

# Self-Assessment versus Objective Proficiency Evaluation in English as a Foreign Language among Italian First-Year University Students

Irina Stan 

Department of Education and Human Sciences, University of Modena and Reggio Emilia, Reggio Emilia, Italy

Email: [irina.stan@unimore.it](mailto:irina.stan@unimore.it)

**How to cite this paper:** Stan, I. (2024). Self-Assessment versus Objective Proficiency Evaluation in English as a Foreign Language among Italian First-Year University Students. *Open Journal of Modern Linguistics*, 14, 543-558.  
<https://doi.org/10.4236/ojml.2024.144028>

**Received:** May 29, 2024

**Accepted:** July 5, 2024

**Published:** July 8, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).  
<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

The aim of this study is to investigate first-year non-English-major Italian University students' ability to assess their English proficiency at the beginning of the academic year based on comparisons to official assessments from objective tests. One thousand and ninety freshmen's Placement Test scores were analysed against self-ratings expressed using the CEFR self-assessment scale for Speaking, Listening, Reading and Writing skills. The objective measure was Pearson's Versant Placement Test (VEPT), which is intended for adults and students over the age of 16 and measures facility in spoken and written English. For the self-assessment tool, we applied the table of descriptors for self-assessment from the 2020 Italian version of the Companion Volume (Council of Europe, 2020: pp. 188-190) to which a <A1 level was added. To analyse the data, descriptive statistics and inferential statistics (Pearson correlation) using statistical package for social science (SPSS) version 22 and Excel were used. Tables and figures were also used to present the data. The results of the 4 correlation analyses we conducted showed that students' self-assessments and Placement Test assessment were statistically significant, positive and high for all four skills abilities (0.55 for Speaking, 0.68 for Listening, 0.50 for Reading, and 0.58 for Writing. Moreover, students tend to underestimate their writing skill and overestimate their reading ability. We can affirm that the self-assessments produced by students participating in our study were accurate. One reason for this high correlation may be the self-assessment tool used in the study, namely the CEFR descriptors. This study provides further evidence of the reliability and accuracy of *Can-Do* descriptors in EFL self-assessment and the findings can be applied by both EFL learners and teachers.

## Keywords

EFL, Objective Proficiency Evaluation, Self-Assessment, Undergraduate,

## 1. Introduction

Bailey (1996) defines self-assessment as “procedures by which the learners themselves evaluate their language skills and knowledge” (Bailey, 1996: p. 227). The literature on self-assessment has typically highlighted several advantages: 1) it is cost-effective and relatively easy to design, administer, and score; 2) it can promote greater learner awareness and self-regulation; and 3) it can motivate students by adding variety to, as well as increased participation in, the assessment process (Ross, 1998). Critics have suggested that self-assessment is not appropriate because learners are not capable of accurately gauging their own abilities, and some asserted that it can lead to lower standards, rewarding students who overestimate their abilities (*ibid.*). In spite of these criticisms, self-assessment has been used regularly either as the sole measure of language development or as a complement to other measures in research on teaching English as a foreign language (EFL). To determine the validity of self-assessment as a proxy for more objective measures, researchers in a variety of fields, including first and second language acquisition have administered both self-assessments and objective measures and have found that learners are largely able to make good judgments of their own abilities (Falchikov & Boud, 1989). It was also found that self-assessments prove most useful when they are tied to tasks that learners are likely to encounter and can imagine themselves experiencing. As Oscarson (1997) pointed out, “Self-assessments are more accurate when based on task content closely tied to students’ situations as potential users of the language in question.” Furthermore, “The evidence is that it is easier for learners to assess their ability in relation to concrete descriptions of more narrowly defined linguistic situations” (p. 183).

Bandura (2006) noted that in order to measure a learner’s self-perceived capability, “items should be phrased in terms of can-do” statements (p. 308). In fact, he asserted that valid self-assessment requires such statements. The levels of the Common European Framework (Council of Europe, 2001) offer a description of the linguistic-communicative competencies broken down into descriptors of progressive difficulty. The Framework was, in this sense, the first manual to provide school and university teachers with a key to interpreting the performance of learners of foreign languages through the use of descriptors of competencies and it also enabled students to recognize their own abilities more precisely.

Back in 2002, Weeden et al. (2002) regarding the question “can students self-assess?” believed that “there is evidence that some learners of all ages do have a degree of skill at self-assessment” (p.84). However, despite the importance of self-assessment, learners are rarely put in charge of rating their own abilities and performance (Luoma & Tarnanen, 2003: p. 440) and it seems that evidence in this regard deserves further investigation.

Research has provided a comprehensive understanding of the advantages and practical applications of self-assessment in language learning. It collectively highlighted the importance of integrating self-assessment practices to foster a more autonomous, motivated, and effective language learning experience. For example, [Butler and Lee \(2010\)](#) examined the impact of self-assessment on young English learners in an Asian context. The results indicate that self-assessment can improve learners' confidence and language proficiency, especially when combined with teacher feedback.

In Italy, English learning and teaching mainly takes place in the classroom, rather than during everyday communication. According to [Little \(2005\)](#), involving learners in the assessment process brings a positive impact on their language learning. Moreover, it has been suggested that self-assessment is generally reliable and valid, and can serve as a useful supplement to traditional assessment methods ([Ross, 1998](#)). In this paper we aim to report the results of an analysis that looked at students' ability to self-assess their English proficiency in each of the four main skills, namely speaking, listening, reading and writing.

