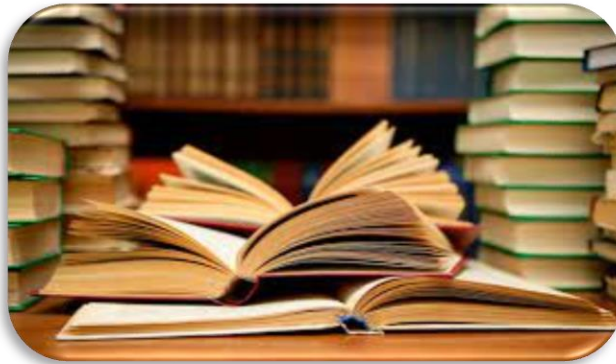## Research Paper

Designing and Validating a Self-Assessment Tool for Improving the Speaking Skill of Iranian EFL Learners

Ali Fathi Karizak[1] , Shahram Afraz[2]* , Fazlolah Samimi[3]

[1]Ph.D. Candidate, Department of English, Qeshm Branch, Islamic Azad University, Qeshm, Iran.
ali.fathi1839@yahoo.com

[2]Assistant Professor, Department of English, Qeshm Branch, Islamic Azad University, Qeshm, Iran.
a.sh32@rocketmail.com

[3]Assistant Professor, Department of English, Bandar-Abbas Branch, Islamic Azad University, Bandar-Abbas, Iran.
fazl.samimi67@gmail.com

ABSTRACT

This study aimed to examine the different components of the speaking self-assessment instrument, and the most and least influential factors in the self- assessment instrument affecting Iranian EFL learners' speaking skill. Finally, utilizing a grounded theory, the model of self -assessment Instrument affecting Iranian EFL learners' speaking skill was designed and validated. Accordingly, 20 language experts and 350 Iranian EFL learners were selected based on convenience sampling. The data were gathered via interviews and questionnaire. The data was analyzed by MAXODA. The findings revealed two broad categories of most influential factors in the self-assessment instrument affecting Iranian EFL learners' speaking skill are classroom practice concerns EFL teachers' actions in EFL speaking classrooms and conceptions referring to their beliefs about assessment in general and self-assessment in particular. In addition to teacher assessment, it was also reported using peer assessment can be influential. The participants explained that sometimes comment on their peers' speaking productions can be conductive. Regarding the assessment techniques, the participants did not provide much information concerning the assessment procedures as no reference was made to any particular assessment techniques or tasks. As far as the least influential factors are concerned, teachers' formative feedback, and assessment criteria were mentioned, implying that they do not have any significant impact on self-assessment instrument affecting Iranian EFL learners' speaking skill. Consequently, the model of self-assessment instrument affecting Iranian EFL learners' speaking skill emerges includes some characteristics and advantages, as well as disadvantages and challenges

Keywords: *Iranian EFL Learners، MAXODA ، Self-Assessment ، Speaking Skill*

این پژوهش به منظور بررسی اجزای مختلف ابزار خودارزیابی گفتاری و عوامل تأثیرگذارترین و کمترین در ابزار خودارزیابی موثر بر مهارت گفتاری دانش‌آموزان ایرانی زبان انگلیسی انجام شد. در نهایت، مدل خودارزیابی موثر بر مهارت گفتاری دانش‌آموزان ایرانی زبان انگلیسی براساس استفاده از تئوری پایه‌گذاری شد. به این منظور، بر اساس نمونه‌گیری آسان، ۲۰ معلم زبان انگلیسی و ۳۵۰ دانش‌آموز زبان انگلیسی ایرانی انتخاب شدند. داده‌ها از طریق مصاحبه‌ها و پرسشنامه جمع‌آوری شدند. تحلیل داده‌ها توسط نرم‌افزار MAXODA انجام شد. نتایج نشان داد که دو دسته‌بندی گسترده از عوامل تأثیرگذارترین در ابزار خودارزیابی موثر بر مهارت گفتاری دانش‌آموزان ایرانی زبان انگلیسی، فعالیت‌های کلاسی در مورد روش‌های معلمان زبان انگلیسی در کلاس‌های گفتاری زبان انگلیسی و نگرش‌ها در مورد خودارزیابی هستند. همچنین نشان داده شد که ارزیابی همتا می‌تواند در زمینه‌های زبان انگلیسی به عنوان زبان خارجی تأثیرگذار باشد. به نظر شرکت‌کنندگان، نظر دادن درباره عملکرد گفتاری همتایان خود می‌تواند موثر باشد. در مورد تکنیک‌های ارزیابی، اطلاعات بیشتری درباره روش‌های ارزیابی ارائه نشد زیرا هیچ اشاره‌ای به هیچ تکنیک یا عملیات خاصی نشده بود. در مورد اجزای کمتر تأثیرگذار، بازخورد فرماتیو معلمان و معیارهای ارزیابی ذکر شد که نشان می‌دهد که هیچ تأثیر معنی‌داری بر ابزار خودارزیابی موثر بر مهارت گفتاری دانش‌آموزان ایرانی زبان انگلیسی ندارند. همانطور که نشان داده شد، مدل خودارزیابی موثر بر مهارت گفتاری دانش‌آموزان ایرانی زبان انگلیسی شامل برخی ویژگی‌ها و مزایا، و همچنین معایب و چالش‌ها می‌شود

واژه های کلیدی: فراگیران زبان انگلیسی به عنوان زبان خارجی؛ MAXODA، خودارزیابی؛ مهارت گفتاری،

INTRODUCTION

During the last two decades, teaching foreign languages shifted drastically, which resulted in the emergence of new assessment methods. Such assessment methods were mainly based on the learner centered paradigms. According to Grabin (2017), the movement from teacher-centered methods to communicative language teaching (CLT) approach indicated the change from conventional assessment approaches to alternative assessment (AA) approaches, highly based on learner-centered.

