Self-Assessment in English Language Teaching and Learning in the Current Decade (2010-2020): A Systematic Review

Mona Hosseini[1], Vahid Nimehchisalem[2*]

[1]Faculty of Education, Monash University, Melbourne, Australia
[2]Department of English, Faculty of Modern Languages and Communication, Universiti Putra Malaysia, Serdang, Malaysia
Email: mona.hosseini@monash.edu, *vahid@upm.edu.my

How to cite this paper: Hosseini, M., & Nimehchisalem, V. (2021). Self-Assessment in English Language Teaching and Learning in the Current Decade (2010-2020): A Systematic Review. *Open Journal of Modern Linguistics, 11,* 854-872.
https://doi.org/10.4236/ojml.2021.116066

Received: October 31, 2021
Accepted: November 27, 2021
Published: November 30, 2021

Abstract

Self-assessment (SA), defined as the evaluation of process and product and the main feature of self-regulated learning, has been the focus of many recent studies. Due to the importance of SA, the purpose of this systematic review is to provide a comprehensive overview of relevant studies regarding the use of SA for teaching and learning English language. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) was employed to collect studies on SA. The search for articles was done with the combination of these keywords: "self-regulation", "self-assessment", "foreign language learning", "student", and "teachers" in EBSCOhost research platform. We found 106 studies related to SA published in period of 2010-2020. After screening the title and abstracts for selecting articles, based on the inclusion and exclusion criteria, 18 articles were selected to be reviewed. These articles were closely related to the focus of this systematic review on the reliability and validity, effectiveness, accuracy as well as the modes of application of SA in English language learning-teaching. The paper starts with reviewing the definitions of SA and further it analyzes and synthesizes the selected studies to reply to the research questions. The paper concludes with the pedagogical implications and recommendations for future studies on various aspects of SA in the field of English language learning and teaching.

Keywords

Purposes of Self-Assessment, Language Assessment, Review, Self-Assessment, Teaching and Learning English

1. Introduction

Educators today are encouraged to develop their learners' autonomy by moti-

vating them to self-regulate and self-invest in their learning experience. For learning to take place more effectively, learners should be involved in the process of their own learning. Proponents of constructivism move much of the responsibility from the teacher to the learner; they no longer see the teacher the central figure who spoon-feeds students in the classroom. In line with this trend, the current paradigm, referred to as Education 4.0, promotes notions like Flipped classroom approach, personalized learning, hands-on learning, and independent learning (Aziz Hussin, 2018). In response, in the area of assessment, alternative methods of assessment are recommended in order to boost learners' higher order thinking skills. More learner-oriented methods of assessment like collaborative assessment, peer assessment, project-based assessment, and individual assessment have emerged. The teacher is expected to be engaged and keep the learner involved in process of assessment in an on-going manner throughout the learning experience (Farhady, 2021). Self-assessment (SA), our main focus in this review article, is one of these learner-oriented assessment methods that have attracted an increasingly large number of practitioners in recent years.

According to Xi and Davis (2016), it was in 1961 that SA appeared in the area of second language acquisition (SLA) area with the publication of Lado's *Language Testing* (Lado, 1961). Since initially language was regarded as a set of discrete elements, language assessment was done structurally, however, later psycholinguistic and sociolinguistic approaches created integrative global measures (Xi & Davis, 2016). With the advent of constructivism, the focus shifted from the summative assessment and learning product to formative assessment and learning process. According to constructivists, knowledge is not gained but constructed by learners, therefore, the teacher is expected to engage learners in self-investing in their learning and to train them to assess their own learning process in addition to the learning product.

In 1976, psychometric testing gave place to educational assessment and testing with the purpose of learning became one of the major goals to pursue in education (Lambert & Lines, 2000). Since then, alternative assessment has been applied as a means to make learning more meaningful in the classroom. Whereas assessment was primarily applied to judge students' learning, focus from assessment of learning (AoL) later shifted to assessment for learning (AfL), whose purpose is enhancing and promoting learning (Lee, 2017; William, 2001 as cited in Lee & Coniam, 2013). AfL helps the teacher explore what the learners have learned and what they have not learned yet. AoL, on the other hand, is the assessment "used to give grades or to satisfy the accountability demands of an external authority" (Shepard, 2000: p. 4).

Among the various modes of alternative assessment, self- and peer-assessment have attracted more attention in recent years since they seem to influence learners' independence and autonomy (Sambell et al., 2006). SA gives learners the opportunity to focus on their learning, manage their progress, and find ways to change, adapt or improve it (Kavaliauskienė, 2004). Some of the purposes of engaging students in SA are to enhance their learning and realization, to help their

academic self-regulation, and to monitor and manage their own learning (Zimmerman & Schunk, 2004). However, it can be quite tough for foreign language learners to self-assess their learning due to the lack of exposure to the target language (Hung, 2019). Although SA is generally believed to have positive effects on students' learning, learners' inaccurate perceptions about their own work and capabilities may create serious consequences such as underestimating their real achievement, feeling of incompetency, and skipping courses (Harris & Brown, 2018).

Some researchers view SA as merely a quantitative evaluation of one's own performance, by counting the number of correct answers (Andrade, 2009; Panadero, 2011). Others consider it a qualitative and efficient way of learning and evaluation (Harris & Brown, 2013). Oscarson (1989) divides SA into two types of performance-oriented and development-oriented types. Performance-oriented SA is summative in nature with a focus on the learners' grade and achievement. Development-oriented SA, on the other hand, is a type of formative assessment in which the focus is on the learner's progress. It encourages the students to reflect on the quality of their work and learning, make a judgment of the degree to which they reflect the specified goal or criteria, recognize strengths and weaknesses in their work, and review their work (Andrade et al., 2008). Moreover, development-oriented SA provides learners with feed-foreword rather than feedback. While feedback helps learners use learning materials in meaningful ways (McGonigal, 2006), it can turn into to feed-forward if the assessment is connected to the feedback comments and provides learners with the information that improves their future learning (Irons, 2004).

The distinction between the two types of SA is also evident from the various definitions provided by the scholars. **Table 1** summarizes some of the definitions of SA. The position of each definition is based on the dichotomy of "performance-oriented" and "development-oriented" SA defined by Oscarson (1989).

