



Trying Self- and Peer-assessment with Intermediate Spanish Students of English

 **Juan Carlos Araujo Portugal,¹**

¹Department of English,
Escuela Oficial de Idiomas,
Burgos, Spain

Corresponding Author: Juan Carlos Araujo Portugal

Phone: +34 6597477368

e-mail: juancarlos.araujo@yahoo.es

Article citation: Araujo Portugal, J. C. (2020). Trying self- and peer-assessment with intermediate Spanish students of English, *Applied Linguistics Research Journal*, 4(2): 40–53.

Received Date: September 2, 2019

Accepted Date: January 14, 2020

Online Date: March 5, 2020

Publisher: Kare Publishing

© 2018 Applied Linguistics Research Journal

E-ISSN: 2651-2629

ABSTRACT

Assessment is a vital element in any teaching-learning process. Traditionally, the teacher has been the only one responsible for it, especially in the case of summative assessment. The information gathered from it allows teachers to make important decisions as regards their students' learning process, e.g. to obtain a certificate, to promote onto a higher level, etc. However, new forms of assessment, mainly self- and peer-assessment, are becoming more and more popular, and are being seen as ways of formative assessment that may help students make informed decisions about their own learning process, while becoming more autonomous at the same time. This paper presents the first experience of self- and peer-assessment of two groups of intermediate Spanish students of English and their main impressions about it. According to them, it seems that most of the students consider the experience of alternative assessment useful, particularly as a way to prepare their oral exam. It is hoped that it will serve as the basis for autonomous work and learning in this sense..

Keywords: Assessment; Assessment experience; Learning process; Peer-assessment; Self-assessment.

1. Introduction

Assessment plays a key role in any teaching-learning process, as if favourable, it may have different consequences – to promote onto a higher level or course, to obtain a certificate that proves that students have reached a specific level, to inform both teachers and students of the latter's progress in their learning process, etc. In this sense, Wang and Wang (2007) believe that the main aim of assessment should be to guide students throughout this process as well as to provide them with feedback.

In a traditional view of education, the teacher has had a central role in the assessment process, and has been the only agent in charge of it, which involved a good deal of burden and responsibility on the teacher. Besides, it has normally been associated with teacher-centred approaches to education, and this assessment carried out by the teacher was mostly summative. However, during the last few years there has been a gradual shift to methodologies which are more student-centred, which means

that students must adopt an active role and take responsibility for their own learning. This has also been reflected in assessment, and despite the maintenance of summative assessment, alternative types of assessment such as self- and peer-assessment have started being implemented. These are common types of formative assessment whose main aim is to inform students of how they are progressing in their learning process, as well as to prepare them to become lifelong learners and more autonomous in their learning process.

2.Literature Review

As regards assessment, there are two main types of assessment, namely, formative and summative assessment. The former is carried out while learning is taking place, whereas the latter tends to happen at the end of the learning process (Whitlock & Nanavati, 2013, p. 34). As these authors highlight, nowadays more importance is attached to what students learn than to what they are taught, as it helps establish how effective the instruction they have received is. Thanks to formative assessment, and by judging students' strengths and weaknesses during their learning process, teachers are able to gather the necessary evidence that will allow them to adjust their teaching so that their students are more likely to succeed in their learning process (Ortega & Minchala, 2017, p. 159).

2.1.Main difficulties in assessment

So as to ensure reliable assessment of students' performance, a wide range of assessment methods and strategies should be used. Until quite recently, in traditional education systems the teacher was the only agent in charge of such an essential process (Birjandi & Tamjid, 2010, p. 211). Apart from this, it must not be ignored that evaluating is subjective (Patri, 2002, p. 110). Specifically referring to language learning, Underhill (1987) highlights that assessing oral performance is even more subjective "because in an oral test what is assessed is the oral 'message' that is communicated between people or conveyed by the speaker" (Patri, 2002, p. 111).

As Patri (2002) also highlights, it is not only its assessment, but speaking itself is particularly difficult if compared to other skills like reading, listening and writing, especially for second language (L2) learners. The reason for this is that students have to focus on many things while they are speaking. For example, they need to pay attention to the content of the message they are trying to convey, their pronunciation, their diction, their body language, they have to establish eye contact, etc. Shore, Shore and Thornton (1992) add that some of these aspects, namely, appropriacy, fluency and clarity, are particularly subjective when it comes to assessing them.

Therefore, it may be easily concluded that teacher assessment can only offer a limited view of students' progress in their learning process, which is why alternative forms of assessment have started being implemented. These new types of assessment attempt to train students to assess themselves, and thus identify their strengths and weaknesses (Birjandi & Tamjid, 2010, p. 211). These new ways of assessment also represent a move from the traditional testing of knowledge to an approach where assessment is learned, and where students are required to adopt an active role in both their learning and assessment processes (Lindblom-ylänne, Pihlajamäki & Kotkas, 2006). This will also help equip students with the necessary skills to become successful lifelong learners (Rowntree, 1987).