The present contribution is part of a wider project carried out at a large University in Italy, aimed at investigating, among others, freshmen EFL students' preliminary language competence in English, based on objective tests, their ability of self-assessment and self-evaluation of their weaknesses and strengths as EFL learners and the extra-curricular use of the language as well as the effects of previous learning experience with English.

## 2. Literature Review

### 2.1. Self-Assessment

Research on self-assessment has its roots back in the 1960s and 1970s when research started comparing self-assessments with predictors of academic achievement. Few studies, however, focused on using self-assessment of language abilities until the late 1970s and the mid-1980s. Most of these studies involved correlational studies aimed at looking at the relationship between self-assessment ratings and objective exams or teacher ratings. [Oscarsson \(1978\)](#) reports a 1977 study by Balke-Aurell of Swedish students studying English as a foreign language. Correlations between self-assessment and instructors' judgments were 0.60 while the self-assessment and formal tests correlated about 0.50. In a study of students at the University of Ottawa for placement into French and ESL courses, [LeBlanc & Painchaud \(1985\)](#) found positive correlations (ranging from 0.39 to 0.53 on subtests and 0.53 on total score) between results on a self-assessment instrument and a standardised English proficiency exam.

Among the more recent correlational studies on second language (L2) self- and proficiency outcomes assessment, [Stansfield et al. \(2010\)](#) investigated the extent to which self-assessment scores from 323 learners of eight different languages could help the National Language Service Corps screen applicants for the program. Each applicant completed a two-part self-assessment composed of a series of can-do statements and a simplified set of Interagency Language Roundta-

ble global skill-level descriptions. The correlation between the applicants' scores on the Defense Language Institute Oral Proficiency Interview (OPI) and their self-assessed global speaking scores was moderate ( $r = 0.49$ ). The authors posited that self-assessment could be used as a low-stakes screening measure, a finding that has been supported by other researchers (Malabonga et al., 2005; Tigchelaar, 2018; Tigchelaar et al., 2017). Concomitantly, other researchers have investigated the use of self-assessment in placement testing and found results similar to Stansfield et al.'s (2010) results.

However, while self-assessment seems to aid screening and placement testing, not all research has rendered such positive conclusions and a certain amount of studies focusing on the use of self-assessment instruments have reported conflicting results. Blanche & Merino (1989) found that self-assessments did not correlate with classroom or test performance. In a study focused on how well self-ratings correlated with an English placement exam, Wesche et al. (1996) found "placement via self-assessment was extremely unreliable" (p. 15).

These conflicting results make developing a self-assessment questionnaire difficult for those exploring the possibility of using self-assessment as part of their placement procedures. However, previous studies have described several factors that may have limited the effectiveness of a self-assessment questionnaire, including the wording on the questionnaire, the language skill being assessed, the level of proficiency of the students, the cultural background of the students, and the type of self-assessment task.

Several years ago, Upshur (1975) suggested that students should be able to respond to questions about their abilities in a language using all their experience with the language.

Indeed, providing concrete descriptors with specific examples can help learners more accurately evaluate their abilities. As previously stated, recent efforts to implement self-assessment in second language education have involved using Can-Do statements connected with tasks that are often the focus of language curricula and that learners ought to expect to encounter in authentic real-world situations. Finally, in line with learner-centred language teaching (Little, 2005) and the use of assessment for language learning (Butler, 2016), Can-Do statements are designed to promote language awareness and independence by allowing learners to document what they can do in their second language. While this is a reasonable first step in the construction of language proficiency descriptors and associated Can-Do statements, without psychometric validation it is not certain that the abilities described in the scales and in the statements are valid or useful for measuring language proficiency. In the case of the European scale (Council of Europe, 2001), extensive empirical work has been done to scale the proficiency descriptors using Rasch modeling.

## 2.2. The CEFR and Self-Assessment

The CEFR is the description of the Council of Europe's language policy, produced for the purposes of increasing collaboration and cooperation between

European educational institutions (Trim, 2007). As a system used to describe communicative language competences, the CEFR intends to be extensive, coherent, and transparent. The CEFR is best known for its descriptors of language proficiency, which are divided into five language sub-skills across six levels. Since the CEFR's publication, it has impacted foreign language education industries both within and outside of the CoE's member states with many identifying it as "[an] international standard for language teaching and learning" (North et al., 2010: p. 6). The CEFR is purported to support a number of facets of language education, including the planning of the content, objectives or assessment criteria of language learning programs and language certification, the selection of materials for self-directed learners and for the evaluation of learning or learner progress (Council of Europe, 2001: p. 7). It also intends to provide a set of learner-centred scales which allow for a standardised assessment of proficiency (North, 2011). The CEFR is also criticised, particularly for its usage in assessment (see Alderson, 2007). Further critiques relate to the lack of support to the purported progression of difficulty inherent to the framework which is neither tied to stages of language acquisition nor evidenced by empirically obtained performance samples (Westhoff, 2007). Nonetheless, one of the CEFR's strengths is that its scales of Can-Do statements permit learners to both define their own abilities in their language of study, and plan the direction of their future studies (Council of Europe, 2001). A learner may read a statement and then make a decision regarding their perceived performance of the communicative task implied by the statement. If the learner believes they can perform sufficiently or proficiently within the area of that Can-Do statement, they may move on to responding to a statement from another skill or a more difficult statement within the same skill.