While SA is defined as performance-oriented and a testing instrument by some scholars (Bailey, 1998; Panadero, 2011; Rust et al., 2003), and as development-oriented and learning material by Brown (2004), others view it as both performance- and development-oriented which could be used as testing instrument as well as learning material (Benson, 2011; Panadero et al., 2016; Brown & Harris, 2013). These three categories of definitions show that beside the dichotomy of being purely a testing instrument or learning material, SA can be included in a third category which fulfills these 2 purposes simultaneously. This could be shown more clearly as in **Table 2**.

Table 2 proposes a hybrid model of SA which rather than defining SA as *either* a testing instrument *or* a learning material, views it as a tool which can simultaneously be used *both* for learning-teaching *and* for testing. Such a hybrid model gives more freedom to the learner and teacher to integrate SA throughout the language learning process rather than limiting it only to a particular point in the process.

Table 1. Definitions of SA.

Reference	Definition	Position
Bailey (1998: p. 227)	"Procedures by which learners themselves evaluate their language skills and knowledge".	SA as a testing instrument
Rust et al. (2003)	The learners' judgment of their work: it covers terms such as "self-appraisal" and "self-evaluation".	
Panadero (2011: p. 78)	"Self-assessment is a descriptive and evaluative act carried out by the student concerning his or her own work and academic abilities".	
Brown (2004)	A representation of formative assessment that involves learners in learning process by focusing on their performance.	SA as a learning material
Benson (2011)	"The qualitative assessment of the learning process and of its final product, realised on the basis of pre-established criteria".	SA as a learning-teaching-testing tool
Harris and Brown (2013: p. 368)	"SA is the ability to assess one's knowledge, learning and performance, a key element in becoming an autonomous learner".	
Panadero et al. (2016: p. 804)	"Wide variety of mechanisms and techniques through which students describe (i.e., assess) and possibly assign merit or worth to (i.e., evaluate) the qualities of their own learning processes and products".	

Table 2. A hybrid model of SA.

SA as a		
learning material	**learning-teaching-testing tool**	**testing instrument**
formative	AND	summative
descriptive		judgmental
feed-forward		grade/feedback
process-oriented		product-oriented
autonomous learning		accountable teaching
development-oriented		performance-oriented
learning/learner-centered		exam/teacher-centered
continuous/on-going assessment		achievement examination
assessment for/as learning (AfL/AaL)		assessment of learning (AoL)
focus on quality of learning in progress		focus on reliability and validity

Education in all stages is moving towards more learner-centered and SA is among the main learner-centered practices which would potentially be beneficial in testing as well as learning processes. However rich in theory, SA has not been widely used in educational contexts in practice. Either as a learning material or testing instrument, SA practices need to be investigated to find out the reasons why they are not extensively applied in practice, and what potential factors would enhance the status of SA as a favorable and reliable tool as well as material in educational contexts in all stages. The purpose of this systematic review is to provide an overview of the current literature on the SA practices in English language teaching and learning. As the number of studies on SA in the domain of English teaching and learning is relatively limited, this review focuses some as-

pects of SA and makes an attempt to answer the following research questions:

1) What issues are involved in using SA as a testing instrument?

2) What makes SA an effective material for improved learning and testing instrument?

2. Methodology

In conducting this systematic review on SA, we applied PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) which is the standardized protocol for conducting systematic reviews as well as delivering vivid reports (Liberati et al., 2009). Following the PRISMA guidelines, checklist and flow diagram, a systematic search strategy with pre-selected search terms and eligibility criteria was applied and EBSCOhost research platform was searched for peer-reviewed articles on SA. In total, 106 articles on SA, published in period of 2010-2021, were retained from EBSCOhost research platform. The search was done through the combinations of different search terms, namely: "self-assessment", "foreign language learning", "second language learning", "student", "teachers", "evaluation", "assessment", "autonomy", and "self-regulation". In the next stage, after screening the title and abstract of the articles based on inclusion and exclusion criteria, articles were selected to be examined for their eligibility. We finally selected 18 articles to be reviewed. These articles were specifically related to the area of SA in English language teaching (ELT), which is the main focus of this paper. To select the articles, the following inclusion criteria were applied: 1) original peer-reviewed studies published in English, 2) studies published within a time span of a decade (2010-2020), 3) the application of SA in EFL/ESL settings. We rarely found articles on SA published earlier than this time span which could fit ESL/EFL setting which is the focus of our study. The process for selecting the articles is provided through PRISMA flow diagram below (**Figure 1**).

3. Results of the Review

We read the 18 selected articles, summarized them to answer 2 research questions. **Appendix 1** presents an overview of the selected studies, and it includes information on 1) authors' name, year of publication, and context; 2) methods applied to conduct the study, participants, and instruments; 3) findings of the studies. The results of analyzing and synthesizing the reviews are presented in 2 sections according to the 2 research questions.

1) What issues are involved in using SA as a testing instrument?

Educational practices have faced many changes and "modern democratic, collaborative and socioculturally oriented teaching strategies" demand learners' active involvement in monitoring and evaluation of their learning which leads to a greater sensitivity to their strength and weaknesses and gives them the chance of reaching their pre-planned goals (Oscarson, 2013: p. 2). Some of the issues involved in using SA as testing instrument investigated in the reviewed articles are the *reliability*, *validity*, and *accuracy* of SA as a testing instrument as well as

