https://doi.org/10.15446/profile.v25n2.103917

# Exploring Criteria for Evaluating In-Service English Language Teachers' Performance

## Exploración de los criterios para evaluar el desempeño de los maestros de inglés en servicio

**Azadeh Hassani**

**Zari Saeedi**

Allameh Tabataba'i University, Tehran, Iran

This case study explored the evaluation criteria applied by six supervisors of a private language institute (three men and three women) through individual in-depth interviews in an English-as-a-foreign-language context. The researchers also collected data from supervisors' observation checklists and written feedback. A thematic analysis resulted in five main themes and two sub-themes. Therefore, a tentative framework was developed, encompassing five criteria: English and Content Knowledge, Teaching Skills, Personal Traits, Fulfilling Workplace Expectations, and Parents'/Learners'/Peers' Feedback. The proposed framework can help increase language supervisors' teacher evaluation literacy.

*Keywords:* language supervisor, language teacher evaluation, performance supervision, supervisors' criteria

Este estudio de caso exploró los criterios de evaluación aplicados por seis supervisores de un instituto privado de idiomas (tres hombres y tres mujeres) en un contexto de inglés como lengua extranjera. Para recoger datos se usaron entrevistas individuales en profundidad, listas de verificación de observación de los supervisores y comentarios escritos. Un análisis temático permitió identificar cinco temas principales y dos subtemas. Así, se desarrolló un marco tentativo que abarca cinco dominios: el inglés y el conocimiento del contenido, las habilidades de enseñanza, los rasgos personales, el cumplimiento de las expectativas del lugar de trabajo y la retroalimentación de los padres/estudiantes/compañeros. El marco propuesto puede ser útil para aumentar el conocimiento de los supervisores de idiomas sobre la evaluación docente.

*Palabras clave:* criterios de los supervisores, evaluación del profesor de idiomas, supervisión del desempeño, supervisor de idiomas

Azadeh Hassani https://orcid.org/0000-0002-5807-3769 . Email: ahassani4@huskers.unl.edu
Zari Saeedi https://orcid.org/0000-0003-2165-8891 . Email: Saeedi.za@atu.ac.ir

Azadeh Hassani is currently a PhD student in Innovate Learning Technologies at the University of Nebraska-Lincoln (USA).

How to cite this article (APA, 7th ed.): Hassani, A., & Saeedi, Z. (2023). Exploring criteria for evaluating in-service English language teachers' performance. *Profile: Issues in Teachers' Professional Development*, *25*(2), 165–181. https://doi.org/10.15446/profile.v25n2.103917

This article was received on July 26, 2022 and accepted on March 24, 2023.

## Introduction

The enduring integration of evaluation and teaching has turned evaluation into an inseparable part of teaching English as a second/foreign language (ESL/EFL), which is also applied in various aspects, including teacher development. In general, there is an evidence-supported relationship between teachers' effectiveness and students' academic achievement (Canales & Maldonado, 2018; Marzano & Toth, 2013; Phillips et al., 2014; Podolsky et al., 2019; Rachmajanti, 2008). Hence, numerous studies have focused on evaluating language teachers' performance, teaching quality, and effectiveness in different contexts, mainly from teachers', lecturers', or teacher trainers' perspectives (Khaksefidi, 2015; Mashhadlou & Izadpanah, 2021; Mazandarani & Troudi, 2017; Mousavi et al., 2016; Ostovar-Namaghi, 2013; Rashidi & Forutan, 2015; Wei & Hui, 2019). Due to teacher evaluation's multifaceted and complicated nature, it is necessary to collect the views of all educational stakeholders, including language supervisors, apart from teachers' perceptions.

Traditionally, these in-house evaluators bear the rather challenging and demanding responsibility of regular supervision of language teachers' performance and performing unpleasant duties—such as giving teachers negative feedback (Bailey, 2006)—to measure and maintain the quality of language institutes' educational services and to improve teaching practices (Chen & Cheng, 2013). Even with the advent of mobile and video-recording technologies, administrators still give more weight to employing supervisors to personally observe and contextually assess language teachers' performance and then discuss issues in teachers' practices during post-observation meetings (Sadeghi & Richards, 2015).

While these supervisory practices would be undoubtedly influential in teachers' professional growth and development (Richards & Farrell, 2005) and in their level of self-esteem and security (Ponticell et al., 2019; Vásquez, 2004), it seems that investigating the evaluative criteria of language supervisors, particularly in EFL contexts, has not received sufficient attention (Akbari & Yazdanmehr, 2012). Mixed feelings and attitudes of EFL teachers toward supervision (Gholaminejad, 2020; Ostovar-Namaghi, 2013) make it harder to determine comprehensive and all-embracing criteria for evaluating them. Thus, by conducting an in-depth qualitative case study in a well-established language institute in the Iranian EFL context, we tried to reveal supervisors' criteria for evaluating language teachers' performance.

## Literature Review

### Language Teacher Supervision

As part of teacher evaluation systems, supervision holds a significant status due to its direct and indirect impacts on teaching quality. Generally, supervision in educational settings has been defined as an interactive, facilitative process aimed at teachers' professional development and improving classroom practice (Ponticell & Zepeda, 2004). To Gebhard (1990), "language teacher supervision is an ongoing process in which the supervisor observes what goes on in the teacher's classroom with an eye toward the goal of improved instruction" (p. 2). Different supervisory approaches have been suggested for evaluating the effectiveness of EFL/ESL teachers (e.g., Bailey, 2006; Freeman, 1982; Gebhard, 1990; Knop, 1980). They primarily differ regarding the supervisor's and the teacher's roles in the supervisory cycle. Bailey (2006) provides a detailed description of various forms of language teacher supervision and performance evaluation. Emphasizing contextual factors, she mentions some evaluative criteria, including (a) evaluators' judgments or opinions (often based on their beliefs and attitudes rather than on predetermined criteria), (b) the teacher's teaching method, and (c) teacher competencies and performance standards.

Regarding the Iranian EFL contexts, supervision procedures mainly conform to the "scientific approach" (Knop, 1980), the "supervisory approach" (Freeman, 1982), the "directive model" (Gebhard, 1990), and the "prescriptive approach" (Bailey, 2006). These approaches assess the quality of teaching against students' achievement and suggest that the supervisor's role is to observe and provide feedback to the teacher. Moreover, teachers have little power and voice in any of these approaches. While the power imbalance ruling the supervisory process is confirmed by both teachers and supervisors (Agheshteh & Mehrpour, 2021), Iranian English language institutes still prefer to rely on their supervisors' judgments to ensure and improve the quality of their educational services.

