assessment
5. Activating students as instructional resources for one another	Key concepts of peer assessment Ways to improve students' assessing their peers' performances Online tools for peer assessment

The In-Service Training Program was carried out in five sessions as determined by the Directorate of National Education. During these sessions, two types of instruments were utilized to gather data: tasks designed for all these courses and a questionnaire.

Tasks

A total of five courses were arranged, and they were based on the five AfL strategies. For all these courses, a task prepared in accordance with the MoNE curriculum was sent to the participants for the first time before the course started, and for the second time at the end of the end of presentation. The first course was about 'Providing feedback that moves learners forward'. A writing task of a student was demonstrated, and the teachers were expected to write feedback for the task. The presentation began after the pre-task was completed. The post-task was conducted for the same task when the presentation finished. In the same way, these pre- and post-task procedures were implemented in all these courses. The second course was about 'Engineering effective classroom discussions and other learning tasks that elicit evidence of student understanding'. For the second course, in line with the item of the curriculum, "Students will be able to write a comment on a topic via social media" (MoNE, 2018b, p.32), the teachers formed three questions that they would pose to their students. The third course was on 'Clarifying and sharing learning intentions and success criteria'. In accordance with the curriculum item "Students will be able to write an argumentative essay including solutions for disadvantaged people's problems" (MoNE, 2018b, p.58), the participants were addressed three questions: 'How can you explain this task?', 'When you are explaining the task, what else can you address?', and 'How can you explain success criteria?' The next course was also about the same item of the curriculum in the scope of 'Activating students as the owner of their own learning', and the teachers were asked to explain the item for this time as a task of self-assessment. For the last session which was about 'Activating students as instructional sources for one another', the teachers were expected to introduce a peer assessment task related to the item 'Students will be able to write a booklet to describe their

hometown' (MoNE, 2018b, p. 53). As mentioned above, the pre- and post-task procedures were carried out during all these sessions for the purpose of comparing two responses of the teachers, and as a result, revealing the effect of the training.

Questionnaire

A questionnaire was prepared to uncover the ideas of the participants about INSET. In this questionnaire, four-point Likert type questions and open-ended items were used. The first Likert type item was *"What do you think about Assessment for Learning?"* in order to disclose their opinions on AfL and the grading points were 'very important', 'quite important', 'important', and 'not important'. *"How do you compare your tasks that you prepared before and after the training?"* was the second question directed to the participants, and the aim, here, was to enable these teachers to make a comparison between their first and second responses to the same task. For the second question, the grading points were 'completely changed', 'changed a lot', 'partially changed', and 'no change'. The third question was about the usefulness of the techniques, and the question was *"What do you think about the practicality of the techniques?"* For the third question, the grading points were 'very practical', 'practical', 'partial practical' and 'not practical'. There were also other open-ended questions about the techniques and the practical implementations: *"Can you share your positive and negative opinions related to implementation of these techniques? Can you give examples related to your students?"* and about a general suggestion for the program: *"Is there any point that you offer to add to or omit from the content of the training?"* Expert opinions were obtained from three academicians including two academicians from the department of English Language Teaching and one from the department of Measurement and Evaluation in Education to validate the questionnaire, and the items were finalized.

Data Analysis

In this study, both qualitative and quantitative analyses were conducted to reveal the effect of the In-Service Training Program. The data obtained as the task responses were evaluated for qualitative analysis. With this purpose, a checklist was formed based on the lists that were prepared by Andersson and Palm (2017) and Andersson et al. (2017). As the first step, the responses provided before and after tasks were read, and activities were described in their responses. The next step was to categorize these activities related to the "big ideas" of Wiliam and Thompson (2008). The rules offered by Andersson and Palm (2017) were also adopted in detecting the differentiations in comparing their first and second answers, and the new responses were distinguished as the answers which were not identified in the first task but found in the second one and which were observed in a modified form in the latter one. The last step was to classify these responses and calculate the scores. For the reliability of the data analysis, one of the researchers analyzed the data alone by following all these steps. After that, the researcher sent the data and the data analysis results to the other researcher, an expert in the department of English Language Teaching and one more expert from the same department to control and confirm the analyses.

In order to detect the participants' opinions on the training program, qualitative and quantitative data were collected. Descriptive statistics were used to analyze quantitative data, and the teachers' responses were shown in pie charts as percentages. As for the qualitative

data analysis, in the light of the previous studies, codes were identified as the 'start list', thus, deductive coding was performed to analyze qualitative data (Miles et al., 2014).

Ethical Issues

The ethics committee permission, numbered 34139 and dated 24.06.2020, was obtained from Anadolu University Scientific Research and Ethical Review Board. Following this, permission of the Directorate of National Education in Afyonkarahisar and the Governorship of the city was also gathered, and the program was arranged in five sessions in three days. The research was conducted with English language teachers working in the high schools based on their voluntary participation in this training program.

Findings

RQ 1. Changes in the task responses of English Language Teachers before and after the In-Service Training Courses

The first research question explored the differences between pre- and post-task responses of the English language teachers. For achieving this aim, teachers' responses to the same task given before and after the training were compared, and the results are presented in Table 3.

Table 3

Results of the Task Responses Before and After the Training

'Assessment for Learning' Activities	Before Training N	After Training N
Clarifying and sharing learning intentions and criteria for success		
1. The teacher identifies general learning intentions.	11	2
2. The teacher makes the learning goals clear by explaining the topic, referring the goals, sharing key words, and examining these words in the context.	0	12
3. The teacher provides samples for getting students to observe strong and weak properties in the texts.	0	6
4. The teacher clarifies the success criteria in general.	10	5
5. The teacher enables students to involve in comprehend these criteria step by step identifying these criteria.	0	2
Engineering effective classroom discussions and other learning tasks that elicit evidence of student understanding		
6. The teacher provides opportunities for students to express their opinions.	4	19
7. The teacher provides opportunities for students to think further on specific points.	1	3
8. The teacher gives chances for students to share their ideas before the whole class discussion.	0	1
Providing feedback that moves learners forward		
9. The teacher points out the strong sides of the student work.	4	6
10. The teacher demonstrates sides to be developed more.	6	7
11. The teacher indicates the weak sides of the student work.	2	1
12. The teacher's comments are unclear and unintelligible.	1	6
13. The teacher prefers suggestions as feedback instead of giving the correct answers directly.	0	9

Table 3*(Continued)*

‘Assessment for Learning’ Activities	Before Training N	After Training N
Activating students as the owner of their own learning		
14. The teacher proposes various self-assessment techniques including ‘portfolio’, ‘self-testing’, and so on.	1	7
15. The teacher selects online tools as a way of self-assessment implementation.	0	1
Activating students as instructional resources for one another		
16. The teacher suggests various peer assessment techniques including ‘homework help board’, ‘end-of-topic questions’, and so on.	0	7

As the results in Table 3 are examined, there are differences found between all the scores of pre- and post-task responses. The biggest differences are seen in the items 1, 2, 6, and 13. The items 1 and 2 were closely related to each other. It is clear in this table that the expressions the teachers used for learning goals included more details such as preferring key words and using words in the contexts after the presentation. Moreover, one of the most striking differences is seen in item 6 which is about the strategy of ‘Engineering effective classroom discussions and other learning tasks that elicit evidence of student understanding’. In this sense, more examples were found to give students chances to share their ideas in the responses of teachers after the training. Another significant difference is observed in item 13 showing a change in the direction of an increase in suggestions as feedback. It is also valuable to pay attention to the changes seen in the last two strategies. Except for a teacher, none of the teachers completing the task provided a self-assessment technique before the training; however, seven teachers offered some ways discussed during these trainings as the post-task responses for both ‘Activating students as the owner of their own learning’ and ‘Activating students as instructional resources for one another’ strategies. In the following paragraphs, each course is evaluated in detail.