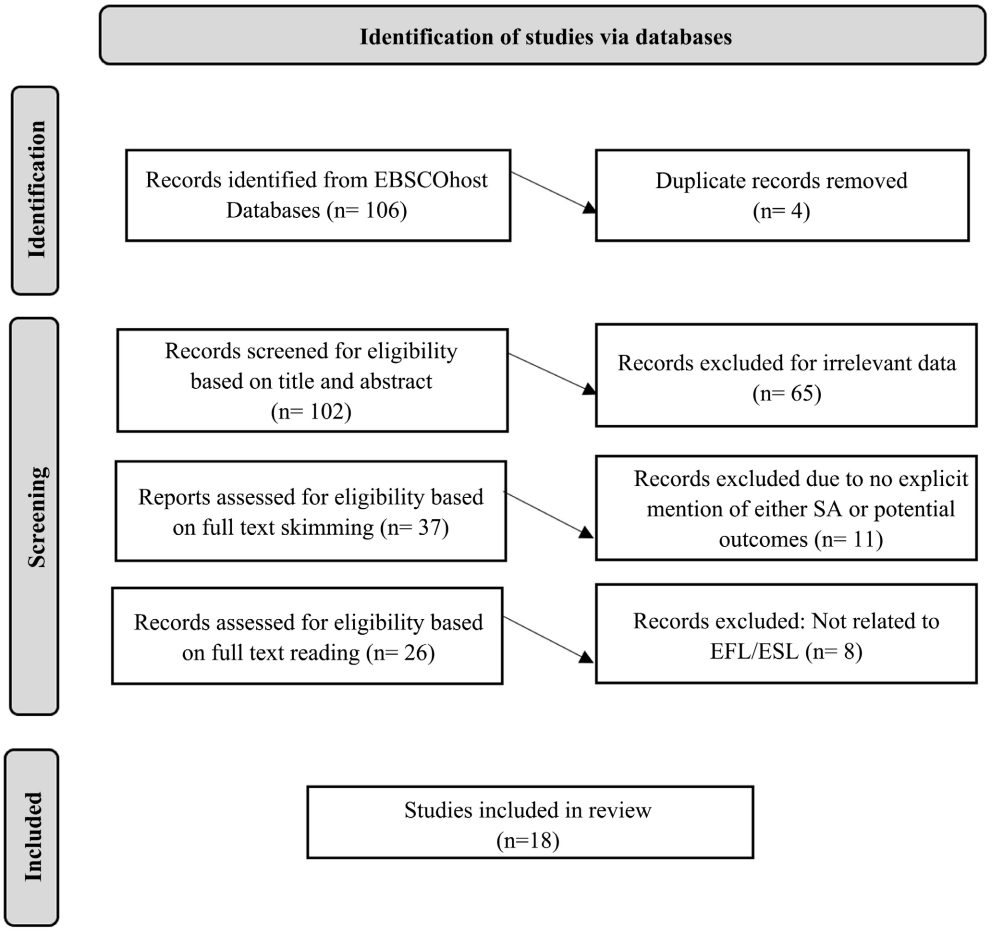

Figure 1. PRISMA flow diagram for article selection.

the *proficiency levels* of the participants who make judgements about themselves.

Reliability, Validity, and accuracy

Many studies on SA have merely focused on its reliability, validity, and accuracy as an instrument for testing and evaluation. While some studies, such as Tigchelaar (2018), on the reliability and validity of SA have recognized it as a reliable measure of language proficiency, others disapprove the accuracy of SA. One of the reasons for questioning the reliability of the SA results is the learners' underestimation or overestimation of their abilities. Learners may overestimate their abilities in reporting the result of SA to keep face in a social context like class (Kuncel et al., 2005). While Dunning et al. (2004) relate SA inaccuracy to learners' optimism about their abilities and their ignorance of important information, some studies (e.g., Ünaldı, 2016) discuss that foreign language learners' underestimation of their language proficiency makes SA a weaker predicator than teacher assessment. Such underestimations according to Ünaldı (2016) have roots in sociocultural issues, particularly the existence of modesty in many eastern societies. Additionally, in comparison of self- and peer-assessment, Samaie et al. (2016) found participants assigned higher grades to their peers than

themselves, which is an indication of learners' underestimation of their own abilities. Therefore, SA seems to be a good complement to teacher assessment, which could be applied to elevate language learners' motivation and help reduce the drawbacks of formal assessment, rather than an individual testing instrument (Ünaldı, 2016).

SA has also been recognized a reliable placement tool in institutions for student enrolments in Intensive English Programs (Summers et al., 2019). The lower level students, however, over-assessed themselves which it can be considered as a weakness of SA. The authors mentioned several reasons for this problem such as students' proficiency level, their experience, cultural background, or even the level of authenticity of the task. Although SA proved to be reliable, there was a low correlation between placement tests and SA which means SA is not valid enough to be solely applied as an instrument for placement. There are different stances on measuring the accuracy and reliability of SA by comparing it to teacher assessment. Although, the significant correlation between SA and objective tests does not support the accuracy of SA (Ashton, 2014), Brown and Harris (2013) reported a weak correlation between learners' rating through SA and teacher's grading of the students' performance. Studies on the reliability of SA relying on the correlation between the student and teacher grading have resulted in contrasting results which could be related to the differences in learners' age, context, subject material, and the way SA is conducted. For instance, regarding the age of learners, while according to Butler (2018) the older learners showed less variability and had more conservative evaluations compared with their younger counterparts, Liu and Brantmeier's (2019) study on self-rating abilities of writing and reading among young English learners reported a significant correlation between the scores of SA reading and writing abilities and the objective tests of reading comprehension and writing production. Accordingly, they believe that young learners are capable of doing SA accurately and SA can be incorporated as part of the curriculum design in foreign language classrooms. Additionally, Bullock (2011) reported that according to the perceptions of majority of teachers, the learners are quite capable of assessing their own work. In the same vein, Birjandi and Hadidi Tamjid (2012) found adult EFL learners capable of conducting reliable SA of their writings, however, Hamer et al. (2015) concluded that SA done by the students is not reliable since they are not competent in the subject matter. This means that the learners may not be qualified enough to judge their performance.

In terms of the accuracy of SA, Ashton (2014) asserts that the significant correlation between SA and objective tests does not support the accuracy of SA. However, when applied as a testing instrument and learning material, SA can fulfil modern notion of education since it gives learners the opportunity of involvement in monitoring, evaluation, awareness, and last but not the least, enhanced learning. There are studies on the accuracy of SA that show positive outcomes. For example, the study done by Summers et al. (2019) examined the

reliability of SA as a tool for placement purposes in institutions for students' enrolment in Intensive English Programs (IEPs). The data of the study were analyzed using CAL validation framework and the study concluded that the difficulty level of the can-do statement in SA were in the same with those of NCSSFL-ACTFL. In addition, it was revealed that the instrument was highly reliable.

