

Secondary School EFL Teachers' Formative Assessment Practices and Their Impact on Learning

Amira Ben Hmida

HVF7+M8V, Korba, Tunisie

Author's postal address : 2350, boulevard Édouard-Montpetit, (14239); Montréal (Québec),
H3T 1J4

E-mail: amira.benhmida@umontreal.ca

Received: May 23, 2022 Accepted: July 8, 2022 Published: July 19, 2022

doi:10.5296/ijssr.v10i2.19889 URL: <https://doi.org/10.5296/ijssr.v10i2.19889>

Abstract

The positive impact of formative assessment (FA) on learning was taken as a conventional wisdom in education for decades, yet the empirical evidence supporting its prospective benefits on learning especially from students' perspectives remains distinctively lacunary. To fill this gap, this research project aimed at examining the FA practices employed by Tunisian EFL teachers and their impact on students' learning by putting students' perspectives at the center of the debate. Semi-structured interviews ($n = 5$) were addressed to 5 secondary school teachers to examine the FA practices they implemented in their classrooms. Students ($n = 100$) were administered an internet-based survey to probe their insights as to the impact of these practices on their learning. To cross-check students' answers to the internet-based survey, comparisons of their test scores between two summative tests (STs) occurring before and after the implementation of the specific FA practices were conducted. Major results showed that EFL teachers referred to providing their students with oral and written feedback, sharing with them the used assessment criteria, and enhancing peer and self-assessment. While students believed that these FA practices are helpful as they enabled them to determine their strengths and weaknesses and to identify what to do to improve their learning, no significant improvement was found in students' test scores between the two STs. Moreover, students seem to always favor their teachers' assessment over that of their peers or themselves. These results challenge the entrenched beliefs about FA as the ultimate tool to enhance students' learning outcomes and open up more venues for further research that land more powerful empirical support for its prospective benefits on learning.

Keywords: formative assessment, summative assessment, feedback, peer-assessment, self-assessment

1. Literature Review: Formative Assessment

Assessment as the set of instruments and procedures whereby evidence about students' achievement is collected, analyzed and interpreted (Brown, 2008) has always been the subject of debate among researchers and practitioners. Over the last decades FA constituted one of the overly emphasized and discussed issues in this area essentially because of its prospective positive impact on both learning and instruction. Since its first denotation in 1976 by Scriven's distinction between summative and formative program evaluation, FA has been the subject of interest among researchers. In this regard Black and Wiliam (1998) defined this concept as "all those activities undertaken by teachers and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged" (p. 10). It is also specified that a classroom assessment practice is considered formative to the extent that the information gathered from assessment is used with the purpose of improving instruction in order to meet students' needs (Popham, 2006). This process is fairly frequent and can occur at any time during instruction as among the reasons justifying teachers' interest in FA is their awareness that summative standardized assessments are not frequent enough to have an impact on their daily instructional decisions contrary to FA (Stiggins, 2005). Hence, information derived from FA is more likely to be used by teachers as well as students in order to make well-informed decisions about the next steps in instruction that are believed to be better grounded than the decisions they would have taken without the evidence that was gathered (Popham, 2006).

Following the same line of reasoning, FA was largely associated with several constructive feedback-related functions as it can serve diagnosis, prediction as well as evaluation purposes (Black & William, 1998) which are all believed to enhance learning and teaching. Moreover, what vividly stands out when implementing FA practices is the role ascribed to students in the assessment of their own performance by means of self-assessment and that of their peers as well using peer-assessment. Within the process of self-assessment, students are expected to evaluate their own work in light of the shared characteristics and standards of quality performance (Brown & Harris, 2013). During this process, learners reflect upon the quality of their own work, compare it to the explicitly provided criteria, and consequently draw conclusions as to one's strengths and weaknesses and make suggestions as to ways for improvement (Andrade & Valtcheva, 2009). The primary purposes of self-assessment are to enhance learning and achievement as well as to promote student self-regulation which is manifested in one's tendency to monitor and manage his/her own learning rather than rely on the teacher as the only agent of assessment (Pintrich, 2000; Zimmerman & Schunk, 2004).

Peer-assessment on the other hand, is the evaluation of a student's output by a fellow student. The major purpose of peer-assessment is the exchange of rich feedback emphasizing the strengths and weaknesses of each one's work. This process is believed to help learners to develop a formative view towards assessment within which the latter is seen as an opportunity to reveal one's gaps in learning mainly for the purpose of overcoming them rather than for scoring or ranking objectives (Brown, 2017). It is also emphasized that students' engagement in peer-assessment activates their metacognitive skills as they are involved in evaluating, justifying, explaining and using their discipline knowledge in order to provide constructive

feedback for their peers' work (Brown, 2017).

While students play a crucial role in the process of FA, teachers remain the main agents behind its successful implementation. In fact, teachers are required to convey to their students the expected learning goals, the quality of a good performance along with the used marking scheme. In fact, Sadler (1989) argued that, if improvement in learning is to take place, students need to have a concept of quality performance similar to the one held by their teachers. In terms of implementation, studies on FA emphasized five main FA practices that teachers need to use if improvement in learning and instruction is to take place namely: 1) Sharing success criteria with learners 2) Classroom questioning 3) Comment-only marking 4) Peer- and self-assessment and 5) Formative use of summative tests (Black et al., 2003).