### 2.3. English as a Foreign Language in Italy

Compulsory education in Italy lasts 10 years (from ages 6 to 16) and covers the eight years of the first education cycle and the first two years of the second cycle. The Italian education system is composed of: preschool (children aged 3 to 6), the first education cycle—divided into primary school (ages 6 to 11) and lower secondary school (ages 11 to 14)—and the second education cycle, which offers two different paths: upper secondary school (ages 14 to 19) and vocational education and training (for students who have completed the first education cycle). Since the 2000s, Italy has adopted and implemented the Council of Europe's recommendation regarding the use of the levels of the Common European Framework of Reference for Languages (CEFR, Council of Europe, 2001) and the study of at least two foreign languages in its curriculum. Today, English is practically taught throughout the entire duration of compulsory education, making Italy one of the European Union (EU) countries with the highest number of years of compulsory foreign language teaching (13 years) and the highest number of hours dedicated to foreign language study (European Commission et al., 2017). When learners leave primary school, they are expected to have

reached a full A1 level. At the end of lower-secondary school, where they are offered a choice of a second foreign language besides English, they should have reached the A2 level in English. Different types of upper-secondary schools often offer a second, and sometimes a third, foreign language. The exit level for English at the end of upper-secondary school is expected to be B2 (European Commission et al., 2023). Nevertheless, the long and compulsory period of study of the English language does not seem sufficient to guarantee Italian students the achievement of an adequate level of competence, so much so as to often generate in them a sense of frustration (and deriving demotivation) at the first occasion in which they are confronted with the real language (Santipolo, 2016).

#### **2.4. Research Gap**

Although the CEFR is best known for its statements and reference levels, many questions as to how the usage of Can-Do statements can help learners work towards their learning goals, develop pathways for future study, help with material selection, or achieve any of the other of the CEFR's goals exist since it is not clear if responding to can-do statements can truly provide an estimate of language proficiency, if even a general one (Green, 2012). Even though the CEFR intends to permit learners to measure or estimate language proficiency, neither is the relationship between self-assessment and actual language ability nor is the performance of the CEFR's Can-Do statements as a self-assessment instrument well-enough understood to be able to provide such a measure (Tavakoli & Ghoorchaei, 2009). The need for the current investigation also stems from these concerns, and aims to investigate issues surrounding the general relationship between self-assessment and language proficiency. Moreover, since it has been suggested that incorporating self-assessment into language courses enhances motivation and goal-oriented learning, the present study tested Italian EFL students' ability to self-assess their language skills, in order to emphasize the potential educational role of self-assessment as a vital component of effective learning beyond and within traditional educational settings.

### **3. Methodology**

Participant were 1090 first-year University students enrolled in various non-linguistic Undergraduate degree programmes at a large-sized University in Northern Italy whose English proficiency entry level was assessed using objective tests at the start of the academic year by the University Language Centre. According to University Regulations, all first-year students who between October and December enroll in three-year and single-cycle degree courses at the University of Milan and do not possess a valid official proof of minimum B1 level must register to the Placement tests. Subsequently, students who do not reach a sufficient level of English are enrolled in a course to achieve the level required by the degree or master's degree course. Participants were aged between 19 and 25 years old ( $M = 21.6$ ,  $SD = 0.7$ ), 63.8% of which were female. 95.3% are native

speakers of Italian while students of other nationalities (Albanian, Argentinian, Brazilian, Bulgarian, Cape Verdean, Chinese, Korean, Dominican, Ecuadorean, Egyptian, Eritrean, Philippine, Japanese, Greek, Indian, Iranian, Luxembourgish, Moroccan, Moldavian, Peruvian, Romanian, Russian, Switzerland, Germany, Ukraine, Uzbek) were all below 1% each. 77.3% of the students have studied English for 13 or more years; 23.7% have studied it for between 5 and 13. 61.3% of the students know 3 languages; 19.8% know 4; 14.3% 2 languages; 4.6% know 5 of them. We focused on first-year students since their EFL proficiency was recently tested which enabled us to focus on participants' self-assessed and objectively tested language ability within the same time frame.

The objective measure was Pearson's Versant Placement Test (VEPT), which is intended for adults and students over the age of 16, measures facility in spoken and written English or, more specifically, how well a person can understand spoken and written English and respond appropriately in speaking and writing on everyday topics, at an appropriate pace in intelligible English. The VEPT is divided into eight automatically scored tasks—Read Aloud, Repeats, Sentence Builds, Conversations, Sentence Completion, Dictation, Passage Reconstruction, and Summary & Opinion—which provide multiple, fully independent measures including phonological fluency, sentence construction and comprehension, passive and active vocabulary use, listening skill, pronunciation of rhythmic and segmental units, and appropriateness and accuracy of writing. The scores for each task are determined by more than one task which strengthens score reliability (Figure 1).

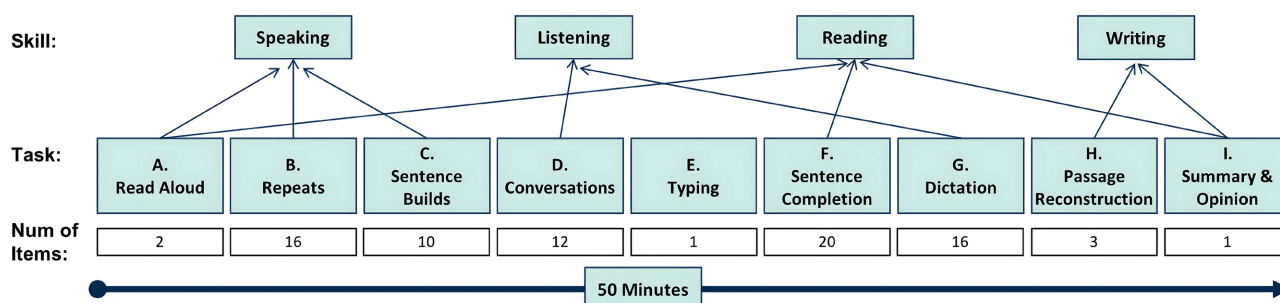


Figure 1. VEPT—Relation of skill scores to tasks

The VEPT provides numeric scores and performance levels and the score report consists of an overall score and four skill scores (speaking, listening, reading, and writing). Versant English Placement Test scores were related to the Common European Framework of Reference for Languages. Table 1 summarises the recommendations made by the expert panelists on the general relationship between scores on the VEPT and the CEFR levels.

