

DEVELOPMENT AND VALIDATION OF AN EFL TEACHER TEACHING QUALITY SCALE

Jamshid GHASEMI (Ilam University, Iran)

jamshid.ghasemi@aol.com

Abstract

The present study aimed to investigate the factors, features and variables that might affect EFL teachers' quality and lead to their success or failure in their professional practice. Despite the importance of this concept, no validated instrument was found to measure these constructs. To fill the gap, this study developed and validated a scale of EFL teachers' quality. A 57 item questionnaire was developed and administered to 334 EFL teachers. We ran Structural Equation Modeling (SEM) to analyze the collected data and test the path model of the study. The reliability and validity of the questionnaires were estimated using exploratory and confirmatory factor analyses. The results indicated acceptable goodness of fit indices ($\chi^2=1/789$, $p < .001$, CFI = 0.942, TLI = 0.906, RMSEA= 0.058, SRMR =0.045, AIC =16600 and BIC =17040). The findings of this study have implications for researchers, as well as language teachers and other practitioners in the field of education and teaching quality.

Keywords: *Scale Development, Structural Equation Modeling, Teaching Quality, Teacher Quality, Teacher Quality Scale*

1. Introduction

The concept of teaching quality has been described by various researchers since the 1960s (e.g., Carroll, 1963). However, according to Chu (1990), evaluation of teaching quality has gained increasing attention recently. University authorities employ quality teaching as one of the main key criteria for judging quality both at the institutional level and the individual academic level for quality assurance, benchmarking, and individual recognition and reward through promotion (Jasman et al., 2013). There are different definitions for quality which express different views of the individual and society. Therefore, it is not possible to talk about quality as a unitary concept. Considering that an entity may have high quality in relation to one factor but low quality in relation to another, quality must be defined in terms of a range of qualities (Fairbrother, 1996).

Teacher quality has been stressed strongly in the recent literature and teachers' professional knowledge is assumed to be a key factor affecting teaching quality (Abell, 2007) and many studies have been conducted since Shulman's study (1986), the fundamental considerations about teachers' professional knowledge. Recent research has shown that teacher quality can influence students' learning ability in a positive or negative manner (Roman, 2014) and motivation affects what learners pay attention to and how effectively they process it (Pintrich & Schunk, 2002). Teaching and teacher quality have numerous definitions and expressions and are sometimes used together or in isolation (Singh & Sarkar, 2015). The teaching quality of teachers in colleges and universities is basic guarantee for the overall teaching quality of colleges and universities (Dong et al., 2019). The evaluation of teaching quality in a scientific and accurate way can identify the