From this perspective the positive impact of FA on both students and teachers has become an unchallenged belief in education. In fact, it was asserted that the implementation of FA in EFL classrooms leads to a noticeable increase in learners' engagement with classroom activities, a substantial boost in their motivation and an embracement of a positive attitude towards learning and self-improvement. While the positive impact of FA practices on learning and instruction became a tacit agreement among researchers and practitioners in reality there is a scarcity of empirical evidence that supports these outcomes. In fact, when it comes to the effectiveness of FA practices, it is mainly Black and William (1998) study that was mostly referred to where researchers reviewed more than 250 articles that investigated the issue of FA reaching the conclusion that the research they reviewed "shows conclusively that FA does improve learning," and that the improvements in students' achievement were "amongst the largest ever reported" (p. 61). On the other hand, in their critical analysis of the literature regarding FA, Dunn and Mulvenon (2009) asserted that a significant number of the selected articles in Black and William (1998) seminal work were criticized on the basis of serious methodological concerns. To mention for example, Fuch and Fuch's (1986) meta-analysis examining the impact of FA practices on student achievement, this study was criticized on the basis of the selected population as 83% were students of special needs (handicapped) which compromises the generalizability of its findings. Another study that Black and William (1998) used to support their conclusion was Whiting, Van Burgh, and Render (1995) that was also criticized on the basis of its sample size (only one teacher was studied). Accordingly, to provide further empirical evidence on the impact of FA on learning, our study attempted to answer the following two research questions:

- 1) What are the formative assessment practices implemented by secondary school EFL teachers?
- 2) What is the impact of the implemented formative assessment practices on students' learning from students' perspectives?

2. Context of the Study

This study was conducted in a high school In Tunisia. In the context of this study (Tunisia) assessment generally and summative tests more precisely are of paramount importance for both students and teachers. As far as the English language is concerned, high school students

are assessed on their mastery of the specific given input that is largely based on grammar and specific content knowledge-six times during an academic year by means of three mid-term and three end of term summative tests. Students' scores in the English exams are more likely to have a serious impact on their overall academic results and can often stand as a barrier to their success. Under these circumstances these tests are considered of high stakes. Concerning EFL teachers' assessment constitutes an impartial part of their jobs as they are required to develop, implement and correct their students' summative tests and to use other FA practices more frequently. Despite the importance attributed to assessment, future EFL teachers in Tunisia are introduced to only some basic assessment courses during their higher education; the extent to which they helped them to improve their assessment competence has never been proved.

3. Methodology

3.1 Data Collection Instruments

In order to answer our two research questions we used a mixed method approach combining qualitative and quantitative data collection instruments following an exploratory sequential design (Creswell, 2009). In the first phase we collected and analyzed data issued from semi-structured interviews, based on which, we developed the questions of our second data collection instrument i.e., the internet-based survey.

Semi-structured interviews

In order to answer the first research question investigating the FA practices that EFL teachers implemented in their classrooms, especially after seeing their students' performance in STs, we used semi-structured interviews consisting of 5 open-ended questions examining the way EFL teachers conceived their classroom-based assessment, the FA practices they implemented in their classrooms, the types of feedback they referred to, and the impact of these practices on their students' learning. Although the interview was guided by this set of questions, clarifications, deviations and openness towards new directions were not rejected but rather considered as further opportunities for in-depth insights and understanding of the topic in hand. The research questions were both formulated and implemented following Cohen, Manion and Morrison (2007) guidelines.

Data collected from the semi-structured interviews was analyzed by means of content analysis which consists of summarizing and reporting written data, while maintaining and conveying its main content and messages (Cohen, Manion, & Morrison, 2007). The process of content analysis involved summarizing, coding, categorizing, and making sense of the categories by establishing links and connections between them to eventually draw sound conclusions regarding the tackled issue (Cohen, Manion, & Morrison, 2007). Accordingly, the summary of each teacher's answer to every interview question constituted the units of analysis for the data, a thematic code was inserted next to each summary of teachers' answers, this code stood for the most representative word or else theme used or expressed by every teacher. Key topics and ideas were elicited and classified under each main thematic code. These ideas constituted the key themes of the text unit. After analyzing teachers' answers to each interview question, teachers who showed significant similarities in their answers were grouped together and data

were discussed accordingly. Data generated from the semi-structured interviews regarding the used FA practices and the implemented type of feedback were used to formulate the questions of the internet-based survey.

Internet based survey

Concerning the second research question investigating the impact of the FA practices used by the five EFL teachers on students' learning from students' perspective, we used an internet-based survey. We chose this instrument due to its practicality as it enabled us to reach a large number of students at a relatively short time. Additionally, students had the freedom to choose when and where to complete the survey which is believed to add more reliability to their answers due to the comfortable setting in which they answered.