2.2.Alternatives forms of assessment

As already mentioned above, new types of assessment mainly have to do with formative assessment, which focuses on the learning potential of assessment (Taras, 2008). Formative assessment may be defined as "assessment that is specifically intended to provide feedback on performance to improve and accelerate learning" (Nicol & Milligan, 2006, p. 64, cited in De Grez, Valcke & Roozen, 2012, p. 129). The two main types of alternative assessment are self- and peer-assessment, where the students adopt a more active role whereas teachers reduce theirs (De Grez *et al.*, 2012; Ozogul & Sullivan, 2009).

Both self- and peer-assessment have gained popularity over the last few years as a result of the increasing interest in learner autonomy. However, as Patri (2002, p. 110) highlights, several studies on these two alternative forms of assessment have revealed that students either over- or underestimate their own performance or that of their partners, which questions the reliability of these new ways of assessment. It seems that best way to equip students with effective assessment skills as regards self- and peer-assessment is to train them as well as to provide them with enough practice (Patri, 2002, p. 126). Birjandi and Tamjid (2010, p. 211) state that this is also becoming common with new approaches to teaching. De Grez *et al.* (2012, p. 139) also highlight the importance of this training and partly justify that the results of their research study were not as good as expected because the participants in it had not received enough training.

Specifically referring to oral skills, Shore *et al.* (1992) pointed out that students find appropriacy and fluency more difficult to assess than other aspects of their oral performance. It also seems that less competent students are less skilful at these forms of alternative assessment than more competent ones (Jafarpur, 1991; Orsmond, Merry & Reiling, 1997; Pond, Ul-Haq & Wade, 1995; Sullivan & Hall, 1997).

Apparently, the most common way to establish to what extent self- and peer-assessment are reliable is to compare them with teacher assessment, and thus to determine to what extent they are similar or they differ (De Grez *et al.* 2012, p. 139). These authors also suggest comparing both self- and peer-assessment.

2.3. Self-assessment

Andrade and Du (2007) define self-assessment as “a process of formative assessment during which students reflect on and evaluate the quality of their work and their learning, judge the degree to which they reflect explicitly stated goals or criteria, identify strengths and weaknesses in their work, and revise accordingly” (p. 160). Birjandi and Tamjid (2010, p. 211) state that some authors like Matsuno (2009) believe that self-assessment is an effective tool in the sense that it helps students understand the goal of the tasks and assignments they are asked to carry out, as well as the criteria according to which they will be assessed, which will ultimately improve their learning process. Similarly, Kavaliauskiene (2004) maintains that self-assessment enables students to reflect on their own progress as well as to establish how this can be changed, adapted or improved. This author adds that, among others, one of the reasons for the implementation of self-assessment is that teachers find it very difficult to monitor the learning needs of their students, which keep changing over time. In order to carry out self-assessment, students need to use specific criteria as well as to apply standards against which they can judge their work (De Grez *et al.*, 2012, p. 130)

Oscarson (1989) identifies six reasons for the use of self-assessment, among which the following may be highlighted:

- a) It fosters learning as it serves as training in evaluating students' performance, which is positive for their learning process.
- b) It allows both students and teachers to be highly aware of the former's abilities, and thus provides both of them with information so as to make informed decisions about which abilities they need to focus on.
- c) It is terribly motivating so that students can set their own learning goals by reflecting on them.
- d) It may bring about benefits in the long run, as in order to become autonomous learners, and thus prepare them for lifelong learning, students must be able to assess to what extent they are progressing in their learning process.

Brantmeier, Vanderplank and Strube (2012) report a number of studies on self-assessment and their findings. For example, Blanche and Merino (1989) examined several studies on self-assessment in language testing in a detailed way, and came to the conclusion that the accuracy of these assessments has to do with the skills which were being assessed, together with the materials employed in the testing context. Blanche and Merino considered students' self-assessment to be

good or very good. However, Ross (1998) proved that there is not uniformity as regards accuracy when students assess themselves.

De Saint-Léger (2009) researched into how self-assessment skills as regards speaking evolved over a semester. This author concluded that the students' abilities in terms of the accuracy of their self-assessment improved throughout this period of time. De Saint-Léger & Storch (2009) pointed out that self-assessment also had a positive influence on the students being more willing to speak in class and it also contributed to an increase in their self-confidence.