Language Proficiency

Since some studies (e.g., Summers et al., 2019) have found learners' low proficiency level a source of unreliability of SA practices, majority of studies on SA as an evaluation material have focused on teachers rather than learners. Teacher's SA is an influential factor in teachers' occupational growth (Ross & Bruce, 2007). Teachers' proficiency and expertise are important issues in EFL context where there is a growing gap between the non-native teachers' current level of language proficiency and the language proficiency level required in the system (Nakata, 2010). Nakata's study on the possibilities of the Classroom Language Assessment Benchmark (CLAB) as a professional development tool for EFL teachers confirms the efficacy of CLAB as a teacher's professional development tool through surveys among teachers. The study also asserts that CLAB, teacher assessment system in Japan, evaluates teachers' proficiency, raises teachers' awareness of their classroom English and help them to improve their English language proficiency as well.

Formative assessment, in the form of observation for new teachers, enables novice teachers to reflect on their professional repertoire, however they can only be observed limited times by the administrator (Snead & Freiberg, 2017). SA enables teachers to have formative assessments independently and without the presence of administrators (Snead & Freiberg, 2017). They discuss that working with Personal Centered Learning Assessment (PCLA) made the participants capable of getting deep insight into their classroom and teaching in a non-evaluative environment, from their own perspective as well as those of their students. SA made the student teachers aware of their weaknesses and strength in teaching. The feedback through PCLA provided them with chances to improve their teaching abilities. Whereas Nakata's (2010) study focused on the instrument that measures teachers' language proficiency level and helps enhance it, Snead and Freiberg (2017) focused on the instrument that examines teachers' awareness of their teaching and class management and it also provides the opportunity to have access to students' perspectives about the teachers.

Borg and Edmett (2019) state that whereas teacher SA is not a novel concept, teacher's evaluation in the field of ELT is still following the top-down approach. In their study they have evaluated a SA Tool (SAT) that is part of an approach to the professional development of English language teachers called "Teaching for Success" by British Council. The study by Borg and Edmett aimed to find out how English teachers rate their competence on SAT and what their views are about SAT in terms of its value, relevance, and content. Borg and Edmett (2019)

concluded that majority of the teachers who took the test, regarded it as a beneficial exercise. Moreover, the study identified some ways to enhance SAT as a SA tool. The study reported some valuable data that can give insights to the researchers who study on SA and can be guidelines for further studies. For instance, 84.6% of the respondents among a total of 1,684 were female, and 57% of the respondents, from 125 countries, were from European countries.

2) What makes SA an effective material for improved learning and testing instrument?

Assessment for learning, formative assessment, aims to enhance learning by providing the learners with information on their learning development while they are still learning (Dragemark-Oscarson, 2009). The studies on SA as a learning material to improve learning as well as an effective testing instrument, have identified repeated and regular practice of SA, making use of rubrics which help the learners in all stages of self-regulated learning, providing feedback, proper employment of technology for SA, making learners aware of the purposes of SA, goal setting, developing SA tasks which help learners practice the assessment criteria, and the combination of self-assessment with teacher and peer-assessment are among the important considerations.

Hung (2019), who examined the effectiveness of repeated (5 trials) SA on EFL learners' oral performance found that students' English speaking proficiency improved particularly in grammar, vocabulary, and use of linking words. Instructors' as well as students confirmed the remarkable improvement in students' fluency. The constant practice of SA of a time period seems to be an effective factor on students' learning. Similarly, Butler and Lee (2010), who conducted a comprehensive study on SA to investigate elementary school students' ability of self-assessment over a period of time, reported improvements in students' SA accuracy and increased confidence through regular unit-based SA. Regular assessment will motivate students to participate in class activities and increase their efforts for learning (Murakami et al., 2012). Murakami et al. (2012) also indicated that the combination of self-, peer-, and teacher-assessment increases students' willingness to involve in speaking activities in class, lifts their language learning effort, and boosts learners' linguistic self-confidence.

There has been a growing interest in formative rubric use in general education in the last decade (Wang, 2016). Wang's study on the contextual analysis of students' perceptions of rubric use in an EFL learning classroom revealed that applying rubrics in SA helped the learners in all three stages of self-regulated learning, namely, forethought, performance, and reflection, defined by Zimmerman and Moylan (2009). Wang (2016) discusses that students can be given chances and power in the design and modification of the rubric criteria. Flexibility in the rubric structure proved to be an important factor in designing an effective rubric since not all students in Wang's (2016) study were interested in welcoming analytic-structured rubrics. Additionally, consideration of wording and score range is highly important to avoid subjectivity in students' self-judgement.

Regarding rubric users' domain knowledge, Wang (2016) recommends using Hattie and Timperley's (2007) three questions on feedback of *where to go*, *what's up*, *and what's next*? Wang discusses that SA has very limited instructional power unless it is provided with sufficient knowledge. Wang (2016) asserts that, similar to other studies (e.g., Andrade et al., 2008), to get the optimum effectiveness of SA rubrics, the students need 2 to 3 sessions to get to know the rubrics.

Adachi et al. (2016) conducted another study that focused on SA as a learning material. They explored the academics' perceptions towards the benefits and challenges of SA as well as peer assessment in higher education. Through the interviews, they extracted the themes, which introduce the benefits of self and peer-assessment. The benefits are the development of transferable skills, cultivating students who are work-ready, promoting active learning, better understanding of standards and assessment criteria, timely, varied and appropriate feedback for students, skills involved in giving and receiving feedback, and less input (and time) required of teachers. However, their study does not only rely on the benefits of SA, and they introduce the challenges of SA as well. Reliability and accuracy of students' judgment skills, perceived expertise, power relations, and time and resource constraints are among the drawbacks they reported. They discuss the opportunities and challenges of self and peer-assessment in online-environment as well. Although the benefits and challenges are categorized, they overlap in some themes. According to Adachi et al. (2016) the existing challenges of SA which they identified in their study to some extent contradict those previously identified by Liu and Carless (2006).