The survey included 10 closed-items that required students to choose one answer from three to more alternatives. The number of questions was reduced to 10 and the questions themselves were concise as much as possible for fear that if they were longer, students would lose patience and interest quickly enough before answering all the questions. The survey was written in English. The first three questions investigated students' gender, age and their perceived English language level. The fourth question described the way students considered their assessment experience, that is to say the feelings they had when being assessed. The fifth question probed into the way students considered their assessment, that is as a chance to learn from their mistakes, an indication of their level or just as a mark. The sixth question examined how students' perceived their teacher's feedback i.e., whether they find it helpful, interesting, not important or boring. The seventh question tapped into students' ability to identify their strengths and weaknesses in light of their teachers' feedback. The eighth question examined students' understanding of their teachers' marking scheme. The ninth question looked into how students would like to be assessed i.e., by themselves, in group with teacher's help or only by the teacher and the last question examined students' potential improvement after receiving the given feedback.

The survey was piloted with five students before its actual use, Lowe (2016). The purpose of the pilot was to unfold, at an early stage, all possible misunderstandings that students might face when answering the questions. In this regard, students participating in the pilot provided us with feedback concerning the questions that hampered their understanding. Consequently, adjustments were made so as to increase the readability of the survey by the participating students. The difficult items were edited to be replaced with simpler and clearer terms and the students were asked to give their opinions regarding every new item, they were encouraged to suggest more options and alternatives as well. The link to the survey was sent to teachers who took the responsibility to send it to their students.

The internet-based survey was conducted with 100 students taught by the same interviewed teachers. Hence, students were divided into two groups according to the same grouping of the teachers by whom they were taught and data were gathered and analyzed accordingly. Data were analyzed by comparing the overall percentage total of each group's answers to every survey question. The percentage total was provided automatically by the survey. No statistical software was used for calculations or presentations as once students' responses entered the

survey; its analysis was processed automatically.

Comparison of students' test scores

In order to cross-check students' answers to the internet-based survey, which examined the impact of the implemented FA practices on their learning from their perspective, we compared students' test scores between two summative tests (one occurring before the implementation of the FA practices and the other one after). Accordingly, 100 students' test scores in the first two consecutive summative tests were obtained. These students belonged to four classes taught by four of the interviewed teachers. The duration of the tests was one hour and they were marked out of 20.

To analyze the obtained data, a paired-samples t-test was performed using SPSS to compare students' test scores between the mid-term and end-of-term STs to see whether there is a significant difference in their test scores between these two tests after the implementation of the specific FA practices by their teachers. On the other hand, since teachers implemented different FA practices with their corresponding groups of students, we wanted to see the impact of these specific practices on each group, hence, we conducted a t-test for independent samples between the two groups of students using SPSS as well.

Prior to conducting the analysis, the correspondence of the data to the requirements of the t-test was checked. In this regard, the four assumptions were satisfied. In fact, the dependent variables consisted of continuous data (test scores), with two scores for each participant (student) who were randomly selected from the population (the total class). Concerning the assumption of normally distributed difference, it was considered satisfied as the skew and kurtosis levels were estimated at -0.110 and 0.544, respectively, which were less than the maximum allowable values for a t-test (i.e., skew < |2.0| and kurtosis < |9.0|) according to Posten (1984). Concerning the test of normality, upon interpreting the Shapiro-Wilk, the p value was 0.773 which is not statistically significant suggesting that the data is normally distributed. Accordingly, the t-test was considered appropriate to use for this analysis.

3.2 Population and Sampling

Semi-structured interviews were conducted with five EFL teachers working in the same High School pertaining to the state of Nabeul (Tunisia). Their profile is described in Table 1 below.

Table 1. Teachers participating in the semi-structured interviews

Teachers	Gender	Age	Teaching experience
Amina	Female	46	21
Boutheina	Female	45	30
Camilia	Female	54	30
Dorra	Female	30	6
Mouhamed	Male	50	25

Regarding the sampling strategy, teachers were selected according to purposive sampling.

Merriam (1998) advocated that purposive sampling is built on the belief that the researcher aims to discover, understand and gain an in-depth knowledge of the subject in hand, therefore must choose a sample from which the maximum can be learned. Hence, these five teachers were selected among a total of nine EFL teachers based on their teaching experience, willingness to collaborate, and the researcher's familiarity with them.

The internet-based survey was conducted with 100 students belonging to 5 classes that were taught by the five interviewed teachers. Teachers were responsible for choosing the classes of students with whom the survey was conducted. The comparison of students' test scores was done by obtaining the same students' test scores in the two consecutive summative tests.

4. Results

4.1 EFL Teachers' Formative Assessment Practices

Results of the semi-structured interviews showed that all teachers conceived assessment as a tool to evaluate their teaching and to measure their students' learning. Alongside these conceptions, Boutheina, Camelia and Dorra conceived assessment as a tool to help their students to improve their learning as well.