Various aspects of language teacher supervision have been targeted in research carried out in EFL/ESL contexts. Conducting two separate studies, Azizpour and Gholami (2021a, 2021b) investigated seven teacher supervisors' and 218 teachers' attitudes toward language supervision. The findings of the former shed light on supervisors' qualifications, responsibilities, and concerns, and those of the latter revealed that, despite the anxiety-provoking nature of the supervision process, many EFL teachers found it beneficial. Hişmanoğlu and Hişmanoğlu (2010) explored non-native and native language teachers' perceptions of educational supervision at three universities. According to the results, most teachers complained about the supervisors' judgmental attitudes.

Moreover, classroom observation was found relatively unfruitful in developing teachers professionally. Akbari and Yazdanmehr (2012) examined Iranian EFL teachers' recruitment and assessment criteria in five institutes. The resulting four-element categorization for teacher assessment included command of English, teaching skills, compliance with the syllabus, and personal/affective features.

Focusing on one aspect of language teacher supervision, Hatamvand et al. (2020) attempted to develop and validate a Language Teacher Observation Scale. After gathering data from 540 Iranian English teachers, a six-factor model of EFL teacher observation was developed, including cognitive considerations, classroom management and teacher behavior, metacognitive considerations, preparing for the lesson, social-interaction considerations, and teacher knowledge. Another study explored the quality standards applied in private language institutes to evaluate EFL teachers' professional competence (Mousavi et al., 2016). Consequently, five standards emerged in the following order, from the most to the least frequent: "describing language and understanding language acquisition and development process," "planning, managing, and implementing instruction," "assessment skill," "cultural competence," and "professionalism skill." These studies represent teacher supervision and evaluation as two processes that go hand in hand. In other words, supervision is conducted based on predetermined, often research-supported criteria. The dynamic nature of a language classroom lays the ground for the language supervisor to devise and apply additional evaluative criteria.

## Language Teacher Evaluation

Despite various definitions, models, and frameworks for teacher evaluation, researchers agree that it should be conducted in a regular and formative manner to ensure the achievement of institutional goals, focused educational improvement, and accountability of educators for their instruction (Phillips et al., 2014). Danielson and McGreal (2000) believe that an effective teacher evaluation system defines the teaching domain coherently and introduces clear standards for acceptable performance. In addition, through trained evaluators capable of making evidence-based decisions, consistent judgments are made in this system to assess all aspects of teaching based on defined procedures. According to Isoré (2009), the objectives of teacher evaluation are twofold: (a) ensuring that

teachers make concerted efforts to enhance student learning and (b) informing teachers of their strengths and weaknesses. The fundamental challenge in supervision and evaluation is activating a supervisory sense in teachers that guides them while teaching (Marshall, 2009). When aiming at developing teaching capacity, teacher evaluation is achieved through supervision conducted by administrators, supervisors, or peers (Hallinger et al., 2014).

Two of the most widely-adopted teacher evaluation frameworks in mainstream education are those of Marzano and Toth (2013) and Danielson (2013). Marzano and Toth's model, which is founded on teacher growth and student achievement, embraces four domains: "classroom strategies and behaviors," "planning and preparing," "reflecting on teaching," and "collegiality and professionalism." Danielson's framework, which mainly depends on classroom observation, is composed of "planning and preparation," "the classroom environment," "instruction," and "professional responsibilities." After all, EFL teachers are teachers in the first place, and to be considered efficient, they should generally embody the desirable characteristics of their counterparts in other subjects.

Bell (2005) explored effective foreign language teaching from the perspectives of 457 postsecondary teachers of French, German, and Spanish. More than 95% of the participants agreed on the teacher's enthusiasm for the target language and culture, competence in using the target language, frequent use of authentic realia and materials, and group work in the classroom. Çelik et al. (2013) sought to reveal 998 undergraduate Turkish students' perceptions regarding the attributes of effective EFL teachers. The results suggested that the significant criteria were personal qualities, content and pedagogy-specific knowledge, professional skills, and classroom behavior. Ninety Iranian students and EFL teachers in Khaksefidi's (2015) study agreed upon 13 components as their primary criteria for an effective EFL teacher. Non-linguistic factors, such as teachers' appearance and discipline, were regarded as crucial as their linguistic abilities. Another study explored the perceptions of Iranian students majoring in English language teaching and translation about effective EFL teachers (Zamani & Ahangari, 2016). The findings reflected students' expectations of effective EFL teachers in having the ability to establish rapport, build up students' confidence, and maintain discipline in the classroom.

Mazandarani and Troudi's (2017) exploratory study investigated Iranian EFL lecturers' perceptions of the qualities of an efficient EFL/ESL teacher. The findings led to an effective teaching model composed of teachers' traits, cognitive and metacognitive qualities, and pedagogical and professional skills. Wei and Hui (2019) compared Vietnamese university students' and two administrators' views about their EFL teachers' performance. The students preferred teachers who promoted classroom interaction and student engagement, while the administrators associated good teaching with focusing on instructional techniques. Griffiths and Tajeddin (2020) enumerated similar characteristics for efficient language teachers. They also point out other standards, such as up-to-date technical knowledge and familiarity with feedback techniques, classroom management practices, instructional strategies, and assessment procedures. A myriad of studies has also sought to explore the characteristics of effective EFL/ESL teachers (e.g., Barnes & Lock, 2013; Brown, 2009; *Demiröz* & *Yeşilyurt*, 2015; Shishavan & Sadeghi, 2009; Tarajová & Metruk, 2020). To our knowledge, almost no study has focused merely on exploring the evaluation criteria of EFL supervisors, who have the most determining power in the supervisory processes. Through an exploratory qualitative study, we tried to explore the main criteria applied by language supervisors in their evaluation of Iranian in-service EFL teachers' performance.

## Method

### Participants and Context of the Study

The context for the current study was a well-established private language institute in Tehran, Iran. Moreover, the head supervisor of this institute was an acquaintance of the first researcher. Therefore, we managed to gain access to the target population of this study. For these reasons, this institute was selected for data collection. Before selecting the participants, the head supervisor contacted the supervisors in different branches to gain their consent for participation. Finally, six supervisors (three men and three women) in five different institute branches agreed to participate in this study. Hence, they were selected through convenience sampling due to their accessibility and availability to the researchers (Ary et al., 2019). The institute's managers agreed to collaborate with us under the condition that the anonymity of the institute was preserved. The data collection procedures were carried out in December 2020.