As mentioned above, there seems to be some mistrust on the validity of students' self-assessment. Specifically, Patri (2002) states that "there is much research evidence that students performing self-assessment tend to overrate low performance and underrate high performance" (p. 120). Related to this, Brantmeier *et al.* (2012) highlight that the findings that they refer to in their article seem to support the "hypothesis that there is a threshold beyond which learners are better able to assess themselves and learn independently" (p. 147). In the case of language learning, these authors specifically mention Stern (1983), who seems to suggest that in order to be reliable self-assessors students must have reached at least an intermediate level.

2.4. Peer-assessment

Falchikov (2007) defines peer assessment as the kind of assessment that "requires students to provide either feedback or grades (or both) to their peers on a product or a performance, based on the criteria of excellence for that product or event which students may have been involved in determining" (p.132). According to Topping (2009), it is "an arrangement for learners to consider and specify the level, value, or quality of a product or performance of other equal-status learners" (pp. 20-21). In order to do so, students have to "use criteria and apply standards to the work of their peers in order to judge that work" (Falchikov, 2005, p. 27).

Planas Lladó *et al.* (2014) list some of the benefits in students' learning process that several research studies have reported as regards peer-assessment. Among them, the following ones may be highlighted:

- It facilitates student participation, learner autonomy and enables students to take responsibility for their own learning.
- It increases students' motivation.
- It helps students understand content better, which means learning is more meaningful and deeper.
- Assessment is another aspect of the learning process, which involves regarding mistakes as a chance to learn rather than a failure.
- It helps develop students' analytical skills.
- It provides students with skills which will prove useful for their career prospects and to become lifelong learners.

Despite the above-mentioned benefits, Planas Lladó *et al.* (2014) also draw attention to some of the negative aspects which some authors have identified with peer-assessment. They mention the following ones as examples:

- The accuracy and validity of the feedback provided has been questioned.
- Students are not sure about this assessment.
- Students find it difficult to give a mark.
- Students tend to overmark or undermark their partners' work.

As regards students' overmarking, Vickerman (2009) points out that this normally happens in those cases where peer-assessment is not carried out in an anonymous way, as students fear to be believed to be penalising or criticising their partners rather than assessing their work objectively. Pond *et al.* (1995) call this overmarking tendency 'friendship marking' or 'decibel marking'. However, when carried out properly, it seems that by assessing the work of their partners, students can understand the subject matter much better (Vickerman, 2009).

This tendency to overmarking seems to be supported by other authors. For example, Woolhouse (1999) remarked that students found it difficult to be honest when judging their partners' work. Similarly, Oldfield and Macalpine (1995) pointed out that students found it difficult to grant their partners low marks from an emotional point of view.

Despite the risk of lack of reliability of peer-assessment because students' work might be overrated, Patri (2002) considers that while interacting with their partners, the assessment of their work becomes more objective, and they can obtain more valuable information about the quality and level of both their performance and that of their partners. This will possibly also allow them to acquire a better understanding of the assessment criteria. This is supported by the findings of Planas Lladó *et al.*'s (2014) research study where about 60% of the students who took part in their research study believed that thanks to peer-assessment they had gained a better understanding of the assessment process.

Once implemented, and in spite of an initial reluctance to peer-assessment due to their lack of experience in it, students seem to show a positive attitude towards it (De Grez *et al.*, 2012; Langan *et al.*, 2005; Lindblom-ylänne *et al.*, 2006; Planas Lladó *et al.*, 2014). This leads De Grez *et al.* (2012) to believe that students will be more willing to take their partners' feedback into consideration and do something with it. This is in accordance with what Gibbs (2006) and Yorke (2003) say about the role of feedback in assessment. According to these authors, apart from being of good quality, students have to do something with it, because otherwise it will be completely useless.

2.5. Description of the assessment experience

This paper is based on an activity of alternative assessment that was carried out with two groups of students of the first year of the Intermediate Level B2 at the *Escuela Oficial de Idiomas* of Burgos, Spain. I was in charge of both groups of students for the whole academic year. The *Escuelas Oficiales de Idiomas* are the officially authorised institutions for the teaching of foreign languages in Spain. They are also language accreditation institutions that issue certificates that state students' language proficiency in each of the language taught in them in accordance with the Common European Framework of Reference for Languages (CEFR).

The students that attend these language schools are young adults and adults. As for their occupations, they are quite varied – secondary school students, university students, working professionals, people who are unemployed, housewives, OAPs, etc. As these studies are not compulsory education, some students hardly ever come to class or stop doing so at any time during the academic year due to family, work or study commitments. However, the tuition fees entitle them to take the official exam at the end of the academic year in June, and to re-sit it in September if they have not passed all the exam papers in June, even if they have not attended the lessons. Besides, students may enrol just to take the official exam under the same conditions as those ones who come to class regularly, but without attending any lessons whatsoever.