Concerning the FA practices teachers implemented in their classrooms, the majority of them referred to making significant changes in their future instruction by including consolidation tasks and exercises that focused on students' gaps in learning. During these tasks and exercises, teachers emphasized that they try to explain to their students why their answers were considered incorrect and reminded them of the course section where they saw the specific area of concern (Boutheina, Camilia and, Dorra). On the other hand, the remaining teachers did not make any changes in their future instruction and they did not refer to any particular tasks. They assumed that they only abided by their pre-established course content even though they expressed the need to implement some FA practices; they were unable to do so due to time constraints (Amina and Mouhamed).

As far as feedback is concerned, the majority of the interviewed teachers referred to providing their students with different types of feedback. In fact, some teachers promoted feedback by means of classroom discussions where they gave their students specific tasks that required group work which were followed by an open classroom interaction. Students were encouraged to discuss their answers among each other and to find the right answer on their own with a minimum of teachers' intervention (Boutheina, Camilia and Dorra). Also, these teachers encouraged their students to exchange their test papers among each other and to detect their peer's mistakes after sharing and discussing with them the used marking scheme. The remaining teachers relied mainly on the given mark as a source of feedback claiming that students only care about their test scores when it comes to assessment (Amina and Mouhamed). Moreover, they provided them with some direct corrective feedback concerning the difficult tasks.

As to the impact of these FA practices on learners' learning from teachers' perspective, the

majority of the interviewed teachers (Boutheina, Camilia and Dorra) assumed that upon implementing these practices, students tend to be more engaged in the assessment process. Teachers also emphasized that these practices entrenched in their students the spirit of collaborative work, acknowledgment of their mistakes and of why they were considered as such and a self-awareness of their actual level in the subject matter. Teachers denoted that these practices rendered the assessment and the learning experience more joyful for their students (Boutheina and Camelia). However, all teachers agreed that not all students benefited from the implemented practices, only the minority of high achievers who cared about learning the subject matter as much as they cared about their marks did. On the other hand, the remaining teachers (Amina and Mouhamed) stressed that most of the students are not generally interested in getting feedback, collaborating with other students, or about improving their learning, they only care about their marks.

4.2 The Impact of FA Practices on Learning: Students' Perspectives

Results of the semi-structured interviews showed that three teachers (Bouteina, Camilia and Dorra) implemented certain FA practices as opposed to the remaining teachers (Amina and Mouhamed). Hence, students were divided in two groups according to the teachers by whom they were taught: Group1 included students taught by Boutheina, Camilia and Dorra, while group2 included students taught by Amina and Mouhamed. Findings of the survey were analyzed by means of comparing the overall percentage total of students' answers to the survey items.

When sending the survey to Group 1 students, 66 students answered (20 male and 46 female), their ages ranging from 17 to 22. Most of these students (34%) viewed the assessment experience as challenging, 23% considered it as stressful, 17% described it as joyful, 16% considered it as scary while 9% described it as motivating. On the other hand, most of the students (37 %) considered the English exam itself as an indication of their level as well as a chance to learn from their mistakes while 26% viewed it only as a mark. Concerning their teachers' feedback, the majority of Group1 students (36%) viewed it as helpful, 32% considered it interesting, 7% considered it boring while 14% thought that it was not important. Regarding students' ability to identify their strengths and weaknesses after receiving their teachers' feedback, most of them, 66%, reported that they are sometimes able to, while 19% believed that they cannot. Only 15% believed that they were able to identify them. Also, 48% of the students reported that they understand their teachers' marking strategy and the reasons why their answers were considered correct or incorrect, 41% believed that they sometimes do while only 11% admitted that they do not. Most of the students (59%) preferred to be assessed only by their teacher, 32% preferred to be assessed in a group with the help of the teacher while only 9% preferred to do it individually. After receiving their teachers' feedback, most of the students 84% believed that they know what to do in order to improve their learning, while only 16% believed that they do not.

Concerning group 2 students (taught by Amina and Mouhamed) 44 students answered the survey (6 male and 38 female) their ages ranging from 17 to 22. Most of these students (35%) described their English assessment experience as stressful, 27% considered it as joyful and

20% viewed it as motivating. A minority of students considered it as challenging (12%) while 5% described it as scary. On the other hand, most of the students (56%), considered the English exam itself as a chance to learn from their mistakes while 24% considered it only as a mark and 19% viewed it as an indication of their level. Concerning the way in which Group2 students considered their teachers' feedback, most of them (42%) viewed it as interesting. Thirty percent considered it helpful, 16% found it boring, while 11% thought that it was not important. Regarding students' ability to identify their strengths and weaknesses after receiving their teachers' feedback, most of them (63%) reported that they are sometimes able to do it, while 27% believed that they cannot, only 9% believed that they were able to. On the other hand, the majority of the students (53%) reported that they understand their teachers' marking strategy and the reasons why their answers were considered correct or otherwise incorrect, 39% said that they are sometimes able to, while only 7% admitted that they do not understand. Most students (68%) preferred to be accessed only by their teacher, 23% preferred to be assessed by their peers with the help of the teacher while only 9% preferred to do it individually. After receiving their teachers' feedback, most of the students (74%) believed that they know what they should do in order to improve while only 25% believed that they do not.