Other contributing element in shifting towards AA approaches includes the restriction of conventional assessment. In fact, the conventional assessment methods did not satisfy the education assessment systems adequately, which results in the contemplating on the new assessment methods such as peer-assessment and self-assessment, affective than conventional assessment methods. As stated by Black and William (2015), the shift from conventional assessment to alternative assessment led to assessment improvement as well as learning development. According to Hargreaves, et al. (2019), new developments in the field of AA have resulted in an interest in AA techniques such as self-assessment and peer-assessment. In the same vein, Esfandiari and Myford (2013) argue that AA can be considered as a fully-fledging outlet and perspective for education, classroom activities, and assessment in the area of language education in the early 1990s. The last decade witnessed growing interest toward learner-centered syllabi, the issues of self-assessment and peer assessment in testing and evaluation contexts. Further, Oscarson (2019) maintained that the development of self-assessment techniques lends support from moving toward the learners' judgment on their abilities and activities.

Self-assessment is considered as a process in which learners are supposed to evaluate their own behavior and assignments and to think about their learning and development in achieving the desired purposes (Chapelle et al., 2015). As mentioned by Fan and Knoch (2019), self-assessment is a formative means that allows the learners to recognize their full potentials, and take the learners' attention to their behavior and learning via a reflection rather than any grade. In language education history, speaking skill possesses a significant position (Fulcher, 2011). During two decades ago, speaking skill came about as sub-discipline of testing and education in its own. For a number of foreign language curricula, speaking skill was a fundamental skill to students' development and future success. Therefore, finding the affective AA techniques such as self-assessment is of a dire need. Galaczi and Taylor (2018) opined that self-assessment allows the students to become cognizant of their ultimate performance so that they can identify their strengths and weaknesses. In the same line, to Hirai and Koizumi (2013), self-assessment fosters the learners' autonomy, which is an important element of autonomous learning. Further, the promotion of self-monitoring and managing the cognitive processes is undeniable, which through conscious control the knowledge development arises (Vygotsky, 1962). On the other hand, lack of understanding of the assessment process and learners being reluctant of assessing thinking is a teacher's duty, training them through various stages of support would be beneficial to learning (Carter & Nunan, 2001; Oskarsson, 1997). Reviewing the related literature revealed that there is no valid self-assessment tool to evaluate Iranian EFL learners' speaking skill. Accordingly, the current study was an endeavor to validate a self-assessment tool for speaking skill of Iranian EFL learners in Iranian context. Accordingly, the current study was an endeavor to validate a self-assessment instrument for speaking skill of Iranian EFL learners in Iranian context. the current study aimed to fill such a gap in literature by answering the following research questions:

What are the different components of the speaking self-assessment instrument?

What are the most and least influential factors in the self- assessment instrument affecting Iranian EFL learners' speaking skill?

Utilizing the grounded theory, what model of self -assessment Instrument affecting Iranian EFL learners' speaking skill emerges?

REVIEW OF RELATED LITERATURE

Self-assessment

A number of definitions have been presented on self-assessment in literature. For instance, Blanche and Merino (1989) argues that self-assessment addresses the learners' information and evaluation about their performance, progress and abilities. They also argued that the learners need to be aware of their potentials after being taught. For Brown (2004), self-assessment is as an "absurd reversal of politically correct power relationships" (p. 32). ELT scholars emphasize that self-assessment is extremely noteworthy since it is so closely associated with the learners' autonomy and self-regulation. According to Harris (1997), self-assessment assists the language learners to monitor their development and abilities, and makes a connection between learning to their real language needs. For some ELT researchers such as Bachman (2000), Fulcher, (2011), Galaczi and Taylor (2018), Oscarson (1989) defined the self-assessment by identifying two categories of self-assessment based their aims including performance-oriented self-assessment, and development-oriented self-assessment. Performance-oriented self-assessment usually assesses the learners' performance at some point in time. However, whereas performance-oriented self-assessment evaluates the participants for a further time to recognize their progress. One of the models which aptly presents the model of self-assessment is Rolheiser's (1996) self-evaluation model.

Figure 1

Rolheiser's (1996) Self-evaluation Model

According to the model, self-evaluation is the blend of the self-reaction and self-judgment element within the model, and if learners are being taught to self-evaluate better they can reach a better learning. As stated by Wicaksana, et al, (2019), teachers must involve during the learning process to assist language learners move on an upward path.

Empirical Studies

Some research studies have been conducted on speaking self-assessment. For instance, in a study by Ariafar and Fatemipour (2013), the impact of the EFL learners' self-assessment on the speaking skill was calculated and reported. Their results revealed that self-assessment enhanced the EFL learners' speaking skill. In addition, the participants showed their positively perception on self-assessment. In a similar study, Abolfazli Khonbi and Sadeghi (2012) examined the impact of self-assessment, peer- assessment, and teacher-assessment methods on general English proficiency of Iranian EFL students. The results of the study showed the differences in the impact of the three methods and revealed that peer-assessment was the most effective. In the same vein, Javaherbakhsh (2010) inquired on the self-assessment impact on the writing skill of Iranian EFL students. The findings indicated self-assessment influenced the learners' writing skill. In another study, the use of self-assessment in EFL learners' writing class was estimated by Meihami and Varmaghani (2013). The results indicated that self-assessment enhanced the writings skill. Their findings verified that self-assessment was effective in EFL learners' writing skill. Bahmani (2014) explored the impacts of self-assessment on the oral skill of Iranian EFL students. The findings confirmed the efficacy of that self-assessment on learners' oral skill.

By conducting a narrative review of empirical research, Fan and Yan (2020) systematically analyzed the articles on language assessment. Their findings showed that the topics most frequently explored by ELT speaking assessment scholars were the components of rater effects, and speaking ability, as well as the models for self-assessment. In another research, Vasileiadou and Karadimitriou (2021) probed into the learners' self-assessment to enhance the learners' overall performance. The findings indicated that self-assessment had a positive impact on learners' general writing performance. Khonamri et al. (2021) examined the self-assessment as a conductive strategy for improving teaching and learning, especially for oral performance. The results showed that self-assessment was affective in improving the learners' oral skill. In sum, the related literature shows there is no single study to design and validate a self-assessment tool for improving Iranian EFL learners' speaking skill.