In order to understand the purpose of the activity that is reported in this paper, it is necessary to point out that so as to pass the course, students have to take and pass four exam papers, namely, reading, listening, writing and speaking. The oral exam or speaking part, which is one with the highest failing rates and which students find particularly difficult to prepare, consists of two tasks – a monologue and a dialogue. Throughout the year, the students do several training exercises consisting of a monologue, a dialogue, or both of them. This paper focuses on the self- and peer-evaluation that the students carried out based on a training exercise for a monologue.

3. Method

In order to carry out this assessment experience, firstly two groups of intermediate students did an oral activity that was based on some guidelines on giving a talk, more specifically focusing on the useful language that may be used for each part of talk giving, once the students have done a listening exercise which served as a model. When the oral activity was completed, the students were introduced to a self- and peer-assessment activity on how they did the oral activity, taking into account six criteria provided by their coursebook. In order to self-assess themselves and assess

their partners' performance they are asked to use a three-point rating scale.

The purpose of this experience was to ascertain whether students who had no or little experience in alternative forms of assessment were able to do it in a fairly reliable way. Apart from this, it also attempted to establish whether after the awareness raising stage, students were capable of appreciating the usefulness of these two types of assessment, and thus encourage them to start putting these new forms of assessment into practice. Therefore, this would be a way to become more autonomous learners and more familiar with a learner-centred rather than teacher-centred approach to language learning.

3.1. Participants

As mentioned above, the assessment experience was carried out with the groups of intermediate students that I was in charge of. In each group, there were 14 students enrolled and when the activity was conducted the attendance rate was quite different in each of them. In the first group, there was an average attendance of 6-7 students per session, whereas in the second group about ten students attended classes regularly.

As this is not compulsory education, some students never came to class or stopped doing so when they found a job, began a course, had to do their placement, etc. This was particularly significant in the case of Group A, where five students stopped coming to class because five of them experienced one of the above-mentioned situations. In the case of Group B, this only happened to one person. In both groups there were some students who had not stopped coming to class, but did not do it very frequently. Moreover, in both groups there was also one student in each group that attended lessons for a few days at the beginning of the course, but did not return afterwards. Finally, in Group B, there was one student that had never come to class. As regards their sex and age group, this information is to be found in Table 1 and Table 2. In order to establish the students' age, 31st December has been taken as reference. For more detailed information about the students, the appendix may be consulted.

Table 1. Make-up of both groups of students as regards their gender

	Males	Females
Group A	2	12
Group B	4	10

Table 2. Make-up of both groups of students as regards their age groups

	Group A	Group B
In their 20s	5	3
In their 30s	3	6
In their 40s	6	2
In their 50s	0	2
In their 60s	0	1

It is worth mentioning that if the students from one group could not come to class on one of the days they enrolled for, they could do it on the corresponding day for the other group, and in this way they did not miss a lesson. This had become particularly common with two students from Group A. As regards their occupation, the students from both groups were university students; unemployed housewives who were looking for a job and who stopped coming to class when they found one or they were required to do a training or retraining course by the National Employment Department; working professionals who work in the afternoon and/or evening rather than in the morning and early afternoon.

3.2. Materials

In order to carry out this assessment experience the textbook for the course – *Voyage B1+* by Roberts, Buchanan & Pathare (2015) – was used. This coursebook includes a section on speaking and writing, and the section devoted to speaking in one of the units focuses on useful language for giving a talk, which may be considered similar to a monologue. These expressions were divided into three parts – the introduction, the main body and the conclusion. Within each part, there were some subdivisions. For example, the introduction contained useful phrases to introduce oneself and the topic, as well as to present the structure of the talk. The main body highlighted expressions to sequence the different topics of the talk, and to indicate the beginning of a new section. As for the conclusion, useful phrases to indicate the start of the conclusion and to invite questions were included. Previous to this section with useful expressions, there was an example of a typical talk or monologue divided into three parts, which served as the basis for three listening exercises.

3.3. Procedure

As stated above, the assessment experience was done using the section of Speaking on giving a talk (Roberts *et al.*, 2015, p. 83) as a basis. As also pointed out, previous to this section there were three listening exercises of an example of a talk, more specifically on the factors that affect decision-making. Therefore, to begin the experience the listening exercises were done in class. After the listening practice, the textbook presented the useful expressions for giving a talk divided into different groups and the students had to decide which subsection of a talk they corresponded to. This was done as a whole group and then they were seen again in the language section of this part of the unit. Immediately after that, there was an exercise where students had to prepare a two- to three-minute talk and deliver it to their partners. In order to do that, some topics for the talk were suggested, but if the students did not find them appealing, they could choose a topic of their preference.