Ultimately, results of the internet-based survey showed that irrespective of the teacher by whom they were taught, all students preferred to be assessed by their teacher and are more likely to learn from their mistakes and to improve their learning in light of their teachers' feedback. Accordingly, to cross-check these results, we compared students' test scores between two summative tests occurring before and after teachers' implementation of the different FA practices. Hence, students were expected to achieve better results in the second test. However, the results of the paired-sample t-test showed that the difference in students' scores between these two tests was not significant. That is, students' test scores in the two tests were mostly the same ($t(99) = -1.248$, $p(1\text{-tailed}) = 0.1075$ which was bigger than 0.05, the Eta was estimated at 0.015 which was a very small effect size based on Cohen (1988)).

Table 2. Paired samples t-test

	Paired differences				t	df	Sig. (2-tailed)	
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower				Upper
Pair1 ST1-ST2	-,35000	2,80354	,28035	-,90628	,20628	-1,248	99	,215

Since the two groups of teachers implemented two different assessment practices and types of feedback, we used an independent-sample t-test to see if there is a significant difference in

students' test scores between the two groups of students. The results of the independent-sample t-test showed that the mean difference of Group 2 students ($M = 0.54$, $SD = 3.22$) was numerically higher than the mean difference of Group 1 ($M = 0.16$, $SD = 2.318$). However, upon analyzing Levene's test for equality of variances, we found that there was no significant difference in students' test scores between the two groups of students (the p value was estimated at $0.099 > 0.05$ suggesting that equal variance assumed. In this respect $p = 0.501 > 0.05$ implying that there was no statistically significant difference between the means of the two groups).

5. Discussion

To start the discussion with EFL teachers' conceptions of assessment, it is commonly believed that what we think about something tends to influence what we do about it, and teachers' conceptions of assessment is no exception. Accordingly, in this study, teachers' use of FA was first examined by a question investigating their conceptions of assessment itself. This argument found its roots, in part, in a larger body of research and theory namely the theory of reasoned action (Ajzen & Madden, 1986; Fishbein & Ajzen, 1975) and its extension to the theory of planned behavior. As far as teachers are concerned, it is believed that teachers' conceptions of assessment strongly influence the way they teach and what their students learn or achieve (Calderhead, 1996; Clark & Lampert, 1986; Pajares, 1992; Thompson, 1992).

Two major conceptions of assessment prevailed in this study. According to the dominant one, assessment was seen as a tool for measuring the success of teaching and learning whereas in the second one it was viewed as a teaching and an improvement aid. When compared to the results of Brown's studies (e.g., Brown, 2002, 2006) some similarities were found. In fact, Brown (2006) study showed that teachers held four major conceptions of assessment. In the first one, assessment was seen as a tool to improve teachers' instruction and students' learning, in the second one, assessment served the purpose of making students accountable for their learning, as to the third one, assessment was used to make teachers or schools accountable and the last one viewed them as irrelevant to teachers' work. We can notice that the first conception is similar to the one held by three interviewed teachers in this study as they conceived assessment as a tool to improve their instruction and students' learning. In this regard, Brown (2002) stated "In the improvement conception, the purpose for assessing students' knowledge, skill, performance, or understanding is to generate accurate information that leads to valid changes in teaching practice or student learning such that improvement in student achievement can be facilitated" (p. 32). Hence assessment was associated with the purpose of making changes in instruction which would cater for learners' improvement. This view towards assessment aligns well with the formative purpose of assessment. Hence, we can suggest that teachers who conceived assessment as an improvement tool are more likely to implement FA practices in their daily classroom-based assessment.

The questions about teachers' FA practices revealed that the majority of the interviewed EFL teachers referred mainly to providing their students with feedback, enhancing classroom discussions, sharing with their students the used marking scheme, and encouraging peer and self-assessment. To start with the latter, it is well established in the literature that peer and

self-assessment are associated with a positive impact on students' learning outcomes (Boud, 1995; Falchikov, 2001; Patri, 2002; Wen, Tsai, & Chang, 2006). A case in point is Black et al., (2003) findings stressing that in practice, peer assessment turned to be an important agent for learners' improvement for several reasons, one among which is that the prospect of the feedback provided in such a practice increases students' motivation to work more carefully and meticulously to refine their answers. In fact, when students are engaged in a process whereby, they provide information to another peer based on their own knowledge or perspective, they tend to care more about the information they give and its wording as well. Moreover, the exchange of information among students is believed to benefit each one of them as it was found that students tend to accept remarks and criticism from their peers that they would not, otherwise, accept from their teachers (Black et al., 2003). It is also assumed that peer-assessment can have a positive impact on students' learning even when the shared information is not totally accurate for this is more likely to trigger more learning when students are engaged in discussions and checking for confirmations (Falshikov, 1996).