METHOD

Participants

This study followed a qualitative research design to investigate a self-assessment tool for developing Iranian learners' the speaking skill. The first group of participants were 20 EFL teachers (including 15 males and 5 females), studying at a language institute (Noandishan Institute) in Bandar Abbas, Iran. They were teaching English at high intermediate and advanced levels, ranging from 34 to 44. The second group of participants included 350 undergraduate learners majoring in English translation in Azad university of Bandar Abbas, studying English for at least three years.

Materials and Instruments

The main data collection instrument applied in this stage of the study was a semi-structured interviews and a questionnaire.

Interviews

The first collection instrument applied in this stage was a semi-structured interview. To provide the first draft of the interview items, the literature on the concept of self-assessment was surfed, and self-assessment underlying factors were identified. To pilot on the interview and prepare the final interview questions, some participants as well as two colleagues were asked to scrutinize the interview items and provide their comments. Accordingly, some interview questions were modified and deleted.

Questionnaire

The questionnaire made from the results of the interview was also utilized to examine the model. The questionnaire made from the results of the interview was also utilized to examine the model.

Procedure

To analyze the data and answer the research questions, the interviews were examined and their codes were identified in response to the design components of the self-assessment model for the development of speaking skills of Iranian language. Table 1 presents the identified codes. Then, after interviewing these participants open, axial and selective codes were extracted. It should be mentioned that theoretical saturation was achieved after the interview.

Data Analysis

The design indicators of the self-evaluation model for the development of speaking skills of Iranian language learners were:

· Help for better educational and administrative performance
· Recognize the problematic areas that exist in schools
· Enhance innovative reflection and behavior
· The influence of the level of interpersonal relationships
· Needs a specific timeframe
· Needs cooperation of all the stakeholders
· Self-assessment reports may not be considered
· Self-evaluation helps a teacher to increase his awareness
· Self-evaluation helps teachers reach a more accurate reality
· Acquaintance of learners about the process and its value
· Provide opportunities for practice in the classroom
· Feedback and opportunities for constant reflection and revision
· Stimulate learners to learn individually
· Insufficient awareness
· Insufficient motivation

·   Insufficient language proficiency
·   Learners' preference for teacher assessment instead
·   Insufficient training with SA
·   Heavy workload
·   Intensive course
·   Lack of time for training
·   May cause pressure, anxiety and nervousness in teachers
·   Cause insecurity and anxiety

The theoretical saturation table for the design indicators of the self-assessment model for the development of speaking skills of Iranian language learners is as follows:

Table 1

The design indicators of the self-assessment model of the development of speaking skills of Iranian language learners

Code System	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Help for better educational and administrative performance																				
Identify the weak points that exist in schools																				
Cultivating new, innovative and extensive thinking and action																				
The influence of the level of interpersonal relationships																				
Needs a specific timeframe																				
Needs cooperation of all the stakeholders																				
Self-assessment reports may not be considered																				
Self-evaluation helps a teacher to increase his awareness																				
Self-evaluation helps teachers reach a more accurate reality																				
Acquaintance of students about the process and its value																				
Provide opportunities for practice in the classroom																				
Feedback and opportunities for constant reflection and revision																				
Stimulate students to learn individually																				
Lack of awareness																				
Lack of motivation																				
Lack of language proficiency																				
Students' preference for teacher assessment instead																				
Lack of training with SA																				
Heavy workload																				
Intensive course																				
Lack of time for training																				
May cause pressure, anxiety and nervousness in teachers																				
Create a feeling of fear, anxiety and insecurity																				

Table 2

*Theoretical saturation table for the design indicators of the self-assessment model*

Code System	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	SUM
Help for better educational and administrative performance	5	5	6	1	2	3	3			4		5		3	4	3	3	2	1		50
Identify the weak points that exist in schools			2	2	4			1		2	4	5	3	4	1	2	5	1			36
Cultivating new, innovative and extensive thinking and action	4	4	4			7	6							7	4	4	4		1		45
The influence of the level of interpersonal relationships	2	1		3	3	3	5	3	3		1	2	2	3	1	1	3	4	3		43
Needs a specific timeframe	5	4		2	2	4	4	3	2	1	1	5		2					1		36
Needs cooperation of all the stakeholders	1	1	2	1	2	10	8	1		8	9	7	7	4	2		1	4	4	2	74
Self-assessment reports may not be considered	1		1		5	3	2	3	2	3	4	8	5	7	3	1	1	1	3	3	56
Self-evaluation helps a teacher to increase his awareness	1	1		4	2	4	2	1	1	3		6	2	4	2	1	1	4	4	2	45
Self-evaluation helps teachers reach a more accurate reality			1	2	4	1	1	4	4	3	7	1	3	2	1			3	5	4	46
Acquaintance of students about the process and its value			1	1	2	1	6	4	4	4	1	1	6	2		1	2	1	1		38
Provide opportunities for practice in the classroom			3	3	6			4	4	3	3	15	6	3				9	2	1	62
Feedback and opportunities for constant reflection and revision	2	2	3	2	3	6	3	6	4	2	9	10	6	1	5	2	1	2	2	2	73
Stimulate students to learn individually			2	2	4	2	1	3	2	2	6	4	3	4	3	1	1				42
Lack of awareness			1	1	4			6	5	4	6	5	3	6	3	3	6	2	2		61
Lack of motivation			3		4			3	1	4	4	5	2	4			5	4	2		41
Lack of language proficiency	3	3		2	6			6	6	2		2	1	2	1		2	1			37
Students' preference for teacher assessment instead	1	1		3	6	6	4	4	9			1									36
Lack of training with SA			1	2	4			6	4	3	3	2	1	1	4	2	3	3	1	1	41
Heavy workload			1	2	1			8	7	1	2	2	8	1	3	1	2	4			43
Intensive course	8	7	8	3								1	3	3	3		3				39
Lack of time for training	8	7	11	6	1	2	1	1		1	3	5		4	4	4	4	2	4	1	69
May cause pressure, anxiety and nervousness in teachers	1	1				7	7	3	2	4	8		5	1	5	4	5		12	2	67
Create a feeling of fear, anxiety and insecurity	2	2	3	2	5		4	2	1	3	2	1	11	1	1	1	7	1	1		50
Σ SUM	44	39	54	39	64	60	48	78	57	65	68	93	64	76	61	34	37	65	57	27	1130