The students were given about five minutes to prepare their talk, and were asked to try to use the expressions seen in this lesson to introduce the different parts of their talk. Instead of giving their talk before the whole group or in small groups, they were asked to do it in pairs. Once both students had finished their talks, they were first asked to self-assess their performance, and afterwards to compare their self-assessment with their partners' assessment on their performance. In order to do so, both students had to focus on the following aspects (Roberts *et al.*, 2015, p. 83):

- I spoke slowly and clearly
- I was happy with my use of grammar
- I used a range of phrases from the key language box (giving a talk)
- I made appropriate eye contact with the audience
- I used positive body language (e.g. I looked confidently and smiled)
- I was happy with my use of vocabulary

For their assessments, they used a 3-point Likert scale where the value assigned to each figure is the following (Roberts *et al.*, 2015, p. 83):

1. Needs improvement
2. OK
3. Excellent

On the day that this assessment activity was carried out, in Group A four students attended the lesson. These students usually came to class and worked with the same partners in those activities done in pairs. As for Group B, the day of the activity taken as reference for this paper, twelve students came to class. It is necessary to mention, that two of these students were from Group A, but they had been coming to the lessons for Group B lately, because it suited them best. Apart from that, one of the students that was in class that day occasionally came to class. As regards their grouping, except for one pair, for the rest of them their partner was the student who they usually worked with when an exercise or activity was carried out in pairs.

In both groups the students worked in pairs, and after finishing their talks or monologues, they self-assessed themselves using the rating scale and taking into account the aspects that were suggested in their textbook. They all had enough time to do it, and also to compare their self-assessment with that which their partners had carried out concerning their performance. As encouraged to do, they justified their assessment for both their self-assessment and that on their partners' performance. Once they had finished doing this in pairs, there was a short discussion as a whole group to find out whether the students' self-assessment coincided with their partners' assessment, or whether any relevant differences were perceived.

To finish this activity, the students were told that this was a very useful learning strategy, as well as training process, both to be done individually (self-assessment) and in pairs (peer-assessment), to prepare their oral exams, particularly the individual turn or monologue. Moreover, they were encouraged to put it into practice from that moment onwards, and to record themselves, so that they could find their self-assessment easier to be conducted.

4. Results

As stated in the previous section, two pairs of students in Group A, and six in Group B took part in the self- and peer-assessment activity that this paper is based on. As far as the results are concerned, these are the most remarkable data obtained from the students' self-assessment:

1. The students tended to be very demanding with themselves, especially as regards their use and command of the grammar and vocabulary used in their talks.
2. They seemed to be more objective as regards how quickly and clearly they spoke, their use of the phrases from the language box, eye contact with their partner and their use of body language.
3. Rather than just saying the grade or figure they had chosen for each descriptor of their self-assessment, most of them justified their choice.

As far as the assessment of their partners' performance is concerned, these are the most remarkable aspects:

1. Most of the students took it seriously.
2. They seemed to be quite objective, especially as regards how quickly and clearly their partners spoke, their use of the phrases from the language box, eye contact with their partner and their use of body language.
3. Most of the students justified the figure they had chosen for each descriptor of the assessment on their partner's performance.

Before presenting the results as such, it is worth mentioning that one pair of students in Group B did not really understand what they were supposed to do as regards the self- and peer-assessment activity. It is important to mention that one of the members of this pair was the student from Group B who occasionally comes to class.

The results of the comparison of the students' self-assessment and their partners' assessment on their performance are shown in Tables 3 and 4. The information is presented in the following way. For each pair, it is stated whether the members were male or female, and their age comes in brackets immediately after their gender. Then, the outstanding results obtained from this comparison are stated.

Table 3. Comparison of self- and peer-assessment Group A

	Gender and age	Result of the comparison
Pair 1	Female (43) – Female (42)	Self-assessment more demanding than peer-assessment
Pair 2	Female (43) – Female (31)	Self- and peer-assessment coincide

Table 4. Comparison of self- and peer-assessment Group B

	Gender and age	Result of the comparison
Pair 1	Female (34) – Male (56)	Don't understand what they are supposed to do
Pair 2	Female (50) – Female (37)	Self-assessment more demanding than peer-assessment
Pair 3	Male (64) – Male (24)	Self- and peer-assessment coincide
Pair 4	Male (23) – Female (46)	Self- and peer-assessment coincide
Pair 5	Female (31) – Female (41)	Self-assessment more demanding than peer-assessment
Pair 6	Female (39) – Female (34)	For one student self-assessment is more demanding than peer-assessment For the other student, self- and peer-assessment coincide

In Group B, pair 3 point out some of the mistakes, mainly grammatical ones, his partner has made while giving his talk, and they agree on the mistakes committed. The members of the other pairs agree that they need to expand their vocabulary and have a better command of grammar, which both members of each pair also seem to agree with.