The supervisors came from different educational backgrounds, and their ages ranged between 28 and 45, with an average of seven years of experience in supervision. Moreover, they were all considered experienced language teachers since they had more than ten years of teaching English to learners with different proficiency levels. Regarding research ethics, all participants' anonymity was guaranteed, so we used pseudonyms in all transcriptions. Two supervisors, Maryam and Sarah, worked in the same branch. All supervisors, except one female supervisor, Tina, were willing to reveal their personal/professional information, as shown in Table 1.

### Instrumentation

To answer the research question comprehensively, we employed a triangulation of instruments. Due to the privacy policies of the institute, we were not allowed to access the teachers' demographic information.

#### Semi-Structured Interviews

Based on our developed interview protocol, we conducted semi-structured interviews with each supervisor to gain an in-depth understanding of their criteria for evaluating language teachers' performance. Six open-ended questions were formulated based on a comprehensive review of the related literature and our knowledge of supervisory procedures in Iranian language institutes. All questions were consulted with

**Table 1.** Supervisors' Demographic Information

| Name | Gender | Age | Educational background | Teaching experience (Years) | Supervision experience (Years) |
|---|---|---|---|---|---|
| Maryam | Female | 45 | BA in English Literature | 20 | 2 |
| Sarah | Female | 40 | MA in TEFL | 10 | 7 |
| Tina | Female | NA | NA | NA | NA |
| Shayan | Male | 32 | BSc in Mining Engineering | 10 | 5 |
| Parsa | Male | 36 | PhD in TEFL | 17 | 14 |
| Sina | Male | 28 | MBA | 12 | 7 |

*Note*. NA = Not applicable, TEFL = Teaching English as a foreign language.

two PhD holders in Teaching English as a Foreign Language (TEFL) and modified or reworded to ensure the credibility of the interview protocol. After interviewing the first participant (Shayan), we revised the questions to make them more transparent and precise and remove any confusion (Ary et al., 2019). The interviews were carried out in the workplaces of the supervisors. The following questions were asked:

1. What are your primary criteria for evaluating in-service EFL teachers' performance?
2. How do you observe EFL teachers' instruction?
3. What is your most important criterion for evaluating EFL teachers' performance?
4. How often do you observe EFL teachers?
5. What procedures do you follow to observe EFL teachers' performance?
6. What factors, apart from those in observation checklists, affect your judgment during observation?

#### Supervisors' Checklists and Written Feedback

To increase the depth of the inquiry, we utilized the observation checklists and supervisors' written feedback, which were the outcome of their observations and post-observation meetings. Under the supervision of Shayan and Parsa, two branches used the same observation checklist consisting of five sections, each including several sub-sections. The checklist contained 15 statements, with responses on a three-point Likert scale (*exceeds expectations*, *meets expectations*, and *needs improvement*), plus some extra space for the supervisor's comments. At the bottom of the checklist, a separate section was dedicated to the "overall impression of teaching effectiveness." The other three branches had designed their observation forms. Taking an entirely qualitative approach, one branch—supervised by Maryam and Sarah—relied merely on the supervisors' comments. The other two—supervised by Tina and the other by Sina—had enumerated several different criteria in their checklists and separate sections for the evaluator's comments and suggestions.

### Data Collection Procedure

The semi-structured interviews were the primary data source, lasting between 25 and 60 minutes, depending on each interviewee's availability. Initially, the meetings were arranged through the institute headquarters office. The interviews took place at five different branches where the participants worked. During each session, the first author guided the discussion by raising clarifying questions, and the interviewee was allowed to elaborate on their evaluative criteria. The interviews—four in English and two in Persian due to the participants' preferences—were tape-recorded and later transcribed verbatim. At the end of each session, the interviewer was provided with the supervisors' observation checklists.

### Data Analysis

We adopted a qualitative approach to analyze interview transcriptions and supervisory documents. First, we reviewed the scripts multiple times. Then the salient features of the content were coded through initial coding (Saldaña, 2016). In this phase, qualitative data was broken down into discrete parts and closely examined to generate tentative and provisional codes. The second data analysis phase was devoted to categorizing codes obtained from the first phase. Hence, significant themes and sub-themes were recognized through focused coding. Two PhD students in TEFL reviewed the extracted themes and sub-themes to increase the reliability of the coding process. Through consultation and revision, some sub-themes were modified or recategorized. Finally, we and the reviewers reached a reasonable level of agreement.

## Findings

The thematic data analysis revealed five main criteria applied by supervisors for evaluating in-service EFL teachers' performance. We devised a tentative framework, consisting of five themes and two sub-themes, for evaluating the performance of in-service

EFL teachers (see Figure 1). This framework subsumes all the issues of concern in the evaluation procedure followed by the supervisors of the investigated institute.

### English and Content Knowledge

All supervisors endorsed the importance of having adequate general English knowledge as an evident and crucial prerequisite for recruiting EFL teachers in the first place. Considering it his first and foremost criterion, Parsa believed that:

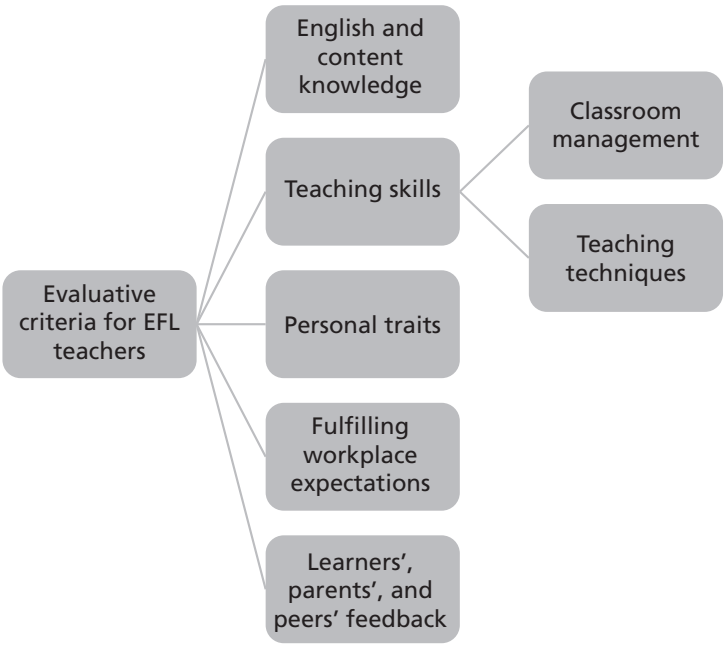
> For the first [criterion], their general English, I myself subcategorize it into seven different groups, and that is their listening skill, reading, writing, speaking, grammatical knowledge, vocabulary command, and pronunciation, though I am not super sensitive to accent.