Despite the above-mentioned prospective benefits of peer and self-assessment, the results of the internet-based survey showed the opposite. In fact, students belonging to both groups favored their teachers' assessment over that of their peers or themselves. Similar findings were obtained in only a few studies to mention Dolin and Evans (2018) who emphasized that students seemed to prefer the feedback they receive from their teacher more than the one they receive from their peers mainly because they considered it more reliable and trustworthy. From teachers' perspective, Harmer (2001) found that although teachers tend to believe and trust their students' ability to assess themselves and give feedback to their peers, students trusted only their teachers' assessment and considered it more accurate. Following this line of reasoning, students are more likely to learn and improve from their teachers' assessment, even when it is restricted to some oral corrective feedback, more than when engaged in peer or self-assessment. These findings challenge the entrenched belief about FA as a tool whereby learners can improve by means of self and peer-assessment. Accordingly, it is not possible to make firm conclusions as to the effectiveness of these practices, more research needs to be conducted to investigate the circumstances under which peer and self-assessment can cater for students' improvement.

The necessity to enable learners to identify their strengths and weaknesses after sharing with them the criteria of what constitutes a quality performance along with the used marking scheme are considered crucial principles of FA that would bring students' learning forward. In this regard, it is stated that "The criteria for evaluating any learning achievements must be made transparent to students to enable them to have a clear overview of the aims of their work and of what it means to complete it successfully" (Black et al., 2003). For these reasons, the above-mentioned practices were adopted by three of the interviewed teachers (Boutheina, Camelia and Dorra). Hence, we expected that their students would achieve better results than the students taught by the remaining teachers. However, the survey showed that all students, irrespective of the teachers by whom they were taught, advocated that they were sometimes able to identify their strengths and weaknesses and to understand their teachers' marking scheme as well. The same results were obtained even though Group2 teachers relied mainly on

providing their students with a minimum of oral corrective feedback and made no explicit attempts to entrench in their learners the ability to evaluate their own work or that of their peers, they never mentioned the practices of sharing their marking schemes or showing their students the characteristics of quality performance either. Ultimately, our results show that students' abilities to identify their own strengths and weaknesses and to elicit their teachers' marking scheme cannot be exclusively associated with the FA practices employed by their teachers.

All the interviewed teachers in this study referred to providing their students with feedback by marking, yet only two of them considered it as a form of feedback. Although, in the literature the term 'feedback' is largely associated with improvement of both learning and instruction, the term 'marking', on the other hand, is mostly associated with the opposite and is depicted in a rather negative way. In fact, Black et al. (2003) asserted that feedback given in the form of grades enhances ego instead of learning involvement. This is explained by the fact that whenever there are grades, students tend to engage in pointless comparisons with other peers where they boost with their marks rather than think of ways for how to improve or to help their peers improve. The internet-based survey showed that both groups benefited from the feedback provided by their teachers as it enabled them to determine what they need to do in order to improve their learning. This implies that students, from both groups, developed an understanding of their mistakes that is to say of why their answers were considered incorrect, they got an insight into their strengths and weaknesses, and they developed the ability to overcome their gaps in learning and by extension to improve (Nicol & McFarlane-dick, 2006). These positive results were obtained from both groups irrespective of the amount or type of feedback employed by their teachers. This can be explained by the sufficiency of the minimum oral corrective feedback that students received. Similarly, it was found that teachers' explicit corrective feedback can be the best way to show students how to improve their performance (Gitsaki & Althobaiti, 2011). Accordingly, the FA practices that teachers implemented and the focus they put on the various types of feedback do not seem to have any significant impact on learning.

Focusing on learners' prospective improvement after receiving their teachers' feedback, students' answers to the internet-based survey suggested that their performance in the following summative test is more likely to improve as they reported that they knew what they should do in order to overcome their gaps in learning and to ameliorate their performance. However, results of the paired-samples t-test showed the opposite as there was no significant difference in learners' test scores between the first and the second summative test. These results cast some doubt as to the truthfulness of students' answers and their ability to determine their level of competence in this subject matter. In fact, there is a significant body of research showing that students tend to overestimate their level (Ünalı, 2016), that is to say they tend to assign themselves high levels of English language proficiency, whereas, in reality, their level is much less than what they think it is. On the other hand, adopting a formative approach to assessment was believed to help learners to assess themselves plausibly and objectively through evaluating their performance and comparing it to the shared criteria. Accordingly, learners not being able to determine their English language level and what they are able to do in this subject matter questions the effectiveness of the FA practices implemented by their

teachers and by extension the prospective benefits of FA itself on learners' ability to assess themselves.

Continuing with the idea of learners' prospective improvement after receiving their teachers' feedback, all the interviewed EFL teachers reported that only the minority of good students who cared about their marks as much as they did about their learning actually improved after receiving their feedback. On the other hand, the results of the internet-based survey showed that most of the students cared about their learning and improvement more than they did about their marks. This dichotomy suggests that teachers might not be well knowledgeable about their students. In fact, research on teachers' knowledge about their students, as important as it seems, was the focus of only a few studies in the literature as asserted by Mayer and Marland (2007). On the other hand, the extensive literature on FA substantiated that by using ongoing FA practices such as quizzes, discussions and observations teachers were more likely to gain an insight into their learners' strengths, weaknesses, level, misconceptions, understanding, etc. which would constitute an exhaustive picture of their students' overall profile. However, even after claiming the use of certain FA practices EFL teachers, in this study, are apparently still not able to make a reliable generalization as to their learners' level.