In light of the above presented results, it may easily be concluded that some of the students that carried out this training activity on self- and peer-assessment seem to be very demanding with their self-assessment, whereas they show a more understanding attitude when assessing their partner's performance. This may seem to be related to what is pointed out by many authors as regards the lack of reliability of peer-assessment. There is only one case when it just seems to be the opposite. As regards eye contact, one member of pair 1 from Group A disagreed with her partner while self-assessing her own performance when she said that she had made appropriate eye contact with her partner, which her partner contradicted. However, for the rest of the descriptors both students were more demanding with their self-assessment than her partner's assessment of their performance. This might have been due to the fact that establishing whether eye contact is made is more objective than most of the other descriptors.

It is worth highlighting that three pairs (one from Group A and two from Group B), as well as one member of pair 6 from Group B, agree with the self- and peer-assessment. This is something that deserves being further researched to establish whether it was just sheer coincide, or they really took the activity seriously and were good at self- and peer-assessment.

5. Conclusions

After doing the activity, it might be concluded that most of the students that took part in it perceived the usefulness of both self- and peer-assessment as part of their learning process, and more specifically as an aid that may help them prepare their oral exam. This seems to be in accordance with authors like De Grez *et al.* (2012), Langan *et al.* (2005), Lindblom-yllänne *et al.* (2006), Planas Lladó *et al.* (2014), who remark students' positive attitude towards peer-assessment. This is the reason why the students were encouraged to continue carrying it out during the course, as well as during their learning process. However, as this activity was only done in class once, there is no evidence whether the students continued doing it outside class or not, and if they did, whether both self- and peer-assessment were implemented or just one of them, and the reasons for this.

As this was the first year that this textbook was used with students of the first year of Intermediate Level B2, I was not aware of this activity until I saw it in the unit where it is included. However, as practice speaking activities serve as training for both the monologue and the dialogue that comprise the final oral exam, it seems to be a good idea to work on this section of talk giving, and

more specifically on the self- and peer-assessment part, after doing the first speaking activity used as a mock oral exam. In this way, students can become familiar with these two forms of alternative assessment and, therefore, see them as formative assessment which should be integrated as part of their learning process as soon as possible.

Moreover, by doing so, it would be much easier to encourage students to continue doing it throughout the year and to conduct it in class more than once, so as to make them aware of its usefulness and as a way to integrate it as part of their learning process. Thus, it would be possible to compare the results with those of De Saint-Léger's (2009) research study, and check whether these students' self-assessing skills also increased as in the case of her research study. By repeating it more than once in class, students would be encouraged to reflect on the learning strategies they may use to improve those skills such as speaking that they tend to find particularly difficult. This reflection may also help them set the foundations for the acquisition of the competence of learning to learn that they will have to develop to become effective lifelong learners, which is one of the ultimate goals of alternative forms of assessment as already mentioned.

Apart from introducing students to self- and peer-assessment as early as possible in the course, it might also be a good idea to modify the rating scale from a three- to a five-point scale. As a result, students can render more accurate assessment as some intermediate figures may prove useful, especially in order to make students aware of how they are progressing throughout the course and what aspects they still have to focus on so as to improve their speaking. The use of a five-point Likert scale seems to be the most common rating instrument for research studies on self- and peer-assessment. For example, Patri (2002) and De Grez *et al.* (2012) use it. Thus, rather than following the three-point scale proposed by the textbook (1. Needs improvement, 2. OK, 3. Excellent), this 5-point rating scale may seem more appropriate:

1. Needs quite a lot of improvement
2. Still needs improvement, but has already improved a bit
3. Acceptable or satisfactory (rather than OK as the textbook suggests)
4. Has already improved quite a lot, but it should still be improved a bit more
5. Excellent

This new scale seems to be more realistic, as well as more informative to students when it comes to self- and peer-assessment, as this wording allows them to focus on more concrete aspects to reflect and focus on, as it states how much a descriptor has improved or whether it is perceived as acceptable or excellent.

As regards the use of different rating scales for both self- and peer-assessment, this does not seem to be necessary so that they are both reliable and effective. For example, Patri (2002, p. 114) clearly states that she used the same questionnaire for teacher-, peer- and self-assessment, and that she only changed the wording in the rating scales to adjust them to the agent in charge of each type of assessment.