Duque Micán and Cuesta Medina (2015) investigated the impact of SA of vocabulary competence on young students' oral fluency. Their study concluded that the students found SA strategies effective in their vocabulary learning process and also acknowledged their own difficulties and strengths regarding their language learning process through reflective practice. SA made them go through the process of examining their weak points and strength in learning and therefore, creating personal commitments of practicing which leads to improvement in vocabulary and oral fluency. Furthermore, the study highlighted the role of goal setting as an important aspect since it was found that SA supported students' ability to evaluate and improve the type of objectives set week after week. The study reported that SA had a positive impact on learners' vocabulary, fluency and learning process in general. The study showed through a cycle of SA, students will be aware of their weaknesses and strengths, create personal commitments, and use strategies to improve their learning. This cycle embraces two edges of SA, being a testing instrument to find drawbacks and strength and simultaneously being learning material to help learners improve.

With the same focus on SA, Mazloomi and Khabiri (2016) investigated SA as a development-oriented formative assessment. In their study, SA is applied as a learning task in which the learners practice the assessment criteria, and they de-

velop their writing skills as well. The analysis of the data in their study revealed that the language learners' writing ability enhanced through applying SA. Their result was in line with the findings of the previous studies such as Sullivan and Lindgren (2002) and Javaherbakhsh (2010). Mazloomi and Khabiri (2016) reported a significant effect of SA on students' general language proficiency through practice and training. Moreover, they pointed out that because of SA training and practice, the learners' SA gains a closer correlation with those of the raters and teachers which adds up to the validity of the learners' self-ratings.

Samaie et al. (2016) who investigated the use of technology for SA concluded that using WhatsApp to perform SA of oral proficiency requires a great amount of effort and time and participants preferred face-to-face interactions for the purpose of assessment and receiving feedback. Therefore, the study suggests that applications such as WhatsApp need to be accompanied with some synchronous visual applications or traditional face-to-face interactions to be beneficial for the purpose of assessment. However, it cannot be generalized that the participants were reluctant to SA through applications since other applications may have different effects on the participants. Regarding the comparisons between self and peer-assessment, their study is in line with previous studies done by Chang et al. (2012) and Lin et al. (2001) who have reported that self-assessors are stricter than peer-assessors are. Unlike what Samaie et al. (2016) concluded in their study, Hung (2019) found that technology plays a positive role in students' SA process. The study revealed that learners improved in some criteria in their English abilities through the practice of repeated SA. Vocabulary and grammar were two aspects of speaking abilities, which were not significantly improved through the practices. The study reported a high level of agreement between students' SA, survey responses, and the raters' assessments. These three aspects agreed on the students' improvement in their SA abilities, their language learning and their mutual enhancement, particularly their ability to evaluate linking words, vocabulary, and to an extent, grammar. The comparisons of students' SA reports showed that there was a positive effect on the use of linking words and vocabulary to report the result of self-evaluations. In contrast to other studies on SA, Hung's (2019) is not limited to one aspect. The study addressed extensive practice with proper guidance, combining evaluation with reflection, and the reciprocal benefits of assessment and language achievement, culminating in using Bandura's observational learning as an operational guide for repeated SA practice. Hung concluded that factors such as sufficient practice, recognition and correction of weaknesses, and real learning outcome created a positive impact on this learning experience.

Scholars discuss that it is only through the learners' awareness of the purposes of SA that self-assessment (SA) will be effective. Harris and Brown (2018) point at two main reasons for the necessity of the training of SA in the classroom. The first reason they mention is that SA is an important part of self-regulation. Secondly, they state that the absence of training and guidance in the process of SA

may result in their poor performance which is varied from the students' real abilities and the standards of curriculum. SA is an important instrument for assessing learners' achievement, giving them diagnostic feedback, and creating critical thinking in learners, therefore, it enables them to assess their learning critically, linking it to their previously acquired knowledge, and applying it in planning for their future learning (Kunan & Jang, 2009). Moreover, teachers claim that SA empowers learners with the awareness of their strength and weaknesses (Bullock, 2011). SA in the form of voice recording and practicing speaking on a personal device, such as mobile phones, is reported to provide several repeated speaking practice opportunities for language learners and hence increasing the students' confidence (Hung, 2019).

4. Reflections

SA has been studied from various perspectives, as a testing instrument, learning material, and testing and learning material. In spite of the abundant number of studies done in this regard, our review shows a lack of studies on the type of SA that plays the role of learning materials and satisfies the requirement of being testing instrument as well. Therefore, more investigations need to be done and more innovative checklists need to be developed to sustain student engagement with SA and foster their development of self-regulated learning. Although many studies have been conducted in this area, the issue of SA and technology deserves more studies due to the importance of technology in life. With the presence of computer-assisted language learning, online learning, learning applications, and mobile assisted language learning in the last decade, it seems to be necessary to study SA in these areas as well. According to previous studies done in this regard, there are areas which have received less attention so far such as studying extensive implementation of mobile or computer SA trials to track the long-term effects of the practice and establish broad stages of development of both assessing and language skills, investigating how to promote positive attitudes towards mobile-assisted assessment applications among learners, and finding out ways to maximize the learning effect of online SA tools particularly in online learning settings. Young learners' SA demands more attention since making them aware of the purposes of SA is challenging. It is crucial to study how young learners' language proficiency level influences their SA process. Moreover, it is necessary to explore the possible approaches by making adjustments to SA in ways that result in enhanced learning and being autonomous learners.

There are many studies on the students' perceptions towards SA. However, there are not enough studies which focus on the ways through which any kind of SA including checklists, rubrics, and online SA are introduced and presented properly to the students. Since students' awareness of the purposes of SA is a highly important issue, more studies are needed to be done in this regard. In many countries, grading is still an important issue in education system, therefore, students would have different perceptions towards assessment of learning

(AoL) and assessment for learning (AfL). Future studies may focus on the differences in students' perceptions about AoL compared to AfL. The literature on SA seems to lack studies that focus on the psychological effects of SA tools. For instance, there are not studies done on whether SA reduces students apprehension in writing or speaking skill and to what extent such instruments can help enhance students' confidence in expressing themselves in speaking or submitting their writings to their teachers. Moreover, the role of teachers as well as students' gender in SA, and SA in various contexts such as less developed countries are the issues that can be investigated in future studies.