For the self-assessment tool, we chose the table of descriptors for self-assessment from the 2020 Italian version of the Companion Volume (see Council of Europe, 2020: pp. 188-190) to which a <A1 level was added. The reason behind such a decision is two-folded. First, because this lower than A1 skill level is already re-

ferred to at the beginning of section 3.5 of the 2001 CEFR (Council of Europe, 2018: p. 32). According to the 2018 Edition of the Volume, the “Pre-A1 represents a “milestone” halfway towards Level A1, a band of proficiency at which the learner has not yet acquired a generative capacity, but relies upon a repertoire of words and formulaic expressions” (Council of Europe, 2018: p. 46). The need for this level below A1 is “important for users as evidenced by the number of descriptor projects which focused on these lower levels” (ibid.). The second reason for including a <A1 descriptor to our self-evaluation scale regards the presence of a Pre-A1 proficiency level within the Versant Placement test that the Language Centre makes use of.

**Table 1.** Recommended mapping of CEFR levels with VEPT overall scores.

VEPT	CEFR
20 - 80	<A1 - C2
20 - 23	<A1
24 - 33	A1
34 - 45	A2
46 - 56	B1
57 - 67	B2

For the sharing of the questionnaire and the data collection we used Google Forms, a free tool that can be accessed from any device. The questionnaire was written in Italian and it consisted of two main parts. The first part included items aimed at getting background information about the students, including name, age, gender and their native country. The second part was designed to elicit information about the students’ assessment of their English language ability in listening, speaking, reading, and writing using the CEFR Can-Do descriptors. Students were asked to choose the descriptors they perceived as best matching their abilities and no reference to the actual CEFR levels was made.

#### 4. Results

To conduct our statistical analysis, we assigned values to the different levels of the CEFR scale: a value of 1 to level Pre-A1, a value of 2 to level A1, and so on, up to a value of 7 (C2). We should point out that this communicative competence scale is not based on equidistant intervals; rather, the continuum is divided into unequal tranches (hence, level A1 covers less distance on the proficiency continuum than level A2). In other words, the more proficient the individual becomes, the higher up the scale they progress and the greater the intervals between levels (Savignon, 1983). However, for the purposes of our investigation, we assigned all the levels the same proportion of the continuum, thus enabling the relevant statistical analyses to be conducted.

To analyse the data, descriptive statistics and inferential statistics (Pearson correlation) were used. This is an exploratory study whose purpose was to ex-

amine whether there were correlations between the scores obtained by students in their self-assessment evaluation of their own proficiency in English as a foreign language, and those obtained through the use of objective tests delivered by the University Language Centre. To this end, Pearson's Product-Moment Correlation was used. We selected the type of study on the basis of the literature dealing with foreign language proficiency assessment, in which correlational studies are the most prolific (Boud & Falchikov, 2007; Oscarsson, 1989; Blanche & Merino, 1989; Liu & Brantmeier, 2019).

The descriptive statistics are displayed in **Table 2**.

**Table 2.** The descriptive statistics of the placement test assessment and students' self-assessments.

	Mean	SD
<b>PT Speaking</b>	3.53	1.196
<b>SA Speaking</b>	4.11	1.173
<b>PT Listening</b>	4.14	1.193
<b>SA Listening</b>	4.47	1.362
<b>PT Writing</b>	4.36	1.085
<b>SA Writing</b>	4.30	1.171
<b>PT Reading</b>	3.87	0.949

At first glance, the above mean values of students' self-assessment of their ability in each skill appear higher than the means of the scores from the Placement test. Moreover, the highest mean difference between self-assessed and objective abilities was found for Reading and Speaking.

The first set of analyses was done to find the degree of go-togetherness of the students' self-assessments and PT-assessments for each of the four skills. The effect sizes of the correlations between students' self-assessments and test assessment are presented in **Table 3**. The value of the effect size of Pearson  $r$  correlation varies between  $-1$  (a perfect negative correlation) to  $+1$  (a perfect positive correlation). The effect size is low if the value of  $r$  varies around  $0.1$ , medium if  $r$  varies around  $0.3$  and large if  $r$  varies more than  $0.5$ .

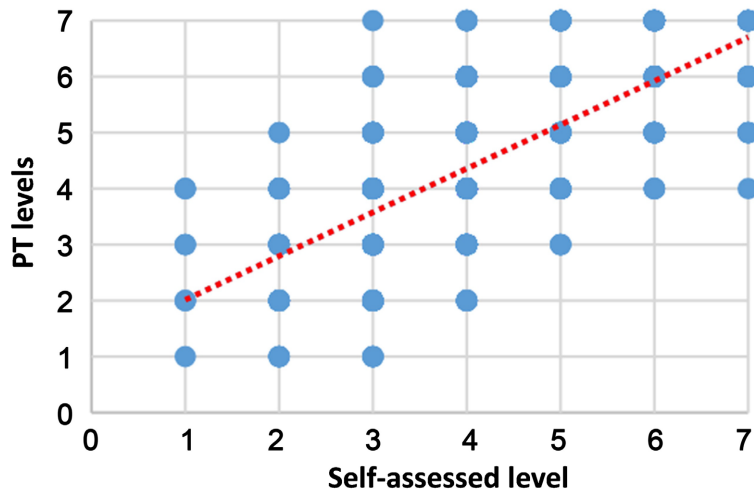
**Table 3.** Correlations effect size.