Maryam associated this feature with the teachers' ability to answer language learners' questions, especially adults who tend to be more goal-oriented in their language learning endeavors. Viewing the EFL teacher as a role model for the learners, Shayan explained the significance of teachers' English pronunciation:

> The other important thing is the teacher's pronunciation, intonation, or enunciation. Because teachers are role models for students. Students usually learn the pronunciation of a word the way that their teachers would pronounce it; they usually go and dig it up in dictionary to see whether the teacher was right or not.

A significant share of evaluators' comments in their checklists was dedicated to teachers' grammatical errors. Moreover, teachers' Persian-accented pronunciation and intonation were highlighted as a sign of their need to attend on-the-job workshops. On the other hand, the supervisors had contradictory opinions about teachers' educational backgrounds and topical knowledge. Shayan, who holds a non-TEFL degree, denied the effects of having an academic major on English language teaching. Sarah, a TEFL degree holder, firmly believed that non-TEFL-degreed teachers, such as electrical engineers, are not familiar with methods and approaches to teaching the English language. Going one step further, she emphasized having a related academic background as one of her

**Figure 1.** In-Service EFL Teacher Supervision Framework



preferences for recruiting teachers. The sub-categories of this theme are presented in Table 2.

**Table 2.** Theme 1: English and Content Knowledge

| |
|---|
| • General English knowledge |
| • Mastering pronunciation, intonation, and enunciation |
| • Topical knowledge |
| • Educational background |
| • IT/Computer literacy |

## Teaching Skills

Regarded as an ability that is even more valued than teachers' educational background, teaching skills were thoroughly addressed by the participants. This theme was broken down into two sub-themes: teaching techniques and classroom management (see Table 3).

### Teaching Techniques

Shayan was the only supervisor who insisted on the rule of a teacher as a facilitator by employing contextualization methods:

> We are in favor of teaching everything through context. Our approach is that all students are going through a context, whether they are listening or reading an article or a conversation, and, actually, students teach themselves. The teacher is just the guide by asking the right questions at the right moment.

Shayan also noted the significance of a teacher's anticipatory skills to predict probable student- or hardware-related issues, such as a CD player breakdown or any questions learners might ask during a given session. Parsa preferred EFL teachers to be creative within the methodological boundaries of their teachers' guides rather than teaching differently from their colleagues since it might induce a sense of insecurity in learners. The concept of "techneme" was only reflected in Parsa's remarks: "Technemes are what makes teachers individual teachers and, like, makes them different. . . . They are like games, [which teachers can use to] present techniques in a creative and individualized way."

Tina magnified the importance of mastering various error correction methods rather than using the traditional ones (i.e., repeating learners' errors and providing them with correct forms). Another essential element repeatedly viewed in the supervisors' written feedback was their emphasis on reducing teacher talking time and increasing student talking time.

### Classroom Management

All the supervisors emphasized the role of classroom management in evaluating teachers' professional competence. Encouraging learners to participate in class/group activities and simultaneously controlling the noise level were among the criteria mentioned by Shayan and Sina. In a similar vein, Parsa indicated the importance of keeping a balance between applying discipline in the classroom and creating a relaxed educational environment for the learners:

> We cannot push the students so hard in order for them to escape, but on the other hand, there should be discipline! So, teachers must use their voices, their looks, and their classroom scoring criteria. They can take help from me as the authority outside the class if they have to in order to manage a class.

All participants had a clear emphasis on teachers' skills in board management. Tina argued that teachers should not block students' vision and use the space efficiently while writing on the board. Sina emphasized that teachers must form the habit of regularly using the board. He reasoned that this habit would help teachers present more examples and grab learners' attention. Another highlighted factor in observation checklists involved teachers' movements in the classroom. In Sarah's words, taking specific postures, such as "walking aimlessly" and "standing like a statue in one corner," were rebuked.

**Table 3.** Theme 2: Teaching Skills

| Sub-themes | |
|---|---|
| **Teaching techniques** | **Classroom management** |
| • Mastering various teaching approaches/ techniques | • Creating a comfortable and friendly atmosphere |
| • Increasing student talking time and decreasing teacher talking time | • Managing learners' arrangement, behaviors, and level of noise |
| • Following teachers' guidebooks | • Showing appropriate reactions |
| • Ability to impart knowledge | • Encouraging learner engagement |
| • Ability to provide remedial help | • Effective board management |
| • Anticipatory skills | • Making efficient eye contact |
| • Time management | • Posture/dynamics |
| • Using dictionaries | • Punishing uncooperative learners fairly |

### Personal Traits

The supervisors held somewhat similar expectations of EFL teachers regarding their traits. All of them seemed to be completely strict about teachers' punctuality. For instance, Sarah associated this feature with teachers' responsibility:

> [It is] important, even if they are late for two minutes, I would go crazy because I say that, as a teacher, you are responsible for the time of the students and in this way, we give permission to our students to be late too.

Maryam, Sarah's colleague, totally agreed with her. Shayan referred to punctuality as his most important criterion for evaluating EFL teachers. Tina evaluated teachers' creativity based on their ability to present various instructional aids, such as exciting games, posters, or flashcards. She said flexible teachers would adapt themselves to students' different characteristics (age, gender, level of proficiency, or personality). Adding a high weight to a teacher's voice, Tina and Parsa highlighted the importance of vocal features such as clarity, strength, authoritativeness, tone, liveliness, and audibility. Tina mentioned their voice quality would attract her attention while observing a teacher first. Parsa believed teachers should manage the class with their voices and avoid speaking monotonously.

According to Shayan, the teacher's appearance was the first but not the most crucial feature that grabbed his attention during observation. He emphasized that since the institute worked under the direct observation of Iran's Ministry of Education, teachers had to follow this organization's clothing regulations. Parsa, in agreement with Shayan, named appearance as one of his primary criteria:

> They all have to smell good, to look good, to look neat, and to be kind. They must be smiling and be likable people, I don't mean that female teachers must have more makeup or male teachers must particularly do something to their faces, but students must like them.

Sina interpreted setting a dress code as violating teachers' privacy; therefore, he was utterly against adopting strict rules for their appearance:

> I highly respect teachers. I will never destroy their dignity by asking, "Why are you wearing this kind of clothes?" And if anything wrong happens that I have never seen during my supervision experience, I will ask one of our secretaries to speak really respectfully with that person.