When the answers of pre- and post-task were compared related to the session ‘Clarifying and sharing learning intentions and success criteria’, differences were detected in the explanation of learning intentions. In this group, seven teachers provided general learning intentions before the presentation as follows:

‘Essay writing task’ (P3), ‘I can talk about disadvantaged people’ (P4), ‘Let’s write an essay on disadvantaged people’ (P5), ‘Give the topic and explain the main points’ (P6), ‘Let’s write an essay on disadvantaged people. Please write your own ideas on the disadvantaged people’s problem. You should write it in an argumentative essay format’ (P7), ‘Let’s write an essay. It must be argumentative’ (P8), and ‘Write an essay including solutions for disadvantaged people’s problems. You should write an essay at least in three paragraphs. You should search on the net about the subject.’ (P11)

When these participants’ answers were examined after the training, some changes were seen. For instance, three teachers mentioned the goals in their explanations: ‘*You will be able to argue surely on solutions for disadvantaged people’s problems by using step by step approach for first identifying, then giving solid examples and finally finding solutions for those people*’ (P3) and identifying key words and using them in the context:

'Use words and ask what they mean or be reminded of them. I can address related topics' (P4), 'Writing and essay about the disadvantaged people's problems and solutions of these problems topic, key words, and examples' (P5), 'We should identify the goals and share keywords. We should examine the words in the context. We may present good works about the topic' (P7), and 'I can explain by using an outline and giving some sample sentences' (P8).

A differentiation was found related to the explanations of success criteria in their responses before and after the session. Especially, the following statements by P9 and P11 are seemed as the prominent samples of the post-task answers:

"Be careful to write a topic sentence. Support your topic with arguments and examples. Finish your paragraph with a strong sentence including your topic sentences and your opinion." (P9); "Be careful to stick to the topic, to strengthen it with supportive sentences. Be careful about use of language and vocabulary" (P9); "Be careful about the integrity of the meaning, vocabulary, and punctuation. Conclude with a strong sentence" (P11)

As for the course 'Engineering effective classroom discussions and other learning tasks that elicit evidence of student understanding', the questions prepared by the teachers before and following the training were compared. The first point revealed that many of these questions sought definite answers, for instance,

'How often do you look at your facebook or instagram' (P2); 'Do you use social media?', 'Which social media tools do you use?' (P3); 'Do you like using social media?', 'How much time do you spend on social media?' (P5); 'Which age group uses the social media the most?' (P7); 'Is the social media necessary for us?', 'Are you addicted to social media?', 'How many hours are you online in social media?' (P8); 'Do you use social media actively?', 'Which accounts do you have?' (P9); 'Can you use a social media organ?', 'Which social media organ do you use the most?' (P10); 'How many social media accounts have you got?' (P11); 'Have you got social media accounts?', 'What's your favorite social media app?', 'How often do you check your social media accounts?' (P13); and 'How many friends have you got on your social networking sites?' (P14).

Although most of these teachers chose to ask yes/no questions, there were four questions offered to students to give them a chance to articulate their opinions: '*Which age group uses the social media the most?*' (P7); '*Is the social media necessary for us?*', '*Are you addicted to social media?*' (P8); and '*Do you use social media actively?*' (P9). Among the questions suggested at the end of the training, more open-ended questions were detected as follows:

'What do you think about social media?' (P1); 'What do you think about your relationship with social media?' (P2); 'In what purpose do you use your mobile phones? (Elicit all the answers in brainstorming)', 'In which different ways can we use social media?', 'In which ways should you write your comment on social media?' (P3); 'What do you think about violence on tv?', 'What can be done to prevent violence in the society?', 'What can you say as a motto to raise awareness?' (P4); 'Which social media tools do you prefer and what is the reason?' (P5); 'What do you think about social media?' (P6); 'What do you think about this?' (P8); 'What kind of topics in social media are you interested in?' (P12); 'What are the good points of social media?', 'What are the bad points of social media?' (P13); 'What do you think about learning processes?', 'Why is this process important?', 'What are the advantages of this activity?' (P15); and 'What is social media?', 'What do you think about it?' (P16).

Thus, these questions could allow teachers to carry out effective classroom discussion. To keep interaction going among the students, one of these teachers proposed: *'In pairs, think and share the advantages of social media'* (P14).

In the feedback session, the place of providing clear and intelligible feedback was underlined several times. One of the teachers (P3) stated that *'That's a good essay for a student at your age'* before the presentation, but the feedback of the same teacher changed after the training as *'You could write longer sentences but still this is also very good. In order to tell your feelings, you could add another paragraph'* (P3). While another teacher (P4) offered that *'The information mentioned above is satisfactory'* and *'The autobiography is motivating'* before the training, the explanations were easier for students to understand just by following the presentation such as *'You should give more information about the motivation that drives you to be an English teacher'* and *'You can also give a tangible explanation about your first English lesson'*. There were two other participants who did not provide feedback before the course. They used clear statements at the end of the training as follows: *'You should be organized about what to write and the order of the ideas before you start writing'* (P6) and *'This text is quite clear, but you can use shorter sentences for the readers'* (P8). When asked how often they used grades as feedback for a transition between the issues using the feedback session, 59 teachers shared their classroom implementations in terms of frequency of occurrence. The most striking result emerged from the data was that a total of 38 participants expressed utilizing the grades as feedback either always (N=3) or often (N=35). The other outcome of the study was that the number of teachers who did not use grades as feedback was only three. Thus, it could be seen that grades were used in a frequent manner when teachers would like to provide feedback. At the end of this session, six teachers provided suggestions rather than giving the correct response although no sample was found before the course:

'You could write longer sentences' (P3); *'You should give more info. about the motivation that drives you to be an English teacher'* (P4); *'You can add how you can use your knowledge in real life'*, *'You can add how your teacher makes creative classroom activities'* (P5); *'You should be organized about what to write and the order of the ideas before you start writing'* (P6); *'This text is quite clear but you can use shorter sentences for the readers'* (P8); and *'While passing a new subject, you should write your text in paragraphs'* (P9).