<b>S PT</b>	$r = 0.55^{**}, p = 0.000$
<b>S SE</b>	
<b>L PT</b>	$r = 0.68, p = 0.000$
<b>L SE</b>	
<b>R PT</b>	$r = 0.50^{**}, p = 0.000$
<b>R SE</b>	
<b>W PT</b>	$r = 0.58, p = 0.000$
<b>W SE</b>	

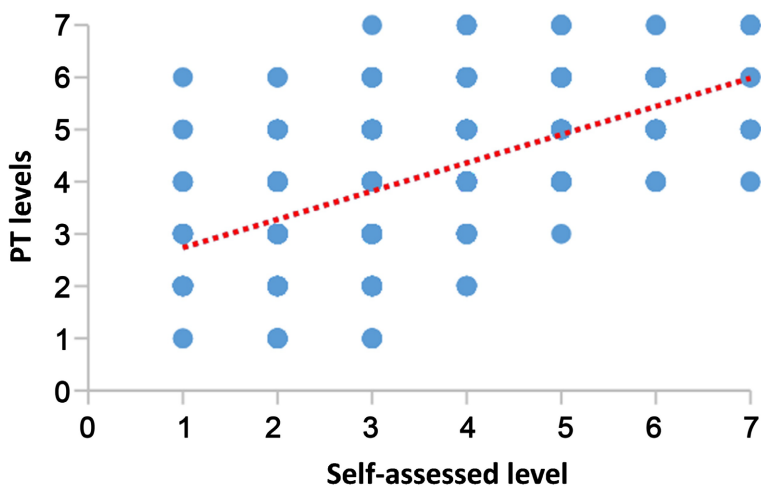
As the means of raw scores in **Table 1** suggested, the results of the 4 correlation analyses we conducted showed that students' self-assessments and Placement test assessment were statistically significant, positive and high for all four skills abilities (0.55 for Speaking, 0.68 for Listening, 0.50 for Reading, and 0.58 for Writing, see **Table 3**).

Furthermore, according to the coefficient  $r$ , students assessed their Listening proficiency in a more accurate way than all the other skills and estimated their Writing skills more accurately than the Reading and Speaking skills. Finally, participants were less accurate when assessing their Reading skills. A graphical representation of the 4 trends is illustrated in **Figures 2-5**.

Moreover, in order to determine whether the type of skill had an effect on students' self-assessment accuracy, we categorised the participants as being over estimators, under estimators or accurate estimators in order to address this final question. The groups were obtained by calculating the difference between students' and PT assessment for each of the four skills.



**Figure 2.** The correlation between speaking assessments.



**Figure 3.** The correlation between listening assessments.

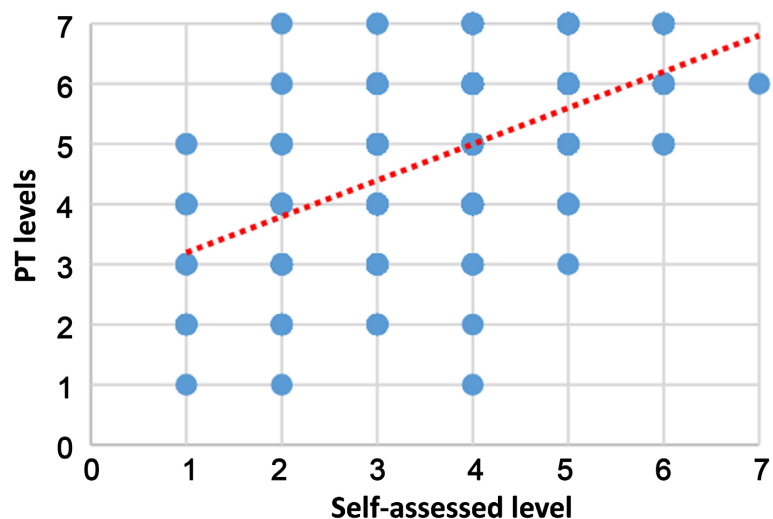


Figure 4. The correlation between reading assessments.