5. Conclusion

This review article started by giving information on the history of SA and further, it introduced the different definitions of SA by different scholars. A chronological overview of these definitions shows that SA was applied as a testing instrument when it was introduced to educational setting, whereas later it started being treated and used as a learning material. The same pattern can be seen in the 18 articles focusing on SA in this paper. It can be implied that the initial studies on SA mostly focused on the reliability and validity of the issue, however, the more recent studies investigate the other aspects such as various instruments of SA, applying SA through technology, the impact of SA on young learners, and using SA as a learning material rather than a testing instrument. As mentioned earlier and according to the themes and categories in this review, there is a third view towards SA which integrates SA as a testing instrument and SA as a learning material that results in SA as a testing instrument which can also make learners aware of their performance quality as well as a learning material that enables them to plan for better learning. In this third view, SA highlights learners' roles both in learning tasks and evaluation (Oscarson, 2013). This is where feedback alongside with feed-forward paves the way for meaningful and long-lasting learning. In addition, the review of these studies shows that those done in the beginning of the decade mostly relied on the quantitative research methods whereas reaching the end of the decade, we can observe that the studies are conducted using mixed methods approach or qualitative research methods. According to all aspects which have been introduced and elaborated in the paper, it can be concluded that nearly all the studies on SA have reported positive effects of SA on students' learning process with only minor problematic issues. It seems that future studies will lead to more professional and beneficial usages of SA for the purpose of better learning.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

Adachi, C., Tai, J., & Dawson, P. (2016). Enabler or Inhibitor? Educational Technology in Self and Peer Assessment. In S. Barker, S. Dawson, A. Pardo, & C. Colvin (Eds.), *Pro-*

ceeding of Ascilite 2016 Show Me The Learning Adelaide (pp. 11-16).

Andrade, H. L. (2009). Students as the Definite Source of Formative Assessment: Academic Self-Assessment and the Self-Regulation of Learning. In H. L. Andrade, & G. J. Cizek (Eds.), *Handbook of Formative Assessment* (pp. 90-105). Routledge. https://doi.org/10.4324/9780203874851

Andrade, H. L., Du, Y., & Wang, X. (2008). Putting Rubrics to the Test: The Effect of a Model, Criteria Generation, and Rubric-Referenced Self-Assessment on elementary School Students' Writing. *Educational Measurement: Issues and Practice, 27,* 3-13. https://doi.org/10.1111/j.1745-3992.2008.00118.x

Ashton, K. (2014). Using Self-Assessment to Compare Learners' Reading Proficiency in a Multilingual Assessment Framework. *System, 42,* 105-119. https://doi.org/10.1016/j.system.2013.11.006

Aziz Hussin, A. (2018). Education 4.0 Made Simple: Ideas For Teaching. *International Journal of Education and Literacy Studies, 6,* 92-98. https://doi.org/10.7575/aiac.ijels.v.6n.3p.92

Bailey, K. M. (1998). *Learning about Language Assessment: Dilemmas, Decisions, and Directions*. Heinle & Heinle.

Benson, P. (2011). *Teaching and Researching Autonomy* (2nd ed.). Routledge.

Birjandi, P., & Hadidi Tamjid, N. (2012). The Role of Self-, Peer and Teacher Assessment in Promoting Iranian EFL Learners' Writing Performance. *Assessment & Evaluation in Higher Education, 37,* 513-533. https://doi.org/10.1080/02602938.2010.549204

Borg, S., & Edmett, A. (2019). Developing a Self-Assessment Tool for English Language Teachers. *Language Teaching Research, 23,* 656-679. https://doi.org/10.1177/1362168817752543

Brown, H. D. (2004). *Language Assessment: Principles and Classroom Practices*. Longman.

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. https://doi.org/10.4135/9781452218649.n21

Bullock, D. (2011). Learner Self-Assessment: An Investigation into Teachers' Beliefs. *ELT Journal, 65,* 114-127. https://doi.org/10.1093/elt/ccq041

Butler, Y. G. (2018). The Role of Context in Young Learners' Processes for Responding to Self-Assessment Items. *The Modern Language Journal, 102,* 242-261. https://doi.org/10.1111/modl.12459

Butler, Y. G., & Lee, J. (2010). The Effects of Self-Assessment among Young Learners of English. *Language Testing, 27,* 5-31. https://doi.org/10.1177/0265532209346370

Chang, C. C., Tseng, K. H., & Lou, S. J. (2012). A Comparative Analysis of the Consistency and Difference among Teacher-Assessment, Student Self-Assessment and Peer-Assessment in a Web-Based Portfolio Assessment Environment for High School Students. *Computers & Education, 58,* 303-320. https://doi.org/10.1016/j.compedu.2011.08.005

Dragemark-Oscarson, A. (2009). *Self-Assessment of Writing in Learning English as a Foreign Language. A Study at the Upper Secondary School Level*. PhD Dissertation, Göteborg Studies in Educational Sciences 277, Acta Universitatis Gothoburgensis. University of Gothenburg.

Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed Self-Assessment: Implications for Health, Education, and the Workplace. *Psychological Science in the Public Interest, 5,* 69-106. https://doi.org/10.1111/j.1529-1006.2004.00018.x

Duque Micán, A., & Cuesta Medina, L. (2015). Boosting Vocabulary Learning through Self-Assessment in an English Language Teaching Context. *Assessment & Evaluation*

in Higher Education, 42, 398-414. https://doi.org/10.1080/02602938.2015.1118433

Farhady, H. (2021). Learning-Oriented Assessment in Virtual Classroom Contexts. *Journal of Language and Communication (JLC), 8,* 121-132.