All supervisors believed that teachers should avoid distractions that impact students' learning (e.g., wearing very tight clothes, noise-making shoes, too

many accessories, or too much makeup/perfume). In Shayan's opinion, establishing rapport with students was regarded as somehow even more important than a teacher's knowledge:

> Another important thing about the observation is the rapport between the teacher and the students, . . . a lot of people may have the knowledge to teach, but they cannot be good teachers because they are incapable of establishing rapport with the students.

Teachers' patience and tolerance were two of Maryam's highly valued features, especially when dealing with troubled or rebellious teenagers or while eliciting responses from learners. Just one of the supervisors, Shayan, talked about the commitment of teachers. He believed that teachers must be committed to their employers and students and do their best to fulfill their responsibilities. Table 4 summarizes this theme and its components.

**Table 4.** Theme 3: Personal Traits

| | |
|---|---|
| • Appearance | • Politeness |
| • Poise | • Tolerance |
| • Confidence | • Establishing rapport with learners |
| • Voice | • Respecting learners |
| • Patience | • Being humorous |
| • Enthusiasm | • Commitment |
| • Punctuality | • Honesty |
| • Flexibility | • Creativity |

### Fulfilling Workplace Expectations

Since fulfilling workplace requirements directly affect teachers' pay raises or contract extensions, they were categorized as a distinctive theme. A degree of inconsistency was observed among the supervisors in terms of their expectations from teachers. For instance, Sina frequently checked teachers' written lesson plans before starting every instructional session. Maryam took a more holistic approach toward teachers' lesson plans: "We'd like them to prepare and write a lesson plan before a term starts, and during the term, about second or third session, we would check all lesson plans and leave some notes."

However, Tina mentioned that because of her constant monitoring, it would not be necessary for the teachers to deliver their lesson plans. All supervisors unanimously confirmed banning the use of L1 in the classroom. However, Maryam and Sarah were more lenient toward using L1 for teaching abstract concepts or complicated vocabulary. They believed that teachers should use Persian to accelerate students' learning process on these occasions.

Sina enumerated a decrease in learners' registration rate to indicate teachers' poor performance. However, Shayan argued that teachers should not be penalized for any increase in attrition rates. Counting various reasons for such a phenomenon, he viewed teachers' performance as one of the probable causes for a decrease in registration. Table 5 depicts the components of this theme.

### Learners', Parents', and Peers' Feedback

Although not explicitly mentioned by all the supervisors, the level of learners' satisfaction was a determining factor in evaluating the teachers. While five supervisors implicitly referred to it, Sina said that learner satisfaction was his first and most significant criterion in teacher evaluation. He used different methods to receive learners' feedback, such as contacting parents to investigate their opinions about the teachers, conducting regular face-to-face interviews with students, and administering surveys at the end of each term. To explain more, he said:

> We call parents—or in case of adult learners, themselves—regularly to give us descriptive feedback. We avoid yes/no questions and challenge them by asking, "How much Farsi was spoken in the classroom?", "How much time was dedicated to explaining grammar?"

**Table 5.** Theme 4: Fulfilling Workplace Expectations

| | |
|---|---|
| • Scoring 6.5 or above in the annual IELTS mock test<br>• Making contact with young learners' parents<br>• Submitting a written lesson plan before each session<br>• Not speaking in L1 and banning its use in the classroom<br>• Active participation in workshops or on-the-job training courses | • Filling out the teacher's progress chart for each session<br>• Giving regular quizzes<br>• Reducing the number of requests for substitute teachers<br>• Grading learners fairly<br>• Learners' attrition rate |

While trying to avoid "spying on teachers," Parsa used Telegram, a social media platform, to constantly communicate with learners or their parents. Moreover, he implemented a peer-observation scheme to gain more professional feedback about teachers' performance. He was the only supervisor who made teachers, regardless of their experience, observe four colleagues each term and report their observations. Maryam monitored the learners' affection toward their teachers, especially outside the classroom. She asked some questions from the learners to make them talk about their teachers. The components of this theme are presented in Table 6.

**Table 6.** Theme 5: Learners', Parents', and Peers' Feedback

| |
|---|
| • Learners'/parents' satisfaction<br>• Learners' outward expression of feelings toward teachers<br>• Results of interviews with learners (both in English and Persian)<br>• Peer-observation reports |

## Discussion

The present study sought to explore the criteria used by language teacher supervisors to evaluate in-service EFL teachers' performance. Thematic analysis of interview transcriptions and supervisory documents revealed numerous criteria categorized into five themes. Several different models and frameworks for teacher evaluation have been developed in both language and mainstream education from the perspectives of various stakeholders (Akbari & Yazdanmehr, 2012; Danielson, 2013; Hatamvand et al., 2020; Isoré, 2009; Marzano & Toth, 2013; Mazandarani & Troudi, 2017; Mousavi et al., 2016; Ostovar-Namaghi, 2013). The emerging themes conform to some of the criteria and standards presented in the literature. Since supervisors' perceptions, constituting the base of the evaluation hierarchy of EFL teachers, were investigated in this research, several similarities and differences were observed, which will be further elaborated.

The first explored theme was "English and Content Knowledge." Although Iranian EFL teachers' command of English is traditionally assessed through oral and written examinations before recruitment (Akbari & Yazdanmehr, 2012), supervisors have constantly assessed it through in-person or videotaped observations. Teacher knowledge for EFL teachers/lecturers is not only constrained to having fluency and accuracy, but it also entails having topical knowledge, as has been pointed out in numerous studies (Barnes & Lock, 2013; Bell, 2005; Çelik et al., 2013; Coniam et al., 2017; Griffiths & Tajeddin, 2020; Mousavi et al., 2016; Park & Lee, 2006; Wei & Hui, 2019; Yazdanipour & Fakharzadeh, 2020). The contradictory opinions of the participants regarding the effects of EFL teachers'

educational background on their teaching quality are also echoed in administrators' ideas in Yazdanipour and Fakharzadeh's (2020) study. According to their findings, having an academic degree in TEFL is not a strict prerequisite for hiring EFL teachers. Information and communications technology literacy has been reported as a favorable feature of effective English teachers (Çelik et al., 2013; Khaksefidi, 2015; Hatamvand et al., 2020; Mazandarani & Troudi, 2017), and half of the participants in the current study mentioned the necessity of having technological knowledge for the EFL teachers and showing flexibility in using online educational platforms. Similarly, Hsu (2017) stresses the need for EFL teachers to acquire highly technical skills to optimize classroom practices.