In light of the above, it can be explicitly stated that teachers' FA practices implemented by means of feedback, peer and self-assessment and sharing the used marking scheme do not necessarily lead to learners' improvement. Accordingly, understanding the causes of students' stagnation in terms of test scores would be necessary if any improvement in their learning and in their test performance is to take place. In this regard, several studies were conducted in an attempt to understand the reasons influencing students' academic performance (Karande & Kulkarni, 2005; Kivimaki, 1995; Rahamneh, 2012). Findings of these studies reported that several factors can negatively affect learners' performance such as their lack of motivation, continuous absence from classes, inattention to school work and test anxiety. Ostensibly, all these factors are related to learners themselves and not directly to their teachers. In this regard, Alami (2016) study concluded that the main reasons behind students' poor academic results in the English test reside within students themselves primarily whereas teachers' role is not prominent. In fact, students ascribed the reasons for their poor results to their lack of revision at home, excessive absences, not understanding the subject and their lack of interest in the subject. Having said this, it can be assumed that teachers' efforts can do little to ameliorate students' performance as long as the latter do not contribute to and engage with these efforts.

6. Conclusions and Recommendations for Future Research

This study on the impact of formative assessment practices on learning proved that despite the fruitful outcomes that have always been associated with the implementation of FA in everyday classroom-based assessment, these positive results cannot be taken for granted, in fact, several factors should be considered before implementing these practices such as students' preferences and level. Accordingly, this study might contribute to changing the entrenched view of FA as a tool to make positive changes in future learning and instruction triggering, hence, teachers to evaluate the effectiveness of the FA practices they implement and to look for the best FA practices that align with their students' needs and preferences. To do so, teachers need to be

allocated the necessary time and resources which allows them more control over their teaching content and assessment practices. Hence, this study might be an eye-opener to decision makers about the necessity to make teaching programs more flexible so as to give EFL teachers more space for creativity.

This study can be seen as the starting point for other research of paramount importance in the field of assessment and education in general. In this regard, as this study proved that teachers' various types and amount of feedback and FA practices had no significant impact on learners' learning outcomes, further research can address the causes underlying students' stagnation in terms of test scores and the specific FA practices which are more likely to cater for their improvement along with the impact of these practices on high versus low achievers as such low achievers would improve their learning and high achievers would maintain their level and motivation. On the other hand, despite the efforts made by EFL teachers to encourage their students to depend on themselves either by means of peer or self-assessment, students always showed a preference towards the centrality of their teacher as the main source of feedback. Hence, more research is needed on how to entrench in students the values of collaborative work on the one hand and autonomy on the other one as they are believed to constitute 21st century essential soft skills which students need to master if any improvement in their life-long learning is to take place.

References

- Ajzen, I., & Madden, T. J. (1986). Prediction of goal-directed behavior: Attitudes, intentions, and perceived behavioral control. *Journal of Experimental Social Psychology*, 22(1), 453–474. [https://doi.org/10.1016/0022-1031\(86\)90045-4](https://doi.org/10.1016/0022-1031(86)90045-4)
- Alami, M. (2016). Causes of poor academic performance among Omani students. *International Journal of Social Science Research*, 4(1), 126–136. <https://doi.org/10.5296/ijssr.v4i1.8948>
- Andrade, H., & Valcheva, A. (2009). Promoting Learning and Achievement Through Self-Assessment. *Theory into Practice*, 48(1), 12–19. <https://doi.org/10.1080/00405840802577544>
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for Learning: Putting it into Practice* (\$ {nombre}er édition). Open University Press.
- Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>
- Boud, D. (1995). *Enhancing learning through self-assessment*. Kogan Page, London.
- Brown, G. T. L. (2002). *Teachers' conceptions of assessment*. Unpublished dissertation, New Zealand, University of Auckland.
- Brown, G. T. L. (2006). Teachers' conceptions of assessment: Validation of an abridged instrument. *Psychological Reports*, 99(1), 166–170. <https://doi.org/10.2466/pr0.99.1.166-170>