The answers given by the teachers to the pre-task in the session *'Activating students as owner of their own learning'* indicated that they were familiar with some ideas about students' assessing their own performances. To illustrate, two teachers offered some ways used in self-assessment techniques: *'Checklist or questionnaires'* (P4) and *'Using checklist'* (P6). The other participant also mentioned a preparation stage of a self-assessment technique as the answer to the question: *'Students will determine success criteria'* (P8). On the other hand, the differentiation was detected in the techniques they suggested just following the training including *'portfolio'* (P1, P3, P4, P10), *'rubric preparation'* (P2, P5, P6), *'self-testing'* (P5), and *'wiser'* (P5), which was explained during the presentation as an online tool for self-testing.

For the last training *'Activating students as instructional resources for one another'*, there was not any relevant suggestion for peer assessment detected among the responses before the session began. The ideas suggested before the presentation was mainly about changing their writing with one pair and completing the work, grading peer's work, or providing

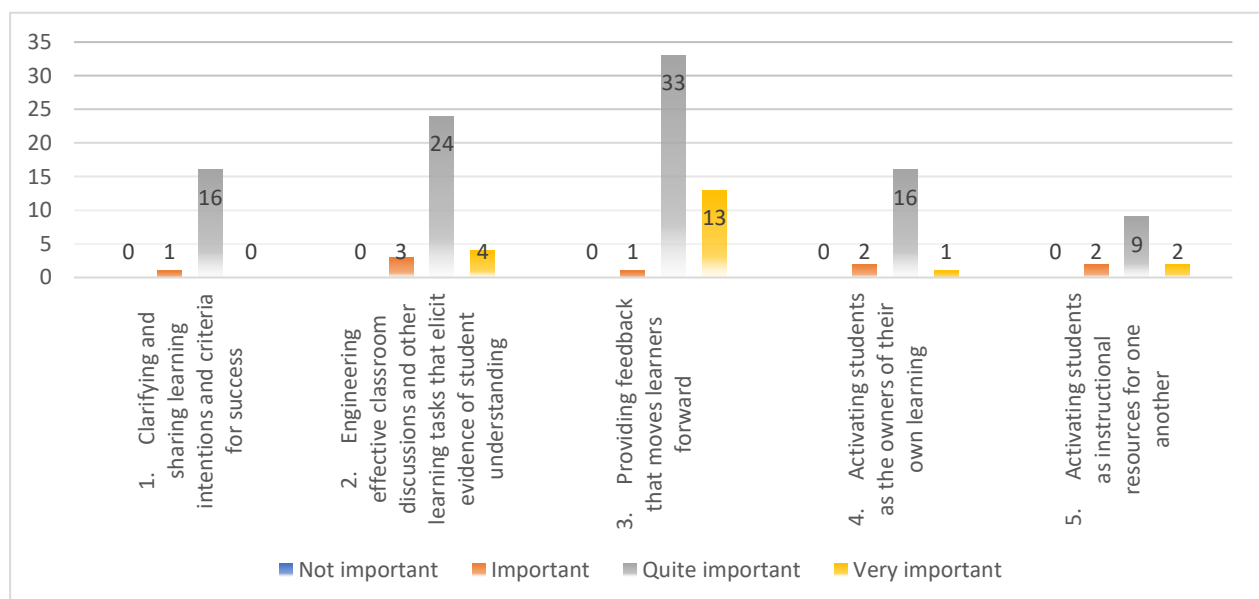
feedback for the other's work. From these answers, an uncertainty was revealed for how students carried out these duties. However, the answers suggested after the training indicated specific techniques about how these teachers would implement peer assessment in their classes including 'end of topic questions' (P1), 'student reporter' (P1, P7), 'homework help board' (P4, P7), and 'error classification' (P7, P10) which were all the issues discussed during the training session.

RQ 2. English Language Teachers' Opinions for the In-Service Training Program

The second research question aimed at revealing the teachers' perceptions related to the In-Service Training Program. The teachers were asked to share their ideas first related to AfL and its impact on their courses with the question 'What do you think about AfL?' Following this general evaluation, they were expected to make a comparison between their performance on the task before and just after the training course, and for this purpose, they were posed the question, 'How do you compare your task that you prepared before and after the training?' It was also among the objectives to disclose their opinions about the practicality of the techniques recommended during the program and the following question was asked 'What do you think about the practicality of these techniques?'