In alignment with other studies, a prominent component of effective teaching was categorized as "Teaching Skills," encompassing two sub-themes: "Teaching Techniques" and "Classroom Management." The participants considered mastering various teaching methods a decisive criterion, similar to the teachers' beliefs in the study by Shishavan and Sadeghi (2009). The supervisors gave significant value to communicative language teaching, which conforms to the teachers' perceptions about an ideal foreign language teacher in Brown (2009). One of the classroom strategies addressed by Marzano and Toth (2013) is "organizing the physical layout of the classroom" (p. 43), which has received the attention of the participants in this study as well. Like the present findings, Danielson (2013) believes that a teacher should skillfully create an environment of rapport and respect in the classroom. Likewise, both teachers and learners in Tarajová and Metruk's (2020) study attached great importance to a robust student–teacher relationship. Almost all participants in the present study agreed that learners should not feel the passage of time. Danielson (2013) also contends that teachers must monitor the smooth functioning of all routines in the classroom.

Teachers' personal qualities are subsumed under the third theme, "Personal Traits." The components of this theme have been mentioned in various contexts as eligible yardsticks for evaluating English language teachers' performance (e.g., Griffiths & Tajeddin, 2020; Khaksefidi, 2015; Mazandarani & Troudi, 2017). While none of the participants named physical features as a decisive factor when hiring or evaluating EFL teachers, it seems that this can be of high priority according to some language administrators, to the extent that they even set strict conditions for their teachers' height (Yazdanipour & Fakharzadeh, 2020). Shishavan and Sadeghi's (2009) findings revealed that EFL teachers agreed more firmly than learners on teachers' appearance as a criterion for judging their effectiveness. Fairness was the highest rank attribute in EFL teachers' traits in the findings of Çelik et al. (2013) and Tarajová and Metruk (2020); however, this feature was not even implied by the supervisors in the present study.

Another emerging theme was "Fulfilling Workplace Expectations." Some of its components have been mentioned in the fourth domain of Danielson's (2013) framework under "professional responsibilities." Despite the participants' different degrees of strictness for banning the use of L1 in the classroom, this approach has been suggested for overcoming classroom management challenges (Todorova & Ivanova, 2020). Apart from learners' and parents' feedback, peer observation was a helpful tool in one branch supervised by Parsa. In his view, teachers and supervisors benefited from this observation since it would increase cooperation among colleagues and expedite supervisors' decision-making process. Motallebzadeh et al.'s (2017) findings confirm the practicality and usefulness of peer-observatory practices for EFL teachers.

The last theme was "Learners', Parents', or Peers' Feedback." Based on the findings of Zarrabi and Brown's (2017) study, the high demand of Iranians to learn the English language has caused a significant

growth in the number of language schools. This phenomenon has created a competitive climate among different institutes, and their administrators try to satisfy the learners as much as possible. The degree of English learners' satisfaction is considered a determining factor in evaluating teachers' performance (e.g., Estaji & Shafaghi, 2018; Hatamvand et al., 2020; Zamani & Ahangari, 2016). Being entirely aware of this trend, all the supervisors devised various approaches to receive the learners' feedback. They relied on their observations and strongly believed in communicating directly with the learners or their parents (Azizpour & Gholami, 2021a). However, English teachers in Chinese schools do not entirely favor students' evaluations since their strictness level highly affects these judgments (Murphey & Yaode, 2010). The emphasis of Sina on learners' attrition rates as his most significant criterion seemed to endorse the concept of "commercialization of education," which has been addressed in Yazdanipour and Fakharzadeh's (2020) study. The same feature has been mentioned as the rate of return in Ostovar-Namaghi's (2013) findings and perceived as an unfair judgment criterion by Shayan, who believed that learners' statistics per se could not be considered a reliable criterion for teacher appraisal.

EFL teacher evaluation has multiple facets, and the effectiveness or quality of teaching is unlikely to be assessed with a single measure, such as classroom observation. A range of data will be needed to include diverging yet achievement-oriented views, attitudes, and perceptions of different stakeholders, namely supervisors, learners, parents, administrators, and teachers. Therefore, all the supervisors utilized multiple measures to evaluate the performance and competence of the EFL teachers. The same approach has been recommended by Borg (2018) to capture the complexity inherent in teaching, provide teachers with opportunities to show their performance and competence and minimize any measure's lack of reliability.

## Conclusion and Implications

We aimed to explore the supervisors' criteria for evaluating in-service EFL teachers' performance. Based on the findings, the participants considered five main criteria while observing EFL teachers' practices. These included English and content knowledge, teaching skills, teacher's traits, fulfilling workplace expectations, and learners'/parents'/peers' feedback. Each of these themes embraced several components that carry specific meanings. Although these findings have been more or less addressed in the extant literature (e.g., Akbari & Yazdanmehr, 2012; Hatamvand et al., 2020; Isoré, 2009; Marzano & Toth, 2013; Mazandarani & Troudi, 2017), they still need to be interpreted based on the context of this research. Certain features of EFL teachers, highlighted in similar studies (Shishavan & Sadeghi, 2009; Tarajová & Metruk, 2020; Yazdanipour & Fakharzadeh, 2020) such as physical appearance and fairness, were not among the determining factors considered by the language supervisors in this study.

Regarding the limitations of this study, the focus was only on one language institute in Tehran. While it is impossible to generalize the findings from this specific context, they may be helpful to highlight language teacher supervisors' evaluative criteria. Moreover, only the supervisors' criteria were investigated, and other stakeholders (e.g., the institute's administrators or the EFL learners and their parents) were ignored. A contrastive study could reveal the matches/mismatches between these two groups. The effects of the supervisors' demographic information on their judgments have not been considered. Finding their biases and tendencies according to various variables, such as gender, age, or professional experience, will provide a more precise and clearer picture of their beliefs and thoughts toward evaluating the performance of EFL teachers (Ashtarian & Weisi, 2016).

Unfortunately, the observed teachers' information was not accessible due to the institute's regulations.

Further studies may benefit from exploring the effects of these features on supervisors' judgments. Most participants used the same evaluative criteria regardless of EFL teachers' teaching experience. It seems that designing an advanced EFL teacher evaluation framework tailored to teachers' experience level will be beneficial in increasing their motivation and reinforcing the fairness of supervisory processes (Gan & Yang, 2018).