Hamer, J., Purchase, H. Luxton-Reilly, A., & Denny, P. (2015). A Comparison of Peer and Tutor Feedback. *Assessment & Evaluation in Higher Education, 40,* 151-164. https://doi.org/10.1080/02602938.2014.893418

Harris, L. R., & Brown, G. T. L. (2013). Opportunities and Obstacles to Consider When Using Peer- and Self-Assessment to Improve Student Learning: Case Studies into Teachers' Implementation. *Teaching and Teacher Education, 36,* 101-111. https://doi.org/10.1016/j.tate.2013.07.008

Harris, L. R., & Brown, G. T. L. (2018). *Using Self-Assessment to Improve Student Learning*. Routledge. https://doi.org/10.4324/9781351036979

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research, 77,* 81-112. https://doi.org/10.3102/003465430298487

Hung, Y. (2019). Bridging Assessment and Achievement: Repeated Practice of Self-Assessment in College English Classes in Taiwan. *Assessment & Evaluation in Higher Education, 44,* 1191-1208. https://doi.org/10.1080/02602938.2019.1584783

Irons, A. (2004). *Enhancing Learning through Formative Assessment and Feedback*. Routledge.

Javaherbakhsh, M. R. (2010). The Impact of Self-Assessment on Iranian EFL Learners' Writing Skill. *English Language Teaching, 3,* 213-218. https://doi.org/10.5539/elt.v3n2p213

Kavaliauskienė, G. (2004). Quality Assessment in Teaching English for Specific Purposes. *English for Specific Purposes World, 3,* 8-17.

Kunan, A. J., & Jang, E. E. (2009). Diagnostic Feedback in Language Assessment. In M. H. Long, & C. J. Doughty (Eds.), *The Handbook of Language Teaching* (pp. 610-627). Blackwell Publishing Ltd. https://doi.org/10.1002/9781444315783.ch32

Kuncel, N. R., Credé, M., & Thomas, L. L. (2005). The Validity of Self-Reported Grade Point Averages, Class Ranks, and Test Scores: A Meta-Analysis and Review of the Literature. *Review of Educational Research, 75,* 63-82. https://doi.org/10.1007/978-981-10-3924-9

Lado, R. (1961). *Language Testing: The Construction and Use of Foreign Language Tests. A Teacher's Book*. McGraw-Hill Book Company.

Lambert, D., & Lines, D. (2000). *Understanding Assessment: Purposes, Perceptions, Practice.* Routledge Falmer.

Lee, I. (2017). *Classroom Writing Assessment and Feedback in L2 School Contexts*. Springer. https://doi.org/10.1007/978-981-10-3924-9

Lee, I., & Coniam, D. (2013). Introducing Assessment for Learning for EFL Writing in an Assessment of Learning Examination-Driven System in Hong Kong. *Journal of Second Language Writing, 22,* 34-50. https://doi.org/10.1016/j.jslw.2012.11.003

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A. et al. (2009). The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies that Evaluate Health Care Interventions: Explanation and Elaboration. *PLoS Medicine, 6,* e1000100. https://doi.org/10.1371/journal.pmed.1000100

Lin, S. J., Liu, Z. F. & Yuan, S. M. (2001). Web-Based Peer Assessment: Attitude and Achievement. *IEEE Transactions on Education, 44,* 13. https://doi.org/10.1109/13.925865

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

Liu, N. F., & Carless, D. (2006). Peer Feedback: The Learning Element of Peer Assessment. *Teaching in Higher Education, 11,* 279-290. https://doi.org/10.1080/13562510600680582

Mazloomi, S., & Khabiri, M. (2016). The Impact of Self-Assessment on Language Learners' Writing Skill. *Innovations in Education and Teaching International, 55,* 91-100. https://doi.org/10.1080/14703297.2016.1214078

McGonigal, K. (2006). Getting More Teaching Out of Testing and Grading. *Speaking of Teaching, The Centre for Teaching and Learning (CTL) Stanford University, 15,* 1-4.

Murakami, C., Valvona, C., & Broudy, D. (2012). Turning Apathy into Activeness in Oral Communication Classes: Regular Self- and Peer-Assessment in a TBLT Programme. *System, 40,* 407-420. https://doi.org/10.1016/j.system.2012.07.003

Nakata, Y. (2010). Improving the Classroom Language Proficiency of Non-Native Teachers of English: What and How? *RELC Journal, 41,* 76-90. https://doi.org/10.1177/0033688210362617

Oscarson, M. (1989). Self-Assessment of Language Proficiency: Rationale and Applications. *Language Testing, 6,* 1-13. https://doi.org/10.1177/026553228900600103

Oscarson, M. (2013). The Challenge of Student Self-Assessment in Language Education. *Voices in Asia Journal, No. 1,* 1-14.

Panadero, E. (2011). *Instructional Help for Self-Assessment and Self-Regulation: Evaluation of the Efficacy of Self-Assessment Scripts vs. Rubrics*. Unpublished Doctoral Dissertation, Universidad Autónoma de Madrid.

Panadero, E., Jonsson, A., & Strijbos, J. (2016). Scaffolding Self-Regulated Learning through Self-Assessment and Peer Assessment: Guidelines for Classroom Implementation. In D. Laveault, & L. Allal (Eds.), *Assessment for Learning: Meeting the Challenge of Implementation* (pp. 311-326). Springer. https://doi.org/10.1007/978-3-319-39211-0_18

Ross, J. A., & Bruce, C. D. (2007). Teacher Self-Assessment: A Mechanism for Facilitating Professional Growth. *Teaching and Teacher Education, 23,* 146-159. https://doi.org/10.1016/j.tate.2006.04.035

Rust, C., Price, M., & O'Donovan, B. (2003). Improving Students' Learning by Developing Their Understanding of Assessment Criteria and Processes. *Assessment and Evaluation in Higher Education, 28,* 147-164. https://doi.org/10.1080/02602930301671

Salehi, M., & Sayyar Masoule, Z. (2017). An Investigation of the Reliability and Validity of Peer, Self-, and Teacher Assessment. *Southern African Linguistics and Applied Language Studies, 35,* 1-15. https://doi.org/10.2989/16073614.2016.1267577

Samaie, M., Nejad, A. M., & Qaracholloo, M. (2016). An Inquiry into the Efficiency of WhatsApp for Self- and Peer-Assessments of Oral Language Proficiency. *British Journal of Educational Technology, 49,* 111-126. https://doi.org/10.1111/bjet.12519

Sambell, K., McDowell, L., & Sambell, A. (2006). Supporting Diverse Students: Developing Learner Autonomy via Assessment. In C. Bryan, & K. Clegg (Eds.), *Innovative Assessment in Higher Education* (pp. 158-168). Routledge.