- Brown, G. T. L. (2008). *Conceptions of Assessment: Understanding What Assessment Means to Teachers and Students*. In *Conceptions of Assessment: Understanding What Assessment Means to Teachers and Students*. <https://doi.org/10.1037/t01348-000>
- Brown, G. T. L. (2017). *Assessment of Student Achievement* (1st ed.). Routledge. <https://doi.org/10.4324/9781315162058>
- Brown, G. T. L., & Harris, L. R. (2013). Student self-assessment. In J. H. McMillan (Ed.), *The SAGE handbook of research on classroom assessment* (pp. 367–393). Thousand Oaks, CA: Sage. <https://doi.org/10.4135/9781452218649.n21>
- Calderhead, J. (1996). Teachers: Beliefs and knowledge. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 709–725). New York: Simon and Schuster Macmillan.
- Clark, C., & Lampert, M. (1986). The study of teacher thinking: Implications for teacher education. *Journal of Teacher Education*, 37(5), 27–31. <https://doi.org/10.1177/002248718603700506>
- Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (6th ed.). London and New York: Routledge. <https://doi.org/10.4324/9780203029053>
- Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed., p. 260). Sage Publications, Inc.
- Dolin, J., & Evans, R. (Eds.). (2018). *Transforming assessment through interplay between practice, research and policy*. Springer. <https://doi.org/10.1007/978-3-319-63248-3>
- Dunn, K. E., & Mulvenon, S. W. (2009). A Critical Review of Research on Formative Assessments: The Limited Scientific Evidence of the Impact of Formative Assessments in Education. *Practical Assessment, Research, and Evaluation*, 14, 1–11. <https://doi.org/10.7275/JG4H-RB87>
- Falchikov, N. (1996). *Improving learning through critical peer feedback and reflection. In Different approaches: Theory and practice in higher education*. Proceedings of HERDSA Conference.
- Falchikov, N. (2001). *Learning together. Peer tutoring in higher education*. Routledge Falmer, London.
- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading, Mass: Addison-Wesley Pub. Co.
- Fuch, L. S., & Fuch, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children*, 53, 199–208. <https://doi.org/10.1177/001440298605300301>
- Gitsaki, C., & Althobaiti, N. (2011). ESL teachers' use of corrective feedback and its effectiveness on learners uptake. *The Journal of Asia*, 7(1), 197–219.
- Harmer, J. (2001). *The practice of English language teaching*. England: Longman.

<https://doi.org/10.1177/003368820103200109>

Karande S., & Kulkarni, M. (2005). Poor school performance. *Indian Journal of Pediatr*, 72(11), 961–971. <https://doi.org/10.1007/BF02731673>

Kivimaki, M. (1995). Test anxiety, below-capacity performance, and poor test performance: Intrasubject approach with violin students. *Personality and Individual Differences*, 18(1), 47–55. [https://doi.org/10.1016/0191-8869\(94\)00115-9](https://doi.org/10.1016/0191-8869(94)00115-9)

Mayer, D., & Marland, P. (2007). Teachers' knowledge of students: A significant domain of practical knowledge. *Asia-Pacific Journal of Teacher Education*, 25(1), 17–34. <https://doi.org/10.1080/1359866970250103>

Merriam, S. (1998). *Qualitative research and case study applications in education* (2nd ed.). San Francisco: Jossey-Bass.

Nicol, D., & MacFarlane-Dick, D. (2006) Formative assessment and self -regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199–218. <https://doi.org/10.1080/03075070600572090>

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, 62(1), 307–332. <https://doi.org/10.3102/00346543062003307>

Patri, M. (2002). The influence of peer feedback on self- and peer assessment of oral skills. *Language Testing*, 19(2), 109–131. <https://doi.org/10.1191/0265532202lt224oa>

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451–501). San Diego, CA: Academic Press. <https://doi.org/10.1016/B978-012109890-2/50043-3>

Popham, W. J. (2006). *Assessment for Educational Leaders*. Pearson/Allyn and Bacon.

Posten, H. (1984). Robustness of the two-sample t test. In D. Rasch & M. L. Tiku (Eds.), *Robustness of statistical methods and nonparametric statistics* (pp. 92–99). The Netherlands. https://doi.org/10.1007/978-94-009-6528-7_23

Rahamneh, K. F. A., (2012). Reasons for the low academic achievements among the students of the main stages in selected schools in the province of Al-Balqa. *Ocean Journal of Social Sciences*, 5(1), 31–40.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(1), 119–144. <https://doi.org/10.1007/BF00117714>

Scriven, M. (1967). *The methodology of evaluation*. Washington, DC: American Educational Research Association.

Stiggins, R. (2005). From Formative Assessment to Assessment for Learning: A Path to Success in Standards-Based Schools. *Phi Delta Kappan*, 87(4), 324–328. <https://doi.org/10.1177/003172170508700414>

Thompson, A. G. (1992). Teachers' beliefs and conceptions: A synthesis of the research. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 127–146). New York: MacMillan.

Ünaldı, I. (2016). Self and teacher assessment as predictors of proficiency levels of Turkish EFL learners. *Assessment & Evaluation in Higher Education*, 41(1), 67–80. <https://doi.org/10.1080/02602938.2014.980223>

Wen, M., Tsai, C., & Chang, C. (2006). Attitudes towards peer assessment: A comparison of the perspectives of pre-service and in-service teachers. *Innov Educ Teach Int.*, 43(1), 83–92. <https://doi.org/10.1080/14703290500467640>

Whiting, B., Van Burgh, J. W., & Render, G F. (1995). *Mastery learning in the classroom. Paper presented at the Annual Meeting of the American Educational Research Association.* San Francisco, CA.

Zimmerman, B., & Schunk, D. (2004). Self-regulating intellectual processes and outcomes: A social cognitive perspective. In D. Dai & R. Sternberg (Eds.), *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development* (pp. 323–349). Mahwah, NJ: Lawrence Erlbaum Associates.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).