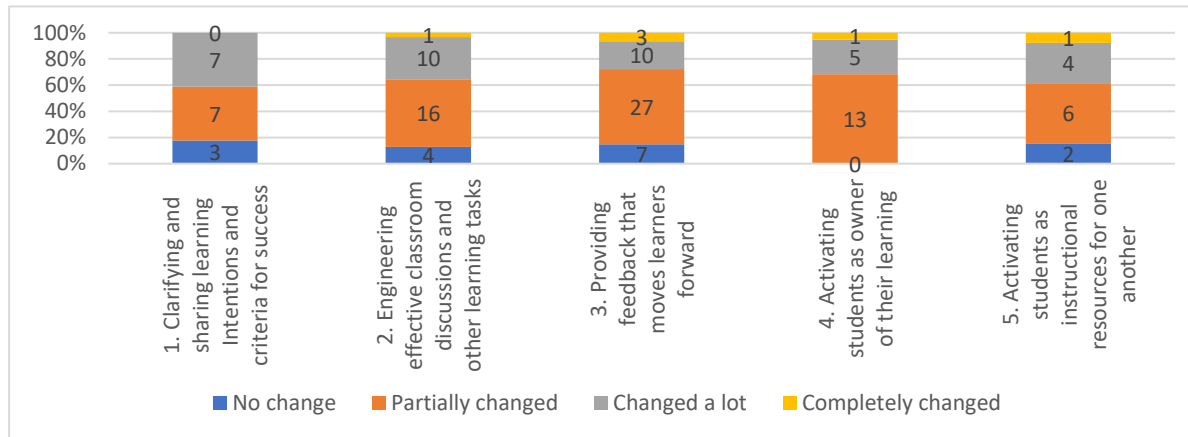
Figure 1

Teachers' Opinions related to Assessment for Learning



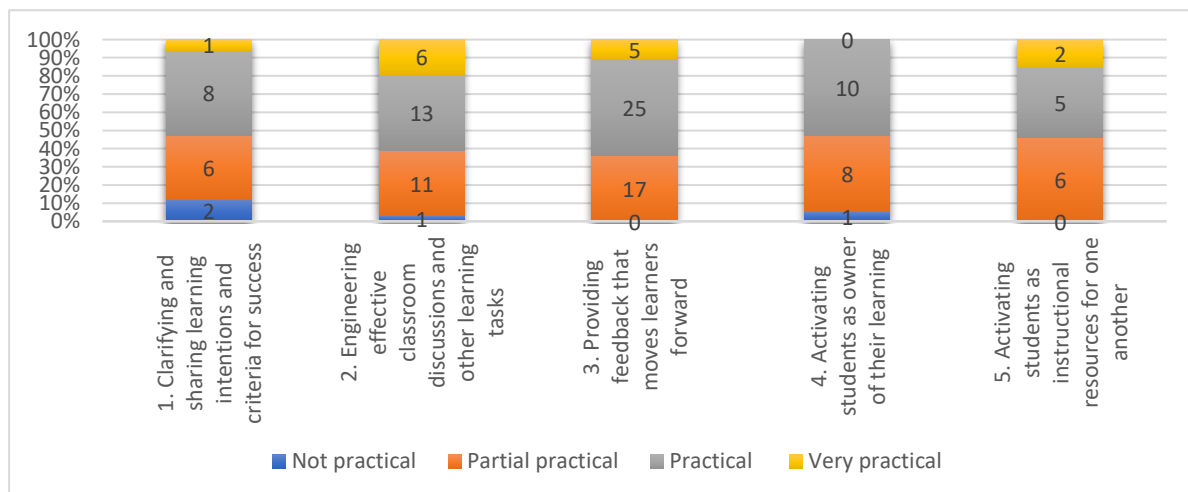
The first questions aimed to reveal these teachers' ideas about AfL at the end of each session. Figure 1 displays the results of the teachers' ideas. A total of 127 responses were gathered, and 98 of these participants thought that AfL was quite important for their teaching. Following this, 20 teachers considered AfL to be 'very important' and 9 teachers 'important'. At the end of these courses, none of the participants stated AfL was 'not important'. Overall, these results indicated that the English language teachers were aware of the value of AfL for their English language courses.

Figure 2
Teachers' Opinions related to the Pre- and Post-Tasks



The next question asked how they compare their tasks prepared before and after the training. As shown in Figure 2, most of the teachers (N=69) reported 'partially changed', a huge group (N=36) indicated 'changed a lot', and another group (N=6) rated 'completely changed'. While many teachers taking a part in these courses detected a change in their pre- and post-tasks, a group (N=16) found 'no change' between their first and second drafts. The group reporting 'no change' was divided into the sessions as follows: 7 teachers in 'Providing feedback that moves learners forward', 4 in 'Engineering effective classroom discussions and other learning tasks', 3 in 'Clarifying and sharing learning intentions and success criteria', and lastly 2 in 'Activating students as instructional resources for one another'. When these teachers were asked to make a comparison between their tasks before and after the courses, a great number of the participants (87.4%) declared a change between two forms in varying degrees. On the other hand, only a minority of the respondents (12.6%) found 'no change' in these two forms.

Figure 3
Teachers' Opinions related to Practicality of the AfL Techniques



The other question was what they thought about the practicality of these techniques. As for this question, 97% of the responses indicated the practicality of the techniques discussed during these sessions in varying degrees. In this group, 61 of them demonstrated these techniques as 'practical', 48 of them as 'partial practical', and 14 of them as 'very practical'. Thus, it can be concluded that nearly all the participants stated an opinion about usefulness of the techniques for their classes.

The participants were also asked to offer further opinions on the applicability of the techniques mentioned during these courses. The following table summarizes the reasons reported by some of the participants demonstrating the inappropriacy of these techniques for their classes.

Table 4

Teachers' Opinions on the Reasons for not Using AfL Techniques

	Clarifying and sharing learning intentions and success criteria	Engineering effective classroom discussions and other learning tasks	Providing feedback that moves learners forward	Activating students as owner of their learning	Activating students as instructional resources for one another	TOTAL
Lack of students' proficiency levels	1	2	5	1		9
Time constraints		2	3			5
Crowded classes			4			4
Problems of students' attendance and motivation			2			2
Lack of interest		2				2
Lack of eagerness for classroom participation		1				1
Characteristics of being shy		1				1
Attitudes of students					1	1

Table 4 presents the results obtained from qualitative analysis of teachers' opinions regarding why they could not use these techniques in their courses. As seen in Table 4, except for the last session 'Activating students as instructional resources for one another', a small number of teachers (N=9) indicated 'lack of students' proficiency levels' as the reason that these techniques would not be relevant for their courses at the end of all the other sessions. Among this group, five teachers showed students' proficiency levels as a cause after the 'Providing feedback that moves learners forward' session. The only teacher who stated this as a reason at the end of the 'Clarifying and sharing learning intentions and success criteria' session claimed that these techniques were utopic for their students due to their level of proficiency. There were also two teachers in 'Engineering effective classroom discussions and other learning tasks' session and one in 'Activating students as owner of their learning'. Following this, 'time constraints' reported by five participants in two of these sessions. At the end of the 'Engineering effective classroom discussions and other learning tasks' session, one teacher expressed the potential usefulness of these techniques on the condition that there was sufficient time. Following the 'Providing feedback that moves learners forward' session, two of these teachers indicated the problems occurring related to the students' attendance and

motivation. Similarly, two teachers reported their lack of interest as a reason for not using these techniques, and one teacher also reported their lack of eagerness for classroom participation at the end of the 'Engineering effective classroom discussions and other learning tasks' session. In the same session, one teacher underlined students' characteristics of being shy as a cause of this situation. One of the participants also ascribed this to their attitudes in the 'Activating students as instructional resources for one another' session. It is apparent from this table that very few teachers gave justification for their inability to carry out these techniques.

At the end of the feedback session, a couple of recommendations were made which could be included in or excluded from the scope of the course. The first one was that a video was demonstrated at the first part of the course which was further suggested by two participants. The second suggestion was about involving students while practicing feedback. To that end, one teacher proposed to add suggestions to increase their involvement in feedback process. The last suggestion was that there should be more examples related to classroom implementation which was also offered by another teacher as containing more practical techniques in 'Activating students as owner of their learning'.

Conclusion, Discussion and Implications

In the recent years, *Assessment for Learning* has drawn considerable attention around the world. Effectiveness of the INSET Program which was developed on AfL and presented to the Turkish EFL teachers was evaluated in this paper.

The results of the questionnaire revealed that most of the teachers stated the changes between their tasks that they suggested at the beginning and just following training sessions, and these differences were detected at varying levels. This finding is consistent with that of DeLuca et al. (2015) who illustrated the changes in the opinions of the teachers as follows:

"I now think of assessment as so much more than mark, it is a guideline for improvement" (p. 129).

"AfL allows for teachers to identify for each particular student what their areas of strength and weaknesses are, allowing them to improve their performance without it factoring into their grade" (pp. 129-130).