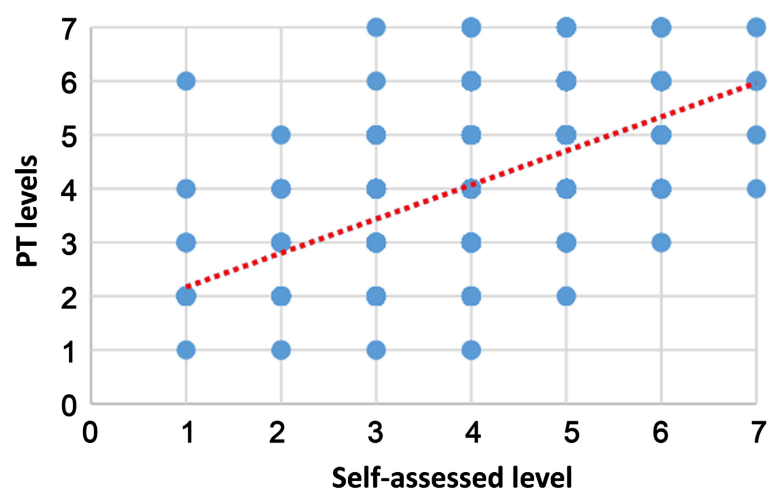
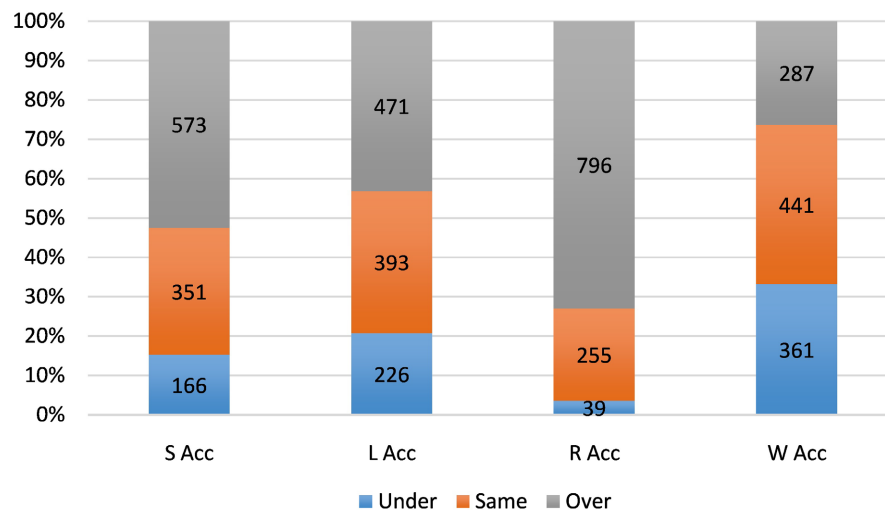


Figure 5. The correlation between writing assessments.

The next step of the analysis involved the estimation of each of the three categories of students between the four skills and within each of them (Table 3). Comparisons illustrated in Figure 6, would seem to suggest that students tend to underestimate their writing skills more than the other skills, while the highest percentage of students who accurately assessed themselves was registered for the same skill. Writing was also the less overrated language ability. Moving to the Reading skill, the percentage of students that underestimated their proficiency in the skill was the lowest among all. Furthermore, the number of participants that rated their Reading skill accurately was again rather low compared to accurate assessments of the other three skills. Quite the opposite, the percentage of students that overestimated their reading ability was the highest among all ratings. As for the Listening skill, the percentage of under estimators is second highest after those of the Writing skill, and the percentage of accurate students seems to follow the same trend. The number of students who overrated the Reading skill

is again second but this time after those that followed the same trend in the Speaking skill assessment. Finally, the percentage of students who underestimated their Speaking skills was second lowest after those that behaved in a similar way when assessing their Reading skill. Accurate estimators were moderate among the other skills while the percentage of participants who overestimated their Speaking skill was second highest after the Reading skill.



**Figure 6.** Assessor type across and within the four ratings.

## 5. Conclusion and Recommendations

Building on past research, this study investigated the extent to which students could accurately evaluate their English proficiency through self-assessment measures.

One of the difficulties related to data analysis in correlational studies is the heterogeneous results found in many of these studies, ranging from very low to very high correlation coefficients, due to a large variety and quality of the instruments used, sample sizes or linguistic dimensions assessed which should always be taken into consideration before data analysis. In this study, significant correlation with a high coefficient ( $r = 0.55, 0.68, 0.50, 0.58$  and  $p = 0.000$ ) was found between the ratings from the self-assessment and the results of the Placement test for each of the four language skills, namely Speaking, Listening, Reading and Writing, which is in line with the findings of other self-assessment correlational studies (Bachman & Palmer, 1989; Blanche & Merino, 1989; LeBlanc & Painchaud, 1985; Oscarsson, 1989; Ross, 1998; Stefani, 1994; Wilson, 1999). In these studies, the correlation coefficients range between  $r = 0.50$  and  $0.60$ , thus answering the primary question posed by the present research.

We can therefore affirm that the self-assessments produced by students participating in our study were accurate. One reason for this high correlation may be the self-assessment tool used in the study, namely the CEFR descriptors. Brantmeier et al. (2012) and Ross (1998) suggest that a high level of specificity

improves the accuracy of self-assessment, as items expressed as “can do” (such as those in the CEFR scale) are more accurate than other possible formats. Equally, when the statements in the self-assessment are specific rather than generalised, and the scale is absolute, the results tend to be more realistic (Ackerman et al., 2002, cited in Sundstroem, 2005). This study provides further evidence of the reliability and accuracy of *Can-Do* descriptors in EFL self-assessment. Some of the benefits associated with Can-do statements, as the CEFR (2001) scale, are that they are positive, concrete, clear and brief.

Self-assessment is currently assuming a larger role in language teaching and learning. The procedure involves students in judging their own learning, particularly their achievements and learning outcomes. Many have argued that teachers should help students construct knowledge through active involvement in assessing their own learning performance, and that students are empowered by gaining ownership of their learning and life-long learning skills. Research on language pedagogy especially recommends that teachers should provide opportunities for students to assess their language level so as to help them focus on their own learning (Dickinson, 1987).

Furthermore, one of the implicit aims of higher education is to enable students to become better judges of their own work. In other words, education has the purpose to develop the capacity for students to make judgments about their own work (Boud & Falchikov, 2007). Such self-evaluation is needed both to enable effective study, so that students can focus on the most important aspects of their work they need to improve, and to build the skills that they will need in any area of work following graduation. If the graduates are not able to make their own judgements about the quality of their work or knowledge, they will be ill equipped for most professional or even non-professional roles. The development of the capacity to make self-judgements about performance tends to be an assumed outcome of higher education. That is, it is taken to be part of any course without the need for specific practice. This is possibly an act of faith, as it is rarely evident in curricula through learning activities or assessment processes (O'Donovan et al., 2008). In contrast to this, research on student self-assessment has suggested that explicit opportunities need to be included for the skill of self-assessing to be developed. We believe that building the capacity to make judgments needs to be an overt part of any curriculum and one that needs to be fostered (Boud & Falchikov, 2007).