Next, open, axial and selective codes were extracted after interviewing these participants. It should be mentioned that theoretical saturation continued after conducting the interview. Then, interviews and coding were finished. It was done in MAXQDA 2018 qualitative data analysis software. The final outcome of the extracted components and indices is as follows.

Figure 2

Final output of teacher opinions for components and indicators for the design indicators of the self-assessment model
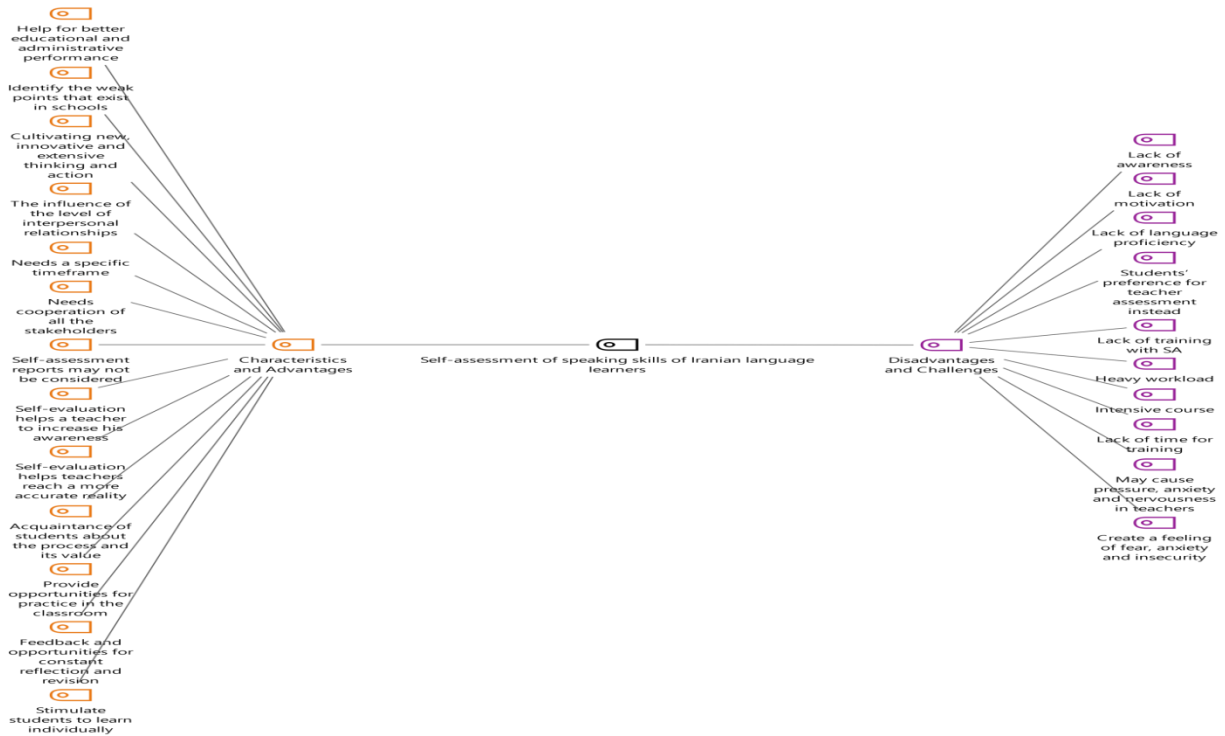

RESULTS AND DISCUSSION

RESULTS

In this study, for gathering quantitative data, 370 questionnaires were distributed among the target population, 350 of which were returned completely. In order to collect data, questions were divided into demographic questions and research questions. After collecting questionnaires, statistical analysis was performed. By using tables and graphs, descriptions of both demographic indices and research questions were presented. Furthermore, for rejection or confirmation of the questions, they were tested on the basis of the results of the survey using inferential statistics. In the first place, the descriptive statistics were examined. Then, confirmatory factor analysis (CFA) was run to examine the questionnaire validity. Finally, Structural Equation Modeling (SEM) with SmartPLS 3 software was used to investigate the research questions.