As for the strategy 'Clarifying and sharing learning intentions and criteria for success', some changes were detected when the participants pre- and post-tasks responses were compared. To illustrate these changes, the teachers were asked to be more specific in their responses after the course such as identifying key words, explaining word meaning depending on the context, and enabling students to understand success criteria gradually. A possible explanation of these results may be the increase in awareness of teachers' implementation of AfL strategies. These results differ from some published studies (Crichton & McDaid, 2016; Nasr et al., 2018). In the study of Nasr et al. (2018), teachers were seen not to be so enthusiastic about explaining learning goals.

For the strategy of 'Engineering effective classroom discussions and other learning tasks', as compared with the task responses before the training, students were given more chances to express their opinions, to think further on some issues, and to share their opinions with a pair before explaining them to the rest of the class according to the responses provided after the course. The questions teachers posed before the training did not enable students to

consider on an issue and form their opinion such as “*How often do you look at your Facebook or Instagram or...?*”, “*Do you use social media?*”, “*Which social media tools do you use?*”, and “*How many friends have you got on your social networking sites?*” However, after the session ended, they could ask several open-closed questions including:

“What do you think about social media and my relation?”, “What about your parents?”, “In what purpose do you use your mobile phones? (elicit all the answers in brainstorming)”, “In which ways should you write your comment on social media?”, “What can be done to prevent violence in the society?”, and “What can you say as a motto to raise awareness?”

It is possible to see the differentiations in the comparison of their responses to the pre- and post-task. As the same with the present study, Yakışık (2021) carried out a study in the EFL high school setting and identified fear of making mistakes as a reason causing students to become anxious. Effective classroom discussion can be maintained by way of giving more chances to students for explaining their opinions and thinking more on specific points, thus teachers’ questions play a crucial role in maintaining an effective classroom discussion.

In the feedback session, there were also some modifications between the responses to the pre-task and post-task. An important finding was that teachers offered fewer vague statements after the training ended. It is encouraging to compare this result with the study conducted by Ratnam-Lim and Tan (2015) who studied the problems related to implementation of an assessment system and concluded that it was necessary to give feedback to be clear. In relation to this session, more suggestions were proposed by the teachers just following the training. The reason of this can be traditional exam dominated culture of the Turkish education system. This result corroborates the outcomes of Boardman and Woodruff (2004) who conducted a study in Texas in the USA, Carless (2005) in Hong Kong, and Brown et al. (2009) in China. In the study of Boardman and Woodruff (2004), high-stake exams’ influence was detected on the practices of new techniques. Carless (2005) also reported difficulties in AfL implementations, and the exam-oriented system was identified as a reason for this. Therefore, the present results are representatives of an emerging trend in AfL implementations.

In relation to the strategy of ‘Activating students as the owners of their own learning’, teachers who took part in this session offered several self-assessment techniques at the end and these were different from the answers suggested before the session started. These results are in agreement with those obtained in Turkish EFL high school contexts by Akdağ and Özkan (2017), Kayacan and Razi (2017), and Yakışık (2021). The outcomes of these study shed light on the following points: positive impact of blogs in developing learners’ writing skills, being autonomous in writing, and enhancing their motivation for writing revealed in the study of Akdağ and Özkan (2017); influence of self-review on learners’ writing skills as a consequence of Kayacan and Razi’s study (2017); and learners’ high levels of self-correction concluded in the study of Yakışık (2021).

For the training of the strategy of ‘Activating students as instructional resources for one another’, as distinct from the previous responses that they offered at the beginning, several new peer assessment techniques were suggested after they were involved in this session. A possible explanation of these results might be the influence of training. These results are in line with the study of Zlabkova et al. (2021), who arrived at a decision about prominent impacts of involving teachers in the peer assessment activities on their formative assessment

comprehension. In the same vein, Kayacan and Razi (2017) concluded that a positive influence of peer feedback was detected on developing writing skills of the students.

For evaluating the INSET program, 127 teachers commented on these sessions and nearly all the participants considered the techniques explained in the scope of this program as practical at different levels, other than only three teachers. It seems possible that these results are due to “teacher commitment”, which was also deduced by Nasr et al. (2019). In the study of Nasr et al. (2019), following the school curriculum rigidly and not having the responsibility to implement AfL were detected as two reasons that teachers working in the high schools did not put these techniques into practice.

The teachers, participants of these sessions, indicated various reasons for not benefiting from these techniques in their classes. The finding of ‘crowded classes’ is supported with the outcomes revealed in the studies of Büyükkarcı (2014) and Nasr et al. (2019). However, this finding contradicts with the results of Nasr et al. (2020), who detected no meaningful distinction for monitoring and scaffolding in respect to class size. These results therefore need to be interpreted with caution, and one possible explanation could be deduced from Xu and Harfitt (2019)’s results that the ways and the degrees teachers exerted themselves to practice these techniques were two factors indicating the success of the implementations in crowded classes (Xu & Harfitt, 2019). ‘Time constraint’ was another finding revealed at the end of the program which is consistent with the data obtained in the studies of Nasr et al. (2019) and Xu and Harfitt (2019). This rather disappointing result might be the consequence of what Deneen et al. (2019) articulated that AfL execution became hard for the teachers on account of large-scale exams commonly used in the high schools. Thus, time management is a crucial point in teaching which should be planned carefully to make AfL as a part of their teaching. ‘Students’ proficiency levels’ was also uncovered, and this result is consistent with the results revealed in the study of Lu and Mustapha (2020) suggesting that students with higher proficiency levels indicated higher levels of desire for accepting and practicing AfL. However, several AfL techniques, some of which were presented in the scope of the training in the current study, can be implemented for groups of students with different proficiency levels. Students’ proficiency levels should not be considered as an impediment depending on the reason that it is possible to benefit from a variety of AfL techniques even for students in kindergarten (William, 2011b). The other factors detected in the current study were ‘problems of students attendance’, ‘motivation’, ‘lack of eagerness for classroom participation’, ‘characteristics of being shy’, and ‘students’ attitudes towards each other’ that are in harmony with the outcomes of the study of Xu and Harfitt (2019).

The present results are significant in at least two major respects. First, the teachers’ responses revealed that teachers had consciousness of the significance of AfL in English language teaching. Second, in this paper, it was also aimed to investigate the opinions of teachers related to the training, and positive results were shed light on at the end of this INSET program. It is required for teachers to become aware of AfL and recent innovations for achieving the improvement of AfL knowledge and practices. These findings may help us to understand the differences in teachers’ responses and opinions about AfL implementations before and after these training sessions. The responsibilities of teachers are to look for the ways to improve their knowledge and skills and to implement new techniques as a part of their courses. The aim of the current study was to increase EFL high school teachers’ AfL knowledge and practical implementations. However, there has been a scarcity on AfL studies in Turkish EFL settings, thus further studies can be conducted with teachers working at various

EFL contexts in Turkey. To develop a full picture of AfL implementation, new studies will be needed that include other stakeholders such as students, parents, and school administrators. In future investigations, it can be suggested to design other programs on AfL with a special focus on teaching practices of pre-service, and it can be a longitudinal study to observe and guide these participants during their first experiences in the real classroom settings.