Shepard, L. A. (2000). The Role of Assessment in a Learning Culture. *Educational Researcher, 29,* 4-14. https://doi.org/10.3102/0013189X029007004

Snead, L. O., & Freiberg, H. J. (2017). Rethinking Student Teacher Feedback: Using a Self-Assessment Resource with Student Teachers. *Journal of Teacher Education, 70,* 155-168. https://doi.org/10.1177/0022487117734535

Sullivan, K., & Lindgren, E. (2002). Self-Assessment in Autonomous Computer-Aided

Second Language Writing. *ELT Journal, 56,* 258-266. https://doi.org/10.1093/elt/56.3.258

Summers, M. M., Cox, T. L., McMurry, B. L., & Dewey, D. P. (2019). Investigating the Use of the ACTFL Can-Do Statements in a Self-Assessment for Student Placement in an Intensive English Program. *System, 80,* 269-287. https://doi.org/10.1016/j.system.2018.12.012

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

Ünaldı, I. (2016). Self and Teacher Assessment as Predictors of Proficiency Levels of Turkish EFL Learners. *Assessment and Evaluation in Higher Education, 41,* 67-80. https://doi.org/10.1080/02602938.2014.980223

Wang, W. (2016). Using Rubrics in Student Self-Assessment: Student Perceptions in the English as a Foreign Language Writing Context. *Assessment & Evaluation in Higher Education, 42,* 1280-1292. https://doi.org/10.1080/02602938.2016.1261993

Xi, X., & Davis, L. (2016). Quality Factors in Language Assessment. In D. Tsagari, & J. Banjeree (Eds.), *Handbook of Second Language Assessment* (pp. 61-76). de Gruyter. https://doi.org/10.1515/9781614513827-007

Zimmerman, B. J., & Moylan, A. R. (2009). Self-Regulation: When Metacognition and Motivation Intersect. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of Metacognition in Education* (pp. 299-315). Routledge.

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). Lawrence Erlbaum Associates.

Appendix 1

Overview of selected studies:

Author(s)/Year of publication/Context	Methods/participants/instrument	Findings
Hung (2019)/Taiwan area	Mixed methods/97 college students/survey and interview	Repeated SA bridged the gap helped the students reflect upon their performance, find their weaknesses, adjust their following talk, and recognize their learning outcomes.
Borg and Edmett (2019)/Instrument was developed by British Council and 1684 teachers from 125 countries participated	Quantitative/2598 individuals/Survey questionnaire	SA has marginal effect on young students learning, however, their SA accuracy grows in repeated SA practice.
Summers et al. (2019)/United State	Quantitative/92 newly enrolled students in ESL program/Questionnaire (SA tests and placement tests)	SA can be used a reliable instrument to discriminate between examinees.
Butler (2018)/Japan	Mixed methods/31 Primary school students/5-point Likert scale Survey and interviews	Students' age plays an important role in SA: older students make more conservative evaluations.
Liu and Brantmeier (2019)/Chinese Mainland	Quantitative/106 Chinese learners/ Reading test, writing tasks, questionnaire	Young learners are capable of doing accurate SA
Snead and Freiberg (2017)/United States	Qualitative/10 student teachers/student feedback, audio recordings of lessons, and the educator's self-assessment.	SA through PCLA enables teachers to self-reflect on their teaching and empowers them to plan their future work.
Adachi et al. (2016)/Australia	Qualitative/13 Australian academics/ Semi-structured interviews	They have reported the benefits and challenges of SA as a learning material
Salehi and Sayyar Masoule (2017)/Iran	Quantitative/32 upper-intermediate EFL learners/5-point Likert scale Questionnaire	Significant correlation between SA and teacher assessment in assessing written production, no significant correlation between SA and teacher assessment in assessing oral production
Samaie et al. (2016)/Iran	Mixed methods/50 EFL learners/ Questionnaires, interviews/Pretest	Participants were not positive towards self and peer-assessment on WhatsApp, and they preferred face-to-face interactions
Ünaldı (2016)/Turkey	Quantitative/239 Turkish EFL learners/ Questionnaire/the comparison of scores in an objective placement test assigned by their instructors and the results of criterion-referenced SA scores.	Although self-assessment is an important predictor of language proficiency, it seems to be much weaker than teacher assessment.
Wang (2016)/Chinese Mainland	Qualitative/80 university students/ reflective journals and case study informants' retrospective interviews	SA rubrics helped the learners in all three stages of self-regulated learning, namely forethought, performance and reflection
Mazloomi and Khabiri (2016)/Iran	Quantitative, Quasi-experimental/60 Translation students/Writing composition	self- and peer-assessment have significant role in enhancing writing skill

Continued

Duque Micán and Cuesta Medina (2015)/Colombia	Mixed methods/24 young-adult learners/Vocabulary pre-test/post-test, Learning log Survey	SA had positive impact on learners' vocabulary, fluency and learning process
Murakami et al. (2012)/Japan	Quantitative/99 students/Pre-test/ post-test design with self-rated survey/Open-ended questionnaire	Regular SA conducted by led to significant increases in the students' frequency of spoken English in the class and a higher level of engagement with English learning outside the classroom
Birjandi and Hadidi Tamjid (2012)/Iran	Quantitative, Experimental/157 TEFL students/Pre-test/post-test	SA makes students involved in learning process and teacher's feedback enhances their writing performance
Nakata (2010)/Japan	Quantitative/21 teachers and student teachers/Likert type questionnaire	The study confirms the developmental effect of a teacher SA tool.
Bullock (2011)/Ukraine	Mixed methods/10 English teachers/ Questionnaire and interviews	All teachers had positive attitudes towards SA and the only negative aspect of SA according to teachers was 'being time-consuming"
Butler and Lee (2010)/South Korea	Quantitative/254 young EFL learners/Pre-test/post-test	SA has marginal effect on students' learning