Figure 3

Standard numbers of research questions
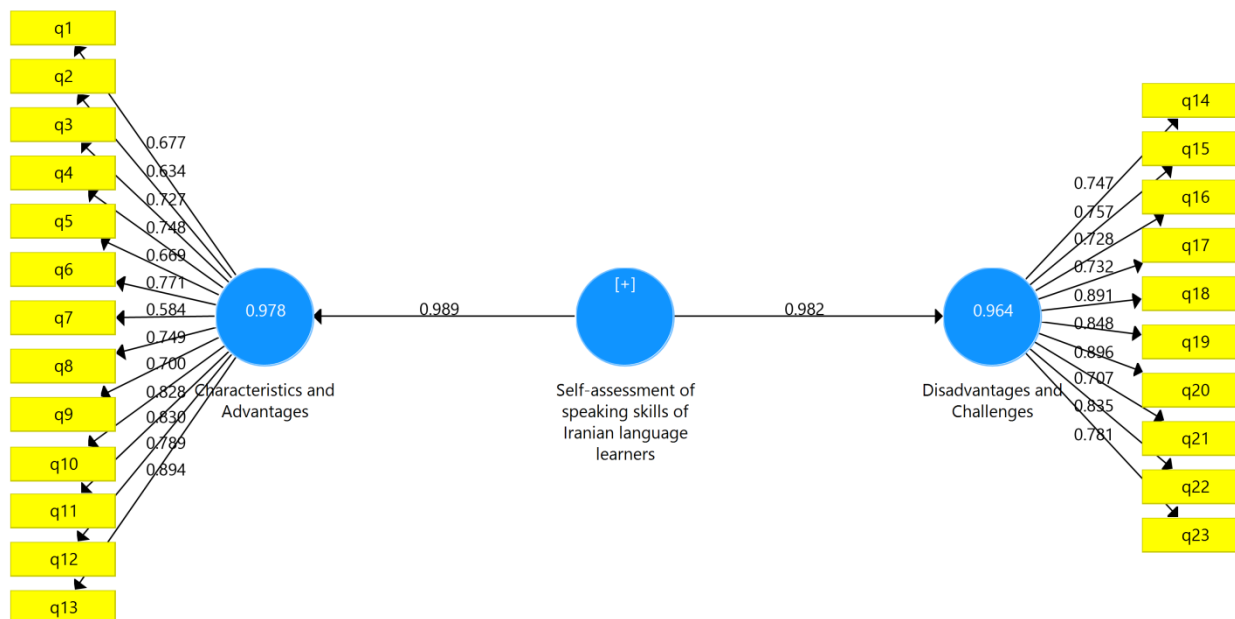
Factor loads are of great importance in interpreting the results of factor analysis. Depending on how accurately the researcher considers omitting questions, criterion values ranging from 0.5 to 0.7 have been introduced but the lowest stated value is 0.4 . This means that questions with factor loads of less than 0.4 should be eliminated. Based on the result, all numbers are above 0.4, so no questions will be omitted.

Figure 4

Significant numbers of research Questions
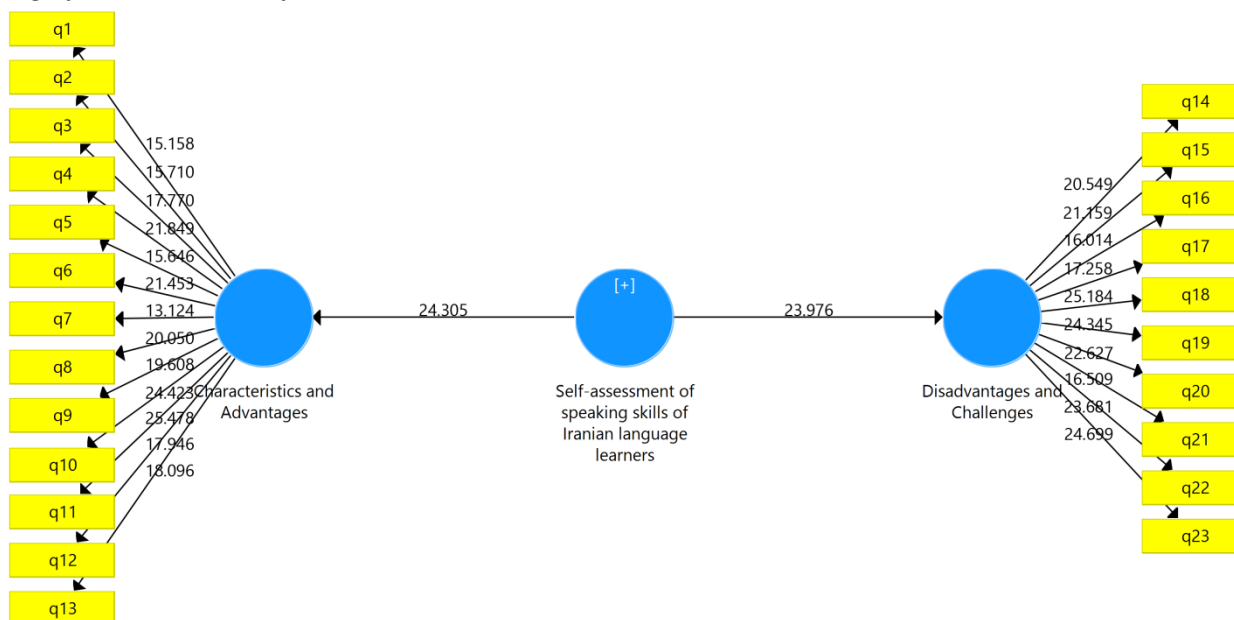
In structural model, only the hidden variables are examined along with the relationships between them. Criterion R^2 determines the impact of an exogenous variable on an endogenous variable. It is important to note that the value of R^2 is calculated only for the dependent (endogenous) constructs of the model, and in the case of exogenous constructs, the value of this criterion is zero. The higher the value of R^2 for the endogenous constructs of a model, the better the fitness of the model. Table 3 presents R^2 coefficient of determination, which was acceptable for the purpose of the study, verifying that the model may predict the features of the endogenous elements (i.e. constructs).

Table 3

The R^2 coefficient of determination

The dependent variable	R^2	Intensity
Characteristics and Advantages	0.978	Strong
Disadvantages and Challenges	0.964	Strong
Average	0.971	Strong

Table 4 presents the predicative quality (Q^2).