Contribution Rate of the Researchers

Both researchers equally contributed to the study.

Support and Acknowledgment

The researchers are grateful to Anadolu University and Afyonkarahisar Provincial Directorate of National Education for the support and permission for this research.

Statement of Conflict of Interest

There is no conflict of interest between researchers.

References

- Akdağ, E. and Özkan, Y. (2017). Enhancing writing skills of EFL learners through blogging. *The Reading Matrix: An International Online Journal*, 17(2), 79-95. Retrieved March 16, 2021, from <https://web.archive.org/web/20180421125808id/http://www.readingmatrix.com/files/17-65f6f3wr.pdf>.
- Alonzo, D. A. (2016). *Development and application of a teacher assessment for learning (AfL) literacy tool*. (Doctoral Dissertation). Retrieved May 4, 2020, from <https://unsworks.unsw.edu.au/entities/publication/2c7b124f-042c-4c54-a792-aa5c7b25f6d3/full>.
- Andersson, C. and Palm, T. (2017) Characteristics of improved formative assessment practice. *Education Inquiry*, 8(2), 104-122. <https://doi.org/10.1080/20004508.2016.1275185>.
- Andersson, C., Boström, E., and Palm, T. (2017). Formative assessment in Swedish mathematics classroom practice. *Nordic Studies in Mathematics Education*, 22(1), 5-20. Retrieved March 2, 2019, from https://ncm.gu.se/wp-content/uploads/2020/06/22_1_005020_andersson.pdf.
- Andrade, H. L. (2010). Students as the definitive source of formative assessment. In H. L. Andrade & G. J. Cizek (Eds.) *Handbook of formative assessment* (pp. 90-105). Routledge.
- Assessment Reform Group (2002). Assessment for learning: 10 principles. Retrieved May 21, 2020, from http://www.hkeaa.edu.hk/DocLibrary/SBA/HKDSE/Eng_DVD/doc/Afl_principles.pdf.
- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 18(1), pp. 5-25. <https://doi.org/10.1080/0969594X.2010.513678>.
- Birenbaum, M., DeLuca, C., Earl, L., Heritage, M., Klenowski, V., Looney, A., Smith, K., Timperley, H., Volante, L., and Wyatt-Smith, C. (2015). International trends in the implementation of assessment for learning: Implications for policy and practice. *Policy Futures in Education*, 13(1), 117-140. <https://doi.org/10.1177/1478210314566733>.
- Black, P. (1999). Assessment, learning theories and testing systems. In P. Murphy (Ed.), *Learners, learning and assessment* (pp. 118-134). Paul Chapman Publishing.
- Black P. (2015). Formative assessment – an optimistic but incomplete vision. *Assessment in Education: Principles, Policy & Practice*, 22(1), 161-177. <https://doi.org/10.1080/0969594X.2014.999643>.
- Black, P. and Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74. <https://doi.org/10.1080/0969594980050102>.
- Black, P. and Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139–148. Retrieved July 2, 2020, from <https://kappanonline.org/inside-the-black-box-raising-standards-through-classroom-assessment/>.
- Black, P., Harrison, C., Lee, C., Marshall, B., and Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Open University Press.
- Black, P. and Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5-31. <https://doi.org/10.1007/s11092-008-9068-5>.
- Boardman, A. G. and Woodruff, A. L. (2004). Teacher change and “high-stakes” assessment: what happens to professional development? *Teaching and Teacher Education*, 20(6), 545-557. <https://doi.org/10.1016/j.tate.2004.06.001>.
- Brown, G. T. L., Kennedy, K. J., Fok, P. K., Chan, J. K. S., and Yu, W. M. (2009). Assessment for student improvement: Understanding Hong Kong teachers' conceptions and practices of assessment. *Assessment in Education: Principles, Policy & Practice*, 16(3), 347–363. <https://doi.org/10.1080/09695940903319737>.
- Büyükkarcı, K. (2014). Assessment beliefs and practices of language teachers in primary education. *International Journal of Instruction*, 7(1), 107-120. Retrieved August 7, 2020, from https://www.e-iji.net/dosyalar/iji_2014_1_8.pdf.
- Carless, D. (2005) Prospects for the implementation of Assessment for Learning. *Assessment in Education: Principles, Policy & Practice*, 12(1), 39-54. <https://doi.org/10.1080/0969594042000333904>.
- Cindrić, I. and Pavić, S. (2017). Tendencies and characteristics of feedback given by primary English language teachers in Croatia. *ELOPE: English Language Overseas Perspectives and Enquiries*, 14(2), 85-102. <https://doi.org/10.4312/elope.14.2.85-102>.
- Coombe, C. (2018). *An A to Z of Second Language Assessment: How Language Teachers Understand Assessment Concepts*. British Council. Retrieved January 7, 2021, from https://www.britishcouncil.org/sites/default/files/a_to_z_glossary_final.pdf
- Creswell, J. W. (2012). *Educational Research: Planning, conducting, and evaluating quantitative and qualitative research*. Pearson Education.