## References


Agheshteh, H., & Mehrpur, S. (2021). Teacher autonomy and supervisor authority: Power dynamics in language teacher supervision in Iran. *Iranian Journal of Language Teaching Research*, *9*(1), 87–106. https://doi.org/10.30466/ijltr.2021.120977

Akbari, R., & Yazdanmehr, E. (2012). EFL teachers' recruitment and dynamic assessment in private language institutes of Iran. *Journal of English Language Teaching and Learning*, *3*(8), 29–51.

Ary, D., Jacobs, L. C., Irvine, C. K. S., & Walker, D. A. (2019). *Introduction to research in education* (10th ed.). Cengage.

Ashtarian, S., & Weisi, H. (2016). A comparative study of the use of affective markers in supervision comments: Case of native versus non-native supervision comments. *RELC Journal*, *47*(2), 161–173. https://doi.org/10.1177/0033688216631172

Azizpour, S., & Gholami, J. (2021a). EFL teacher-supervisors' attitudes toward supervision in Iranian language schools. *Iranian Journal of English for Academic Purposes*, *10*(1), 16–34.

Azizpour, S., & Gholami, J. (2021b). Exploring EFL teachers' attitudes toward teacher supervision in Iranian language schools. *Journal of Applied Linguistics and Applied Literature: Dynamics and Advances*, *9*(2), 105–125. https://doi.org/10.22049/jalda.2021.27221.1300

Bailey, K. M. (2006). *Language teacher supervision: A case-based approach*. Cambridge University Press. https://doi.org/10.1017/CBO9780511667329

Barnes, B. D., & Lock, G. (2013). Student perceptions of effective foreign language teachers: A quantitative investigation from a Korean university. *Australian Journal of Teacher Education*, *38*(2), 19–36. https://doi.org/10.14221/ajte.2013v38n2.2

Bell, T. R. (2005). Behaviors and attitudes of effective foreign language teachers: Results of a questionnaire study. *Foreign Language Annals*, *38*(2), 259–270. https://doi.org/10.1111/j.1944-9720.2005.tb02490.x

Borg, S. (2018). *Teacher evaluation: Global perspectives and their implications for English language teaching. A literature review*. British Council.

Brown, A. V. (2009). Students' and teachers' perceptions of effective foreign language teaching: A comparison of ideals. *The Modern Language Journal*, *93*(1), 46–60. https://doi.org/10.1111/J.1540-4781.2009.00827.X

Canales, A., & Maldonado, L. (2018). Teacher quality and student achievement in Chile: Linking teachers' contribution and observable characteristics. *International Journal of Educational Development*, *60*, 33–50. https://doi.org/10.1016/J.IJEDUDEV.2017.09.009

Çelik, S., Arikan, A., & Caner, M. (2013). In the eyes of Turkish EFL learners: What makes an effective foreign language teacher? *Porta Linguarum*, *20*, 287–297. https://doi.org/10.30827/digibug.18114

Chen, C. W.-Y., & Cheng, Y.-S. (2013). The supervisory process of EFL teachers: A case study. *TESL-EJ*, *17*(1), 1–21.

Coniam, D., Falvey, P., & Xiao, Y. (2017). An investigation of the impact on Hong Kong's English language teaching profession of the Language Proficiency Assessment for Teachers of English (LPATE). *RELC Journal*, *48*(1), 115–133. https://doi.org/10.1177/0033688216687455

Danielson, C. (2013). *The framework for teaching evaluation instrument*. The Danielson Group.

Danielson, C., & McGreal, T. L. (2000). *Teacher evaluation to enhance professional learning.* ASCD, Educational Testing Service.

Demiröz, H., & Yeşilyurt, S. (2015). Effective foreign language teaching: Perceptions of prospective English

language teachers. *Universal Journal of Educational Research*, *3*(11), 862–870. https://doi.org/10.13189/UJER.2015.031112

Estaji, M., & Shafaghi, M. (2018). Teacher evaluation in EFL context: Development and validation of a teacher evaluation questionnaire. *Issues in Language Teaching*, *7*(2), 147–187.

Freeman, D. (1982). Observing teachers: Three approaches to in-service training and development. *TESOL Quarterly*, *16*(1), 21–28. https://doi.org/10.2307/3586560

Gan, Z., & Yang, C. C. R. (2018). How prepared are the preservice ESL teachers to teach: Insights from university supervisor feedback. *The Journal of Asia TEFL*, *15*(1), 99–117. https://doi.org/10.18823/asiatefl.2018.15.1.7.99

Gebhard, J. G. (1990). The supervision of second and foreign language teachers. *Eric Digest*, 2–3.

Gholaminejad, R. (2020). When the evil pops in: Exploring the unheard voices of teachers working in private language schools in Iran concerning supervisory observation. *International Journal of Leadership in Education*, 1–24. https://doi.org/10.1080/13603124.2020.1740795

Griffiths, C., & Tajeddin, Z. (Eds.). (2020). *Lessons from good language teachers*. Cambridge University Press. https://doi.org/10.1017/9781108774390

Hallinger, P., Heck, R. H., & Murphy, J. (2014). Teacher evaluation and school improvement: An analysis of the evidence. *Education Assessment, Evaluation and Accountability*, *26*, 5–28. https://doi.org/10.1007/S11092-013-9179-5

Hatamvand, A., Khany, R., & Samaie, M. (2020). Using PLS path modeling for assessing a language teacher observation scale. *International Journal of Applied Linguistics*, *30*(3), 367–389. https://doi.org/10.1111/ijal.12288

Hişmanoğlu, M., & Hişmanoğlu, S. (2010). English language teachers' perceptions of educational supervision in relation to their professional development: A case study of Northern Cyprus. *Novitas-ROYAL (Research on Youth and Language)*, *4*(1), 16–34.

Hsu, L. (2017). Examining EFL teachers' technological pedagogical content knowledge and the adoption of mobile-assisted language learning: A partial least square approach. *Computer Assisted Language Learning*, *29*(8), 1287–1297. https://doi.org/10.1080/09588221.2016.1278024

Isoré, M. (2009). *Teacher evaluation: Current practices in OECD countries and a literature review* (OECD Education Working Paper No. 23). OECD Publishing. https://doi.org/10.1787/223283631428

Khaksefidi, S. (2015). Foreign language teaching in Iran: A model for effective EFL teaching in the Iranian context. *Theory and Practice in Language Studies*, *5*(5), 1060–1071. https://doi.org/10.17507/TPLS.0505.22

Knop, C. K. (1980). The supervision of foreign language teachers. *Teachers College Record*, *81*(6), 186–207. https://doi.org/10.1177/016146818008100611

Marshall, K. (2009). *Rethinking teacher supervision and evaluation: How to work smart, build collaboration, and close the achievement gap* (1st ed.). Jossey-Bass.