Table 4

Predictive Quality(Q^2)

The dependent variable	Q^2	Intensity
Characteristics and Advantages	0.369	Strong
Disadvantages and Challenges	0.351	Strong
Average	0.360	Strong

In goodness of fit (GOF), according to Schmidt (2004), the values of 0.01, 0.25 and, 0.36 are given as weak, medium and strong, respectively. Therefore, the results confirmed the fitness of the speaking self-assessment model in Iranian EFL context.

$$GOF = \sqrt{\overline{Communality} \times \overline{R^2}} = \sqrt{0.360 \times 0.971}$$
$$= 0.59$$

DISCUSSION

The current study was set to examine the different components of the speaking self-assessment tool, and the most and least influential factors in the self- assessment tool contributing to Iranian EFL learners' speaking skill development. It also explored the self-assessment model contributing to speaking skill of Iranian EFL students which might emerge from the data. According to the reported answers, influential factors were categorized into two thematic elements as personal conceptions and classroom activities. The former addresses the attitudes towards speaking self-assessment, and the latter refers to the EFL instructors' actions in EFL speaking contexts. Not only teacher assessment, but also peer assessment was important for the participants. However, the results indicated the inefficacy of the assessment techniques

and procedures in enhancing speaking skill. As far as the least influential factors are concerned, teachers' formative feedback, and assessment criteria were mentioned, implying that they do not have any significant impact on self- assessment model affecting Iranian EFL learners' speaking skill. As shown, the self-assessment model affecting speaking skill of Iranian EFL learners' emerges includes some characteristics and advantages, as well as disadvantages and challenges. Each one is enumerated for the ease of reading:

I.   Characteristics and Advantages:

   a. To result in a better theoretical and pedagogical implications
   b. To recognize the problematic areas that exist in schools
   c. To enhance innovative and creative reflection and behavior
   d. to be affected by the level of interpersonal interactions
   e. needs a specific timeframe
   f. needs cooperation of all the stakeholders
   g. might go unnoticed.
   h. increase teachers' understanding of their teaching direction.
   i. self-assessment can help teachers step away from an idealized depiction of performance and into a more accurate reality.
   j. Enhance the students' knowledge on speaking skill assessment
   k. Give much more opportunities for speaking practice
   l. Give feedback to reflect and revise.
   m. stimulate learners to learn individually

II.   Disadvantages and Challenges
   a. Insufficient of awareness
   b. Insufficient of motivation
   c. Insufficient of language proficiency
   d. Learners' preference for teacher assessment instead
   e. Lack of training with SA
   f. Burdensome workload
   g. Insufficient training time
   h. may cause pressure, anxiety and nervousness in teachers
   i. to cause insecurity and fear

CONCLUSION AND IMPLICATIONS

These findings of the study are supported by those of previous studies conducted in other countries (e.g. Alek et al. 2020a; Benettayeb-Ouahian, 2016; Broadfoot et al., 2002; Bullock, 2011; Szőcs, 2015). These results are in line with Benettayeb-Ouahian's (2016) study in which teachers favored the importance of learners' active engagement in self-assessment design and implementation. Her findings also revealed that whenever the language learners' self-reflective practice was entangled with a supportive context they can obtain their full speaking skill potential. The results can be also explained by Vygotsky's (1978) socio-cultural theory. According to the theory, some mediating factors prepare and give motivation to

the learners to be responsible for their learning within ZPD. Self-assessment can be integrated in the EFL context practice, however; the curriculum policies might not allow further information on the procedures and techniques of the implementations of self-assessment. In the same vein, according to Broadfoot et al. (2002), Further, the correct utilization of self-assessment needs teachers' essential assessment knowledge and literacy, which are dependent upon a supportive context that assist instructors to put their theoretical knowledge of self-assessment into practice.

The findings also indicated EFL teachers' knowledge on the different challenges in using self-assessment. They mentioned that these challenges were related to time limitations, policy and curriculum issues. The EL teachers maintained that they might be unable to change some of the current assessment policies, and made attempt to cope with these challenges by collaborative collegiality and put self-assessment practices and experience in their teaching profession. Based on the model emerged from the data, EFL teachers argued that getting their English students to communicate in English with an acceptable proficiency of grammar, vocabulary, pronunciation accuracy and fluency was their biggest concern. Alek et al. (2020a) also pointed out that self-assessment can be considered as a technique to evaluate students' oral skills. They drew into the conclusion that self-assessment might enhance future oral results and engagement of the students. As far as the advantages of the self-assessment of speaking skill are concerned, learners obtain feedback on their learning, it raises their autonomy, and they enhance critical thinking to assess themselves. Further, language students felt that they would control their own learning progress, it enhanced the value of responsibility in them, it assists them to eradicate the fear about the teacher's behaviors and reactions in the classrooms.

In concluding, EFL teachers are required to engage language learners in the elaboration of the procedures of conducting self-assessment in EFL settings to provide time to discuss the criteria of their performance to equalize the educational aims with the learner's needs and preferences. In addition, EFL teachers are supposed to show learners how to utilize the self-assessment criteria they made to evaluate their own performance. Moreover, EFL teachers presented feedback of self-assessment utilization. Additionally, EFL teachers are supposed to assist language student to set up new learning aims based on their current self-assessment literacy. As mentioned, they are supposed to utilize the self-assessment when they have sufficient time.

More importantly, to apply self-assessment, some contributing factors should be taken into account to utilize self-assessment technique in speaking skill such as learners' social or cultural background, gender, age, external and internal motivation etc. In fact, evaluating the students' actions in classroom instruction is one of the most important tasks which teacher should carry out. Having high assessment enables teachers to make informed decisions which take account of both teaching and learning. On the other hand, teachers' low assessment may damage the education quality (Popham, 2009). As mentioned by Gronlund and Linn (1990), EFL teachers who evaluate the instruction might initially find out their language students' needs, observe learning and educational processes, recognize the learners' learning difficulty, and confirm learning achievement. Teachers should have the skills to choose an appropriate assessment method to make decisions about the quantity and quality of their instruction. Having said this, language teachers are bound to should assessment by adopting various ways such as participating in pertinent workshops on assessment, reading research reports and books, and pursuing academic

degrees. Given the paramount importance of assessment, teacher educators should incorporate assessment topics into their pre- and in-service teacher education programs. The programs should entail both theoretical discussion of assessment topics and the practical use of them. The programs should empower the teachers to apply assessment results to make decisions about each learner, to plan for instruction, to develop the curriculum, and to improve the school.