- Crichton, H. and McDaid, A. (2016). Learning intentions and success criteria: Learners' and teachers' views. *The Curriculum Journal*, 27(2), 190–203. <https://doi.org/10.1080/09585176.2015.1103278>.
- Çelik, K. ve Filiz, S. B. (2018). Ortaöğretim İngilizce Dersi Öğretim Programı'nın (2014) Eisner Modeline göre değerlendirilmesi. *Eğitim ve Toplum Araştırmaları Dergisi*, 5(1), 50-67. Retrieved October 17, 2020, from <https://dergipark.org.tr/en/download/article-file/496590>.
- Davison, C. (2019). Using Assessment to Enhance Learning in English Language Education. In X. Gao (Ed.) *Second handbook of English language teaching*, (pp. 433-454). Springer.
- De Lisle, J. (2015). The promise and reality of formative assessment practice in a continuous assessment scheme: The case of Trinidad and Tobago. *Assessment in Education: Principles, Policy & Practice*, 22(1), 79-103. <https://doi.org/10.1080/0969594X.2014.944086>.
- DeLuca, C., Klinger, D., Pyper, J., and Woods, J. (2015). Instructional rounds as a professional learning model for systemic implementation of assessment for learning. *Assessment in Education: Principles, Policy & Practice*, 22(1), 122-139. <https://doi.org/10.1080/0969594X.2014.967168>.
- DeLuca, C., Chapman-Chin, A., and Klinger, D. A. (2019). Toward a teacher professional learning continuum in assessment for learning. *Educational Assessment*, 24(4), 267-285. <https://doi.org/10.1080/10627197.2019.1670056>.
- Deneen, C., Fulmer, G. W., Brown, G. T., Tan, K., Leong, W. S., and Tay, H. Y. (2019). Value, practice and proficiency: Teachers' complex relationship with assessment for learning. *Teacher and Teacher Education*, 80, 39-47. <https://doi.org/10.1016/j.tate.2018.12.022>.
- Flórez Petour, M. T. (2015). Systems, ideologies and history: A three-dimensional absence in the study of assessment reform processes. *Assessment in Education: Principles, Policy & Practice*, 22(1), 3-26. <https://doi.org/10.1080/0969594X.2014.943153>.
- Fulmer, G. W., Lee, I. C., and Tan, K. H. (2015). Multi-level model of contextual factors and teachers' assessment practices: An integrative review of research. *Assessment in Education: Principles, Policy & Practice*, 22(4), 475-494. <https://doi.org/10.1080/0969594X.2015.1017445>.
- Ghaffar, M. A., Khairallah, M., and Salloum, S. (2020). Co-constructed rubrics and assessment for learning: The impact on middle school students' attitudes and writing skills. *Assessing Writing*, 45, 1-15. <https://doi.org/10.1016/j.asw.2020.100468>.
- Harrison, C. (2010). Peer and self-assessment. In P. Peterson, E. Baker & B. McGraw (Eds.), *International encyclopedia of education*, (3rd ed. pp 231-235). Elsevier.
- Hatipoğlu, Ç. (2015). Diversity in language testing and assessment literacy of language teachers in Turkey. Paper presented at the 3rd ULEAD Congress, International Congress on Applied Linguistics: Current Issues in Applied Linguistics, Çanakkale, Turkey.
- Hattie, J. and Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112. <https://doi.org/10.3102/003465430298487>.
- Hayward, L. (2015). Assessment is learning: The preposition vanishes. *Assessment in Education: Principles, Policy & Practice*, 22(1), 27-43. <https://doi.org/10.1080/0969594X.2014.984656>.
- Heitink, M. C., Van der Kleij, F. M., Veldkamp, B. P., Schildkamp, K., and Kippers, W. B. (2016). A systematic review of prerequisites for implementing assessment for learning in classroom practice. *Educational Research Review*, 17, 50-62. <https://doi.org/10.1016/j.edurev.2015.12.002>.
- Hopfenbeck, T. N., Flórez Petour, M. T., and Tolo, A. (2015). Balancing tensions in educational policy reforms: Large-scale implementation of assessment for learning in Norway. *Assessment in Education: Principles, Policy & Practice*, 22(1), 44-60. <https://doi.org/10.1080/0969594X.2014.996524>.
- James, M., Black, P., Carmichael, P., Conner, C., Dudley, P., Fox, A., Forst, D., Honour, L., MacBeath, J., McCormick, R., Marshall, B., Pedder, D., Procter, R., Swaffield, S., and Wiliam, D. (2006). *Learning how to learn: Tools for schools*. Routledge.
- Jonsson, A., Lundahl, C., and Holmgren, A. (2015). Evaluating a large-scale implementation of Assessment for Learning in Sweden. *Assessment in Education: Principles, Policy & Practice*, 22(1), 104-121. <https://doi.org/10.1080/0969594X.2014.970612>.
- Karaman, P. and Şahin, Ç. (2017). Adaptation of teachers' conceptions and practices of formative assessment scale into Turkish culture and a structural equation modeling. *International Electronic Journal of Elementary Education*, 10(2), 185-194. Retrieved March 16, 2020, from <https://www.iejee.com/index.php/IEJEE/article/view/320/310>.
- Kayacan, A. and Razi, S. (2017). Digital self-review and anonymous peer feedback in Turkish high school EFL writing. *Journal of Language and Linguistic Studies*, 13(2), 561-577. Retrieved February 12, 2020, from <https://dergipark.org.tr/tr/download/article-file/440790>.
- Kitchen, H., Bethell, G., Fordham, E., Henderson, K., and Li, R. R. (2019). *OECD Reviews of Evaluation*

- And Assessment in Education: Student Assessment in Turkey*. OECD Publishing.
- Klenowski, V. (2009). Assessment for learning revisited: an Asia-Pacific perspective. *Assessment in Education: Principles, Policy & Practice*, 16(3), 263-268. <https://doi.org/10.1080/09695940903319646>.
- Lu, W. and Mustapha, S. M. (2020). Relationship between assessment for learning and English learning motivation of university students in blended learning environment. *Journal of Education and Practice*, 11(5), 65-71. Retrieved April 10, 2021, from <https://core.ac.uk/download/pdf/304991495.pdf>.
- Mede, E. and Atay, D. (2017). English language teachers' assessment literacy: The Turkish context. *Ankara Üniversitesi TÖMER Dil Dergisi*, 168(1), 43-60. Retrieved June 17, 2021, from <https://dergipark.org.tr/en/download/article-file/780021>.
- Miles, M. B., Huberman, A. M., and Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Sage Publications.
- Ministry of National Education [MoNE] (2018a). *Güçlü yarınlar için 2023 Eğitim Vizyonu [2023 Education vision for strong tomorrows]*. Retrieved May 4, 2022, from https://www.gmka.gov.tr/dokumanlar/yayinlar/2023_Egitim%20Vizyonu.pdf.
- Ministry of National Education [MoNE] (2018b). *Ortaöğretim İngilizce Dersi Öğretim Programı [Secondary Education English Language Teaching Curriculum]*. Retrieved March 2, 2021, from <http://mufredat.meb.gov.tr/ProgramDetay.aspx?PID=342>.
- Nasr, M., Bagheri, M. S., Sadighi, F., and Rassaei, E. (2018). Iranian EFL teachers' perceptions of assessment for learning regarding monitoring and scaffolding practices as a function of their demographics. *Cogent Education*, 5(1), 1-29. <https://doi.org/10.1080/2331186X.2018.1558916>.
- Nasr, M., Bagheri, M. S., Sadighi, F., and Rassaei, E. (2019). Iranian EFL teachers' assessment for learning practices and barriers: Do textbooks taught and teaching context matter? *Cogent Arts & Humanities*, 6(1), 1-28. <https://doi.org/10.1080/23311983.2019.1646691>.
- Nasr, M., Bagheri, M. S., and Sadighi, F. (2020). Iranian English language teachers' perceptions of monitoring and scaffolding practices of Assessment for Learning: A focus on gender and class size. *International Journal of Foreign Language Teaching and Research*, 8(29), 75-90. Retrieved May 2, 2021, from <https://www.researchgate.net/publication/344931240>.
- Nation, I. S. P. and Macalister, J. (2010). *Language curriculum design*. Routledge/ Taylor & Francis.
- Ölmezer-Öztürk, E. and Aydın, B. (2018). Toward measuring language teachers' assessment knowledge: Development and validation of Language Assessment Knowledge Scale (LAKS). *Language Testing in Asia*, 8(1), 1-15. <https://doi.org/10.1186/s40468-018-0075-2>.
- Ölmezer-Öztürk, E. and Aydın, B. (2019). Investigating language assessment knowledge of EFL teachers. *Hacettepe University Journal of Education*, 34(3), 602-620. <https://doi.org/10.16986/HUJE.2018043465>.
- Öz, H. (2014). Turkish teachers' practices of assessment for learning in the English as a foreign language classroom. *Journal of Language Teaching & Research*, 5(4), 775-785. <https://doi.org/10.4304/jltr.5.4.775-785>.
- Popham, W. J. (2011). *Transformative assessment in action: An inside look at applying the process*. ASCD.
- PDST, Leaving Certificate Applied, Teacher Handbook. (2022). Retrieved June 19, 2023, from <https://www.pdst.ie/sites/default/files/LCA%20Teacher%20Handbook%202021%20d%20%281%29.pdf>.
- Purpura, J. E. (2016). Second and foreign language assessment. *The Modern Language Journal*, 100, 190-208. <https://doi.org/10.1111/modl.12308>.
- Ratnam-Lim, C. T. L. and Tan, K. H. K. (2015). Large-scale implementation of formative assessment practices in an examination-oriented culture. *Assessment in Education: Principles, Policy & Practice*, 22(1), 61-78. <https://doi.org/10.1080/0969594X.2014.1001319>.
- Stobart, G. and Gipps, C. (2010). Alternative assessment. In P. Peterson, E. Baker & B. McGraw (Eds.), *International encyclopedia of education* (3rd eds. pp. 202-208). Elsevier.
- Swaffield, S. (2011). Getting to the heart of authentic assessment for learning. *Assessment in Education: Principles, Policy & Practice*, 18(4), 433-449. <https://doi.org/10.1080/0969594X.2011.582838>.
- Topping, K. and Ehly, S. (1998). Introduction to peer-assisted learning. In K. Topping, & S. Ehly (Eds), *Peer-assisted learning* (pp. 1-23). Routledge.
- Topping, K. J. (2009). Peer assessment. *Theory into Practice*, 48(1), 20-27. <https://doi.org/10.1080/00405840802577569>.
- Tsagari, D. and Banerjee, J. (2016). *The handbook of second language assessment*. In D. Tsagari & J. Banerjee (Eds.). *Handbook of second language assessment* (pp. 1-10). De Gruyter Mouton.