If the above is the case then the following questions arise: How might such capacity for judgement be encouraged? And, does engagement in making such judgements over time improve capacity for doing so? If it is true that self-assessment can have a place in placement procedures, many unanswered questions still remain. Further research is needed on self-rating of language ability in the four skill areas. To best implement systematic self-assessment across a program of study, students, instructors, and program directors need to understand the appropriate purposes of self-assessment (for lower-stakes decision making); how to

approach and implement self-assessment (with training); how to document, store, and map self-assessment scores over time (to monitor achievement and growth); and how to continuously improve upon the process through continued research and sharing.

### Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

### References

- Ackerman, P. L., Beier, M. E., & Bowen, K. R. (2002). What We Really Know about Our Abilities and Our Knowledge. *Personality and Individual Differences*, *33*, 587-605. [https://doi.org/10.1016/s0191-8869\(01\)00174-x](https://doi.org/10.1016/s0191-8869(01)00174-x)
- Alderson, J. C. (2007). The CEFR and the Need for More Research. *The Modern Language Journal*, *91*, 659-663. [https://doi.org/10.1111/j.1540-4781.2007.00627\\_4.x](https://doi.org/10.1111/j.1540-4781.2007.00627_4.x)
- Bachman, L. F., & Palmer, A. S. (1989). The Construct Validation of Self-Ratings of Communicative Language Ability. *Language Testing*, *6*, 14-29. <https://doi.org/10.1177/026553228900600104>
- Bailey, K. M. (1996). Working for Washback: A Review of the Washback Concept in Language Testing. *Language Testing*, *13*, 257-279. <https://doi.org/10.1177/026553229601300303>
- Bandura, A. (2006). Guide for Constructing Self-Efficacy Scales. In F. Pajares, & T. C. Urdan (Eds.), *Self-Efficacy Beliefs of Adolescents* (pp. 307-337). IAP Information Age Publishing.
- Blanche, P., & Merino, B. J. (1989). Self-Assessment of Foreign-Language Skills: Implications for Teachers and Researchers. *Language Learning*, *39*, 313-338. <https://doi.org/10.1111/j.1467-1770.1989.tb00595.x>
- Boud, D., & Falchikov, N. (2007). Developing Assessment for Informing Judgement. In D. Boud, & N. Falchikov (Eds.), *Rethinking Assessment for Higher Education: Learning for the Longer Term* (pp. 181-197). Taylor & Francis.
- Brantmeier, C., Vanderplank, R., & Strube, M. (2012). What about Me? Individual Self-Assessment by Skill and Level of Language Instruction. *System*, *40*, 144-160. <https://doi.org/10.1016/j.system.2012.01.003>
- Butler, Y. G. (2016). Self-Assessment of and for Young Learners' Foreign Language Learning. In M. Nikolov (Ed.), *Assessing Young Learners of English: Global and Local Perspectives* (pp. 291-315). Springer International.
- Butler, Y., & Lee, J. (2010). The Effects of Self-Assessment among Young Learners of English. *Language Testing*, *27*, 5-31. <https://doi.org/10.1177/0265532209346370>
- Chen, Y. M. (2008). Learning to Self-Assess Oral Performance in English: A Longitudinal Case Study. *Language Teaching Research*, *12*, 235-262. <https://doi.org/10.1177/1362168807086293>
- Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. <https://rm.coe.int/1680459f97>
- Council of Europe (2018). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. <https://rm.coe.int/cefr-companion-volume-with-new-descriptors-2018/1680787989>