EFL/ESL instructors need to possess the skills to communicate evaluation results with language students, parents, and other policy makers. They are supposed to report assessment results to learners and parents. Further, language teachers are gradually requested to report the assessment results. To communicate language students' assessment results to others, language teachers should use the assessment concepts aptly as well as to communicate the meaning, restrictions, and utilization of the assessment results. Furthermore, teachers sometimes need to protect their own assessment process and its interpretation, and they may also need to help the public interpret the assessment results appropriately. The following sections provide the theoretical and pedagogical implications of the study. The study expanded the ways of assessing EFL learners self-assessment by considering the various aspects of self-assessment in terms of the participants' attitudes. It also directed research towards investigating EFL teachers' self-assessment within larger contexts using various research studies. Additionally, this study might contribute to field of ESL and EFL self-assessment and educational assessment. In addition, teachers' knowledge about self-assessment types, procedures, and precepts might be a great assistance for them to achieve and complete their teaching mission appropriately. Further, through being aware of effective self-assessment procedures, EFL instructors would be able to recognize and document whether or not suitable progress has been shown by learners in the classroom. Moreover, the theoretical implications of the current study would contribute to the literature on self-assessment and verify Taylor's (2013) assumptions that different stakeholders in language assessment have different needs.

The study has some pedagogical implications. First, it addresses the EFL teachers, and the self-assessment the teachers need to be improved. Therefore, pre-service teacher education programs should emphasize the teachers as assessors. Second, the results have implications for in-service teachers, effective rater training basement program should be initiated. Pre- and in-service education is important because educating pre- and in-service teachers establish one of the most essential aspects in the fitness of language testing and self-assessment. As the third implications, the school principals should provide language teachers with support for implementing self-assessments by allowing them to be flexible in the assessment. Fourth, EFL teachers should take the significance of the self-assessment into consideration. Fifth, the findings of this study indicated that teacher self-assessment is limited and that it still has its deficits in spite of increased TEFL teacher training in assessment. Furthermore, the findings of this study can immensely assist stakeholders such policy makers and material developers to recognize the difference between reliable and unreliable self-assessment and to do activities consistent with suitable self-assessment. Moreover, novice TEFL and Non-TEFL teachers can make use of the findings of this study to enhance their understanding on self-assessment. In addition, another implication of the present study is that self-assessment needs to be considered more seriously in both pre-service and in-service TEFL teachers' assessment courses. Given the paramount importance of self-assessment, ministry of education should place a higher value on fostering self-assessment of all teachers in general and language teachers in particular. Also, this ministry is in charge of supporting and developing teachers' assessment.

Accordingly, educational assessment reforms are required for language teachers and current educational context. For EFL teachers, more self-assessment courses are needed to be held in the curriculum of language teacher education and the content should pave the way to equate theory and practice. Depending on the results, it can be suggested that EFL teachers should be given some self-assessment training courses. The necessary precautions to solve learners' needs should be taken. While training has a significant impact on self-assessment, future research may need to examine the objectives, tasks, contents and results (essentially the effectiveness) of self-assessment courses provided by teacher preparation institutions and professional development programs. Additionally, in the current study only EFL learners were taken into consideration. Therefore, it is suggested that in the future studies various groups of the participants such as ESL teachers be examined. Further, the sample of this study was small, so using a larger sample might lead to better and generalizable findings. This study also showed lack of teacher-involved assessment. Thus, it is suggested that further investigation on teacher-involved assessment be conducted. It is usually hard to impose change without knowing the underlying factors behind resistance or lack of practice. Future research should also look at how prepared EFL teachers are to use learner-involved self-assessment.

References

Ahmadi, A., & Sadeghi, E. (2016). Assessing English language learners' oral performance: a comparison of monologue, interview, and group oral test. *Language Assessment Quarterly,. 13*, 341–358. doi: 10.1080/15434303.2016.1236797

Alderson, J. C. (2005). *Diagnosing foreign language proficiency: The interface between learning and Assessment*. Bloomsbury.

Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation.* Cambridge University Press.

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford University Press.

Bachman, L. F., & Palmer, A. S. (1996). *Language assessment in practice: Designing and developing useful language tests*. Oxford University Press.

Bonk, W. J., & Ockey, G. J. (2003). A many-facet Rsch analysis of the second language group oral discussion task. *Lang. Test. 20*, 89–110. doi: 10.1191/0265532203lt245oa

Bosker, H. R., Pinget, A.-F., Quené, H., Sanders, T., & De Jong, N. H. (2013). What makes speech sound fluent? The contributions of pauses, speed and repairs. *Lang. Test. 30,* 159–175. doi: 10.1177/0265532212455394

Brown, A. (2003). Interviewer variation and the co-construction of speaking proficiency. *Lang. Test. 20*, 1–25. doi: 10.1191/0265532203lt242oa

Carter, R., & McCarthy, M. (2017). Spoken grammar: where are we and where are we going? *Appl. Linguistics 38,* 1–20. doi: 10.1093/applin/amu080

Chapelle, C. A., Cotos, E., & Lee, J. (2015). Validity arguments for diagnostic assessment using automated writing evaluation. *Lang. Test. 32,* 385–405. doi: 10.1177/0265532214565386

Chapelle, C. A., Enright, M. K., and Jamieson, J. (2010). Does an argument-based approach to validity make a difference? *Educ. Meas. Iss. Pract. 29*, 3–13. doi: 10.1111/j.1745-3992.2009. 00165.x

Chapelle, C. A., Enright, M. K., & Jamieson, J. M. (2008). *Building a validity argument for the test of English as a foreign language*. Routledge; Taylor & Francis Group.

Eckes, T. (2005). Examining rater effects in TestDaF writing and speaking performance assessments: a many-facet Rasch analysis. *Lang. Assess. Q: Int. J.* 2, 197–221. doi: 10.1207/s15434311laq0203_2

Eckes, T. (2011). *Introduction to many-facet Rasch measurement.* Peter Lang.