- Umar, A. M. A. T. (2018). The impact of Assessment for Learning on students' achievement in English for Specific Purposes: A case study of pre-medical students at Khartoum University: Sudan. *English Language Teaching*, 11(2), 15-25. <http://doi.org/10.5539/elt.v11n2p15>.
- Vattøy, K. D. (2020). Teachers' beliefs about feedback practice as related to student self-regulation, self-efficacy, and language skills in teaching English as a foreign language. *Studies in Educational Evaluation*, 64, 1-10. <https://doi.org/10.1016/j.stueduc.2019.100828>.
- Wiliam, D. (2005). Keeping learning on track: Formative assessment and the regulation of learning. In C. J. Anderson & T. Spencer (Eds.) *Making mathematics vital: Proceedings of the twentieth biennial conference of the Australian Association of Mathematics Teachers* (pp. 20-34). Australian Association of Mathematics Teachers.
- Wiliam, D. (2010). An integrative summary of the research literature and implications for a new theory of formative assessment. In H. L. Andrade, R.E. Bennett & G. J. Cizek (Eds.), *Handbook of formative assessment in the disciplines* (pp.18-40). Routledge.
- Wiliam, D. (2011a). What is assessment for learning? *Studies in Educational Evaluation*, 37(1), 3-14. <https://doi.org/10.1016/j.stueduc.2011.03.001>.
- Wiliam, D. (2011b). *Embedded formative assessment*. Solution Tree Press.
- Wiliam, D. (2018). *The handbook for embedded formative assessment*. Solution Tree Press.
- Wiliam, D. and Thompson, M. (2008). Integrating assessment with learning: What will it take to make it work? In C. A. Dwyer (Ed). *The future of assessment* (pp. 53-82). Routledge.
- Willis, J. (2011). *Towards learner autonomy: An assessment for learning approach* (Doctoral dissertation) Retrieved May 17, 2021, from <https://eprints.qut.edu.au/45498/>.
- Wu, X. M., Zhang, L. J., and Dixon, H. R. (2021). Implementing Assessment for Learning (AfL) in Chinese university EFL classes: Teachers' values and practices. *System*, 101, 1–12. <https://doi.org/10.1016/j.system.2021.102589>.
- Wylie, E. C. and Lyon, C. J. (2015). The fidelity of formative assessment implementation: Issues of breadth and quality. *Assessment in Education: Principles, Policy & Practice*, 22(1), 140-160. <https://doi.org/10.1080/0969594X.2014.990416>.
- Xu, Y. and Harfitt, G. (2019). Is assessment for learning feasible in large classes? Challenges and coping strategies from three case studies. *Asia-Pacific Journal of Teacher Education*, 47(5), 472-486. <https://doi.org/10.1080/1359866X.2018.1555790>.
- Yakışık, B. Y. (2021). EFL learners' preferences and emotions about oral corrective feedback at secondary education in Turkey: Are there gender and grade-level differences? *Indonesian Journal of English Language Teaching and Applied Linguistics*, 6(1), 103-122. Retrieved January 9, 2022, from <https://ijeltal.org/index.php/ijeltal/article/view/889/pdf>.
- Yan, Z. (2014). Predicting teachers' intentions to implement school-based assessment using the theory of planned behaviour. *Educational Research and Evaluation*, 20(2), 83-97. <https://doi.org/10.1080/13803611.2013.877394>.
- Yan, Z., Li, Z., Panadero, E., Yang, M., Yang, L., and Lao, H. (2021). A systematic review on factors influencing teachers' intentions and implementations regarding formative assessment. *Assessment in Education: Principles, Policy & Practice*, 28(3) 228-260. <https://doi.org/10.1080/0969594X.2021.1884042>.
- Yücel, E., Dimici, K., Yıldız, B. ve Bümen, N. (2017). Son 15 yılda yayımlanan ilk ve ortaöğretim İngilizce dersi öğretim programları üzerine bir analiz. *Ege Eğitim Dergisi*, 18(2), 702-737. Retrieved January 19, 2021, from <https://dergipark.org.tr/tr/download/article-file/369885>.
- Zlabkova, I., Petr, J., Stuchlikova, I., Rokos, L., and Hospesova, A. (2021). Development of teachers' perspectives on formative peer assessment. *International Journal of Science Education*, 43(3), 428-448. <https://doi.org/10.1080/09500693.2020.1713418>.