- Council of Europe (2020). *Quadro comune europeo di riferimento per le lingue: Apprendimento, insegnamento, valutazione*.  
<https://rm.coe.int/quadro-comune-europeo-di-riferimento-per-le-lingue-apprendimento-inseg/1680a52d52>
- Dickinson, L. (1987). *Self-Instruction in Language Learning*. Cambridge University Press.
- European Commission, EACEA, & Eurydice (2017). *Key Data on Teaching Languages at School in Europe—2017 Edition*. Publications Office of the European Union.
- European Commission, EACEA, & Eurydice (2023). *Key Data on Teaching Languages at School in Europe—2023 Edition*. Eurydice Report, Publications Office of the European Union.
- Green, A. (2012). *Language Functions Revisited: Theoretical and Empirical Bases for Language Construct Definition across the Ability Range*. Cambridge University Press.
- Falchikov, N., & Boud, D. (1989). Student Self-Assessment in Higher Education: A Meta-Analysis. *Review of Educational Research*, 59, 395-430.  
<https://doi.org/10.3102/00346543059004395>
- LeBlanc, R., Painchaud, G., & Painchaud, G. (1985). Self-Assessment as a Second Language Placement Instrument. *TESOL Quarterly*, 19, 673-687.  
<https://doi.org/10.2307/3586670>
- Little, D. (2005). The Common European Framework and the European Language Portfolio: Involving Learners and Their Judgements in the Assessment Process. *Language Testing*, 22, 321-336. <https://doi.org/10.1191/0265532205lt311oa>
- Liu, H., & Brantmeier, C. (2019). “I Know English”: Self-Assessment of Foreign Language Reading and Writing Abilities among Young Chinese Learners of English. *System*, 80, 60-72. <https://doi.org/10.1016/j.system.2018.10.013>
- Luoma, S., & Tarnanen, M. (2003). Creating a Self-Rating Instrument for Second Language Writing: From Idea to Implementation. *Language Testing*, 20, 440-465.  
<https://doi.org/10.1191/0265532203lt267oa>
- Malabonga, V., Kenyon, D. M., & Carpenter, H. (2005). Self-Assessment, Preparation and Response Time on a Computerized Oral Proficiency Test. *Language Testing*, 22, 59-92.  
<https://doi.org/10.1191/0265532205lt297oa>
- North, B. (2011). Describing Language Levels. In B. O’Sullivan (Ed.), *Language Testing: Theories and Practices* (pp. 33-59). Palgrave Macmillan.
- North, B., Ortega, A., & Sheehan, S. (2010). *A Core Inventory for General English*. British Council/EAQUALS.  
[https://www.eaquals.org/wp-content/uploads/EAQUALS\\_British\\_Council\\_Core\\_Curriculum\\_April2011.pdf](https://www.eaquals.org/wp-content/uploads/EAQUALS_British_Council_Core_Curriculum_April2011.pdf)
- O’Donovan, B., Price, M., & Rust, C. (2008). Developing Student Understanding of Assessment Standards: A Nested Hierarchy of Approaches. *Teaching in Higher Education*, 13, 205-217. <https://doi.org/10.1080/13562510801923344>
- Oscarson, M. (1997). Self-Assessment of Foreign and Second Language Proficiency. In C. Clapham, & D. Corson (Eds.), *Encyclopedia of Language and Education*, vol. 7: *Language Testing and Assessment* (pp. 175-187). Kluwer Academic Publishers.
- Oscarsson, M. (1978). *Approaches to Self-Assessment in Foreign Language Learning*. Pergamon
- Oscarsson, M. (1989). Self-Assessment of Language Proficiency: Rationale and Implications. *RELC Journal*, 19, 75-93.
- Ross, S. (1998). Self-Assessment in Second Language Testing: A Meta-Analysis and Analysis of Experiential Factors. *Language Testing*, 15, 1-20.

- Santipolo, M. (2016), L'inglese nella scuola italiana. La questione negata della varietà modello. In C. A. Melero Rodríguez (Eds.), *Le lingue in Italia, le lingue in Europa: Dove siamo, dove andiamo* (pp. 177-191), Edizioni Ca' Foscari Digital Publishing.
- Savignon, S. J. (1983). *Communicative Competence: Theory and Classroom Practice*. Addison-Wesley Publishing Company.
- Stansfield, C. W., Gao, J., & Rivers, W. P. (2010). A Concurrent Validity Study of Self-Assessments and the Federal Interagency Language Roundtable oral Proficiency Interview. *Russian Language Journal*, 60, 301-317.
- Stefani, L. A. J. (1994). Peer, Self and Tutor Assessment: Relative Reliabilities. *Studies in Higher Education*, 19, 69-75. <https://doi.org/10.1080/03075079412331382153>
- Sundstroem, A. (2005). Self-Assessment of Knowledge and Abilities: A Literature Study. *EM*, 54, 1-36.
- Tavakoli, M., & Ghoorchaei, B. (2009). On the Relationship between Risk-Taking and Self-Assessment of Speaking Ability: A Case of Freshman EFL Students. *The Journal of Asia TEFL*, 6, 1-27.
- Tigchelaar, M. (2018). Exploring the Relationship between Self-Assessment and OPIC Ratings of Oral Proficiency in French. In P. Winke, & S. M. Gass (Eds.), *Foreign Language Proficiency in Higher Education* (pp. 153-173). Springer. [https://doi.org/10.1007/978-3-030-01006-5\\_9](https://doi.org/10.1007/978-3-030-01006-5_9)
- Tigchelaar, M., Bowles, R. P., Winke, P., & Gass, S. (2017). Assessing the Validity of ACTFL Can-Do Statements for Spoken Proficiency: A Rasch Analysis. *Foreign Language Annals*, 50, 584-600. <https://doi.org/10.1111/flan.12286>
- Trim, J. L. M. (2007). *Modern Languages in the Council of Europe 1954-1997: International Co-Operation in Support of Lifelong Language Learning for Effective Communication, Mutual Cultural Enrichment and Democratic Citizenship in Europe*. Council of Europe.
- Upshur, J. (1975). Objective Evaluation of Oral Proficiency in the ESOL Classroom. In L. Palmer, & B. Spolsky (Eds.), *Papers on Language Testing 1967-1974* (pp. 53-65). TESOL.
- Weeden, P., Winter, J., & Broadfoot, P. (2002). *Assessment: What's in It for Schools?* RoutledgeFalmer.
- Wesche, M., Paribakht, T. S., & Ready, D. (1996). A Comparative study of Four Placement Instruments. In M. Milanovic, & N. Seville (Eds.), *Performance Testing, Cognition and Assessment* (pp. 199-210). Cambridge University Press.
- Westhoff, G. (2007). Challenges and opportunities of the CEFR for Reimagining Foreign Language Pedagogy. *The Modern Language Journal*, 91, 676-679. [https://doi.org/10.1111/j.1540-4781.2007.00627\\_9.x](https://doi.org/10.1111/j.1540-4781.2007.00627_9.x)
- Wilson, K. M. (1999). Validity of Global Self-Ratings of ESL Proficiency Based on an FSI/ILR-Referenced Scale. *ETS Research Report Series*, 1999, i-64. <https://doi.org/10.1002/j.2333-8504.1999.tb01811.x>