Ellis, R. (2015). Introduction: complementarity in research syntheses. *Appl. Linguistics 36,* 285–289. doi: 10.1093/applin/amv015

Fan, J., & Knoch, U. (2019). Fairness in language assessment: what can the Rasch model offer. *Lang. Test. Assess. 8,* 117–142. Available online at: http://www. altaanz.org/uploads/5/9/0/8/5908292/8_2_s5_fan_and_knoch.pdf

Fulcher, G. (2000). The 'communicative' legacy in language testing. *System, 28,* 483–497. doi: 10.1016/S0346-251X(00)00033-6

Fulcher, G. (2015). Assessing second language speaking. *Lang. teaching, 48*, 198–216. doi: 10.1017/S02614448140 00391

Fulcher, G., Davidson, F., & Kemp, J. (2011). Effective rating scale development for speaking tests: performance decision trees. *Lang. Test. 28*, 5–29. doi: 10.1177/0265532209359514

Galaczi, E., & Taylor, L. (2018). Interactional competence: conceptualisations, operationalisations, and outstanding questions. *Lang. Assess. Q. 15,* 219–236. doi: 10.1080/15434303.2018.1453816

Galaczi, E. D. (2008). Peer-peer interaction in a speaking test: the case of the First Certificate in English examination. *Lang. Assess. Q. 5,* 89–119. doi: 10.1080/15434300801934702

Gan, Z. (2012). Complexity measures, task type, and analytic evaluations of speaking proficiency in a school-based assessment context. *Lang. Assess. 9,* 133–151. doi: 10.1080/15434303.2010.516041

Hirai, A., & Koizumi, R. (2013). Validation of empirically derived rating scales for a story retelling speaking test. *Lang. Assess, 10,* 398–422. doi: 10.1080/15434303.2013.824973

Iwashita, N. (2006). Syntactic complexity measures and their relation to oral proficiency in Japanese as a foreign language. *Lang. Assess, 3,* 151–169. doi: 10.1207/s15434311laq0302_4

Jang, E. E., Wagner, M. & Park, G. (2014). Mixed methods research in language testing and assessment. *Annu. Rev. Appl. Linguistics, 34*, 123–153.

Kim, A. A., Chapman, M., Kondo, A., & Wilmes, C. (2020). Examining the assessment literacy required for interpreting score reports: A focus on educators of K-12 English learners. *Language Testing, 37*(1), 54–75.

Klebanov, B. B., Ramineni, C., Kaufer, D., Yeoh, P., & Ishizaki, S. (2019). Advancing the validity argument for standardized writing tests using quantitative rhetorical analysis. *Language Testing, 36*(1), 125–144.

Longabach, T., & Peyton, V. (2018). A comparison of reliability and precision of subscore reporting methods for a state English language proficiency assessment. *Language Testing, 35*(2), 297–317.

McNamara, T. (2006). Validity in language testing: The challenge of Sam Messick's legacy. *Language Assessment Quarterly, 3(*1), 31–51.

Mertler, C. A., & Campbell, C. (2005). *Measuring teachers' knowledge & application of classroom assessment concepts*: *Development of the" assessment literacy inventory.* Online Submission.
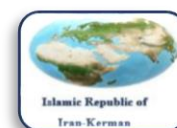
Ng, W. S., Xie, H., & Wang, F. L. (2018). *Enhancing teacher assessment literacy using a blended deep learning approach*. Paper presented at the International Conference on Blended Learning.

Ockey, G. J. (2014). *Exploratory factor analysis and structural equation modeling*. In A. J. Kunnan (Ed.). The companion to language assessment: Abilities, contexts and learners volume III (pp. 140-160). John Wiley & Sons, Inc.

Razavipour, K. (2013). Assessing assessment literacy: Insights from a high-stakes test. *Research in Applied Linguistics, 4(1)*, 111–131.

Salimi, E. A., & Farsi, M. (2018). An Investigation of assessment literacy among native and nonnative English teachers. *Journal of English Language Teaching and Learning, 10*(22), 49–62.

Schils, E., van Der Poel, M., & Weltens, B. (1991). The reliability ritual. *Language Testing, 8*(2), 125–138.

Tommerdahl, J., & Kilpatrick, C. D. (2014). The reliability of morphological analyses in language samples. *Language Testing, 31*(1), 3–18.

Weideman, A. (2019). Assessment literacy and the good language teacher: four principles and their applications. *Journal for Language Teaching, 53*(1), 103–121.

Weideman, A. (2019a). Degrees of adequacy: the disclosure of levels of validity in language assessment. *Koers, 84*(1), 1–15.

Weideman, A. (2019b). Validation and the further disclosures of language test design. *Koers, 84*(1), 1–10.

Xu, Y. (2018). Assessment in the language classroom: teachers support learner learning. *Language Assessment Quarterly, 15*(4), 423–425.

Youn, S. J. (2020). Managing proposal sequences in role-play assessment: Validity evidence of interactional competence across levels. *Language Testing, 37*(1), 76–106.

Biodata

Mr. Ali Fathi Karizak is a Ph.D. candidate of TEFL at Islamic Azad University Qheshm Branch, Iran. He has been involved in teaching English for some years at universities and language institutes.His main areas include language teaching, teaching theories and skills, and testing
Email:*ali.fathi1839@yahoo.com*

Dr. Shahram Afraz (Corresponding Author) is an associate professor of applied linguistics. Shahram Afraz, assistant professor of TEFL, is the faculty member of Islamic Azad University, Qeshm Branch. He has been teaching English for twenty years at several universities .His areas of interest are English language teaching, linguistics, syllabus design, and testing.He has published more than forty articles and three books till now
Email:*a.sh32@rocketmail.com*

Dr. Fazlolah Samimi is an Assistant Professor of TEFL. He has published several papers in national and international journals.
Email:*Fazl.Samimi67@gmail.com*