Apart from this, it would be really useful to be able to know how students regard these new kinds of alternative assessment, and whether they actually use them outside class, and if they do, to what extent this is reflected in the marks they achieve in their oral exams in comparison to previous years when they did not use formative assessment. Besides, it may help determine whether implementing these new forms of assessment makes students feel more confident when they have to face an oral exam, which most of them find terribly stressful regardless of their age and jobs.

Finally, it must be highlighted that this experience was carried out with two rather small groups of students, and even though the results seem to be in accordance with previous research studies, they are somehow limited. Therefore, it would be advisable to replicate the experience with other groups of students of the same level, or even a higher or lower level of proficiency in the language to corroborate the results presented in this paper.

References

- Andrade, H., & Du, Y. (2007). Student responses to criteria-referenced self-assessment. *Assessment & Evaluation in Higher Education*, 32(2), 159-181.
<https://doi.org/10.1080/02602930600801928>
- Birjandi, P., & Tamjid, N. H. (2010). The role of self-assessment in promoting Iranian EFL learners' motivation. *English Language Teaching*, 3(3), 211-220.
<https://doi.org/10.5539/elt.v3n3p211>
- Blanche, P., & Merino, B. J. (1989). Self-assessment of foreign-language skills: Implications for teachers and researchers. *Language Learning*, 39(3), 313-338.
<https://doi.org/10.1111/j.1467-1770.1989.tb00595.x>
- Brantmeier, C., Vanderplank, R., & Strube, M. (2012). What about me? Individual self-assessment by skill and level of language instruction. *System*, 40(1), 144-160.
<https://doi.org/10.1016/j.system.2012.01.003>
- De Grez, L., Valcke, M., & Roozen, I. (2012). How effective are self-and peer assessment of oral presentation skills compared with teachers' assessments? *Active Learning in Higher Education*, 13(2), 129-142.
<https://doi.org/10.1177/1469787412441284>
- De Saint Léger, D. (2009). Self-assessment of speaking skills and participation in a foreign language class. *Foreign Language Annals*, 42(1), 158-178.
<https://doi.org/10.1111/j.1944-9720.2009.01013.x>
- De Saint Léger, D., & Storch, N. (2009). Learners' perceptions and attitudes: Implications for willingness to communicate in an L2 classroom. *System*, 37(2), 269-285.
<https://doi.org/10.1016/j.system.2009.01.001>
- Falchikov, N. (2005). *Improving assessment through student involvement. Practical solutions for aiding learning in higher and further education*. New York: Routledge Falmer.
- Falchikov, N. (2007). The place of peers in learning and assessment. In D. Boud and N. Falchikov (Eds.), *Rethinking assessment in higher education: Learning for the longer term* (pp. 128-143). Oxford: Routledge.
- Gibbs, G. (2006). How assessment frames student learning. In C. Bryan and K. Clegg (Eds.), *Innovative assessment in higher education* (pp. 23-36). London: Routledge.
- Jafarpur, A. (1991). Can naive EFL learners estimate their own proficiency? *Evaluation & Research in Education*, 5(3), 145-157.
<https://doi.org/10.1080/09500799109533306>
- Kavaliauskienė, G. (2004). Quality assessment in teaching English for specific purposes. *English for Specific Purposes World*. Available at: http://www.esp-world.info/Articles_13/article%20QUALITY%20ASSESSMENT%20IN%20ESP.htm.
- Langan, A. M., Wheeler, C. P., Shaw, E. M., Haines, B. J., Cullen, W. R., Boyle, J., Penney, D., Oldekop, J., Ashcroft, C., Lockey, L., & Preziosi, R. F. (2005). Peer assessment of oral presentations: effects of student gender, university affiliation and participation in the development of assessment criteria. *Assessment and Evaluation in Higher Education*, 30, 19-34.
<https://doi.org/10.1080/0260293042003243878>
- Lindblom-ylänne, S., Pihlajamäki, H., & Kotkas, T. (2006). Self-, peer-and teacher-assessment of student essays. *Active Learning in Higher Education*, 7(1), 51-62.
<https://doi.org/10.1177/1469787406061148>
- Matsuno, S. (2009). Self-, peer-, and teacher-assessments in Japanese university EFL writing classrooms. *Language Testing*, 26(1), 75-100.
<https://doi.org/10.1177/0265532208097337>
- Nicol, D. J., & Milligan, C. (2006). Rethinking technology-supported assessment in terms of the seven principles of good feedback practice. In C. Bryan and K. Clegg (Eds.), *Innovative Assessment in Higher Education* (pp. 64-77). London: Taylor and Francis Group Ltd.
- Oldfield, K. A., & Macalpine, J. M. K. (1995). Peer and Self-assessment at Tertiary Level – an experiential report. *Assessment in Higher Education*, 20(1), 125-132.
<https://doi.org/10.1080/0260293950200113>