Marzano, R. J., & Toth, M. D. (2013). *Teacher evaluation that makes a difference: A new model for teacher growth and student achievement*. ASCD.

Mashhadlou, H., & Izadpanah, S. (2021). Assessing Iranian EFL teachers' educational performance based on gender and years of teaching experience. *Language Testing in Asia*, *11*, 1–26. https://doi.org/10.1186/s40468-021-00140-7

Mazandarani, O., & Troudi, S. (2017). Teacher evaluation: What counts as an effective teacher? In S. Hidri & C. Coombe (Eds.), *Evaluation in foreign language education in the Middle East and North Africa* (pp. 3–28). Springer. https://doi.org/10.1007/978-3-319-43234-2_1

Motallebzadeh, K., Hosseinnia, M., & Domskey, J. G. (2017). Peer observation: A key factor to improve Iranian EFL teachers' professional development. *Cogent Education*, *4*(1), 1277456. https://doi.org/10.1080/2331186X.2016.1277456

Mousavi, M. A., Atai, M. R., Babaii, E. (2016). Exploring standards and developing a measure for evaluating

Iranian EFL teachers' professional competence in the private sector. *Iranian Journal of English for Academic Purposes*, *5*(2), 30–59.

Murphey, T., & Yaode, Q. (2010). A China initiative: Portfolio-based teacher development and appraisal. In C. Coombe, M. Al-Hamly, P. Davidson, & S. Troudi (Eds.), *Evaluating teacher effectiveness in ESL/EFL contexts* (pp. 119–133). The University of Michigan Press.

Ostovar-Namaghi, S. A. (2013). Language teachers' perceptions of evaluation criteria in Iran. *The Reading Matrix*, *13*(1), 45–60.

Park, G.-P., & Lee, H.-W. (2006). The characteristics of effective English teachers as perceived by high school teachers and students in Korea. *Asia Pacific Education Review*, *7*, 236–248. https://doi.org/10.1007/BF03031547

Phillips, K., Balan, R., & Manko, T. (2014). Teacher evaluation: Improving the process. *Transformative Dialogues: Teaching & Learning Journal*, *7*(3), 1–22.

Podolsky, A., Kini, T., & Darling-Hammond, L. (2019). Does teaching experience increase teacher effectiveness? *A review of US research. Journal of Professional Capital and Community*, *4*(4), 286–308. https://doi.org/10.1108/JPCC-12-2018-0032

Ponticell, J. A., & Zepeda, S. J. (2004). Confronting well-learned lessons in supervision and evaluation. *NASSP Bulletin*, *88*(639), 43–59. https://doi.org/10.1177/019263650408863905

Ponticell, J. A., Zepeda, S. J., Lanoue, P. D., Haines, J. G., Jimenez, A. M., & Ata, A. (2019). Observation, feedback, and reflection. In S. J. Zepeda & J. A. Ponticell (Eds.), *The Wiley handbook of educational supervision* (pp. 251–279). Wiley Blackwell. https://doi.org/10.1002/9781119128304.CH11

Rachmajanti, S. (2008). Impact of English instruction at the elementary schools on the students' achievement of English at the lower secondary school. *TEFLIN Journal*, *19*(2), 160–185.

Rashidi, N., & Forutan, A. (2015). On the Iranian in-service and pre-service language teachers' perceptions of educational supervision concerning their professional development. *Issues in Language Teaching*, *4*(2), 55–27. https://doi.org/10.22054/ILT.2015.7225

Richards, J. C., & Farrell, T. S. (2005). *Professional development for language teachers: Strategies for teacher learning*. Cambridge University Press. https://doi.org/10.1017/CBO9780511667237

Sadeghi, K., & Richards, J. C. (2015). Teaching spoken English in Iran's private language schools: Issues and options. *English Teaching: Practice & Critique*, *14*(2), 210–234. https://doi.org/10.1108/ETPC-03-2015-0019

Saldaña, J. (2016). *The coding manual for qualitative researchers* (3rd ed.). Sage.

Shishavan, H. B., & Sadeghi, K. (2009). Characteristics of an effective English language teacher as perceived by Iranian teachers and learners of English. *English Language Teaching*, *2*(4), 130–143. https://doi.org/10.5539/ELT.V2N4P130

Tarajová, R., & Metruk, R. (2020). Characteristics of an effective EFL teacher: Perspectives of Slovak EFL students and teachers. *European Journal of Education Studies*, *7*(12), 78–86.

Todorova, R., & Ivanova, I. (2020). Classroom management and good language teachers. In C. Griffiths & Z. Tajeddin (Eds.), *Lessons from good language teachers* (pp. 133–150). Cambridge University Press. https://doi.org/10.1017/9781108774390.014

Vásquez, C. (2004). "Very carefully managed": Advice and suggestions in post-observation meetings. *Linguistics and Education*, *15*(1–2), 33–58. https://doi.org/10.1016/J.LINGED.2004.10.004

Wei, W., & Hui, S. K. F. (2019). Evaluating teacher performance in language learning classes: The gap between students and department administrators. *Teachers and Teaching*, *25*(4), 486–500. https://doi.org/10.1080/13540602.2019.1625763

Yazdanipour, N., & Fakharzadeh, M. (2020). Effective language teachers' characteristics as perceived by English language private institutes' administrators in Iran. *The Qualitative Report*, *25*(8), 3111–3128. https://doi.org/10.46743/2160-3715/2020.4566

Zamani, R., & Ahangari, S. (2016). Characteristics of an effective English language teacher (EELT) as perceived by learners of English. *International Journal of Foreign Language Teaching and Research*, *4*(14), 69–88.

Zarrabi, F., & Brown, J. R. (2017). English language teaching and learning analysis in Iran. *International Journal of Business, Human and Social Sciences*, *8*(10), 3485–3493. https://doi.org/10.5281/zenodo.1340254

## About the Authors

**Azadeh Hassani** has an MA in TEFL from Allameh Tabataba'i University (Iran). Her research interests include second language teaching and learning, computer-assisted language learning, and teacher education.

**Zari Saeedi** received her PhD from the British University of Trinity (UK) and is an Associate Professor at Allameh Tabataba'i University (Iran). She has presented and published papers on various topics, including educational neuro/psycholinguistics and brain-functioning and cognitive language learning.