- Orsmond, P., Merry, S., & Reiling, K. (1997). A study in self-assessment: tutor and students' perceptions of performance criteria. *Assessment & Evaluation in Higher Education*, 22(4), 357-368.
<https://doi.org/10.1080/0260293970220401>
- Ortega, D. P., & Minchala, O. E. (2017). Assessing students in an authentic and ongoing manner in the English classroom. *Theory and Practice in Language Studies*, 7(3), 159-165.
<https://doi.org/10.17507/tpls.0703.01>
- Oscarson, M. (1989). Self-assessment of language proficiency: Rationale and applications. *Language Testing*, 6(1), 1-13.
<https://doi.org/10.1177/026553228900600103>
- Ozogul, G., & Sullivan, H. (2009). Student performance and attitudes under formative evaluation by teacher, self and peer evaluators. *Educational Technology Research and Development*, 57(3), 393-410.
<https://doi.org/10.1007/s11423-007-9052-7>
- 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>
- Planas Lladó, A., Feliu Soley, L., Fruquell Sansbelló, R. M., Arbat Pujolras, G., Pujol Planella, J., Roura Pascual, N., Suñol Martínez, J. J., & Montoro Moreno, L. (2014). Student perceptions of peer assessment: an interdisciplinary study. *Assessment & Evaluation in Higher Education*, 39(5), 592-610.
<https://doi.org/10.1080/02602938.2013.860077>
- Pond, K., Ul-Haq, R., & Wade, W. (1995). Peer review: a precursor to peer assessment. *Innovations in Education and Training International*, 32(4), 314-323.
<https://doi.org/10.1080/1355800950320403>
- Roberts, R., Buchanan, H., & Pathare, E. (2015). *Voyage B1+*. Coursebook with DVD. Oxford: Oxford University Press.
- Ross, S. (1998). Self-assessment in second language testing: A meta-analysis and analysis of experiential factors. *Language Testing*, 15(1), 1-20.
<https://doi.org/10.1177/026553229801500101>
- Rowntree, D. (1987). *Assessing students: How shall we know them?* (2nd edition). London: Kogan Page.
- Shore, T. H., Shore, L. M., & Thornton, G. C. (1992). Construct validity of self-and peer evaluations of performance dimensions in an assessment center. *Journal of Applied Psychology*, 77(1), 42-54.
<https://doi.org/10.1037/0021-9010.77.1.42>
- Stern, H. H. (1983). *Fundamental concepts of language teaching*. Oxford: Oxford University Press.
- Taras, M. (2008). Summative and formative assessment: Perceptions and realities. *Active Learning in Higher Education*, 9(2), 172-192.
<https://doi.org/10.1177/1469787408091655>
- Topping, K. J. (2009). Peer assessment. *Theory into Practice*, 48(1), 20-27.
<https://doi.org/10.1080/00405840802577569>
- Underhill, N. (1987). *Testing spoken language: A handbook of oral testing techniques*. Cambridge: Cambridge University Press.
- Vickerman, P. (2009). Student perspectives on formative peer assessment: an attempt to deepen learning? *Assessment & Evaluation in Higher Education*, 34(2), 221-230.
<https://doi.org/10.1080/02602930801955986>
- Wang, H., & Wang, Y. (2007). The addition of an affect test and self-assessment into ESL writing assessment: Process and effect. A case study in a non-English major postgraduate ESL writing class. *Asian EFL Journal*, 20(5), 299-316.
- Whitlock, B., & Nanavati, J. (2013). A systematic approach to performative and authentic assessment. *Reference Services Review*, 41(1), 32-48.
<https://doi.org/10.1108/00907321311300866>
- Woolhouse, M. (1999). Peer assessment: the participants' perception of two activities on a further education teacher education course. *Journal of Further and Higher Education*, 23(2), 211-219.
<https://doi.org/10.1080/0309877990230205>

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

<https://doi.org/10.1023/A:1023967026413>

Appendix

Students' Gender And Age

Group A

	Gender	Age
Student 1	Female	46
Student 2	Female	43
Student 3	Female	42
Student 4	Female	30
Student 5	Female	31
Student 6	Female	25
Student 7	Male	24
Student 8	Female	42
Student 9	Female	24
Student 10	Male	24
Student 11	Female	43
Student 12	Female	46
Student 13	Female	37
Student 14	Female	24

Group B

	Gender	Age
Student 1	Male	23
Student 2	Male	56
Student 3	Male	21
Student 4	Female	50
Student 5	Female	37
Student 6	Female	39
Student 7	Female	41
Student 8	Female	46
Student 9	Male	64
Student 10	Female	31
Student 11	Female	34
Student 12	Female	31
Student 13	Female	23
Student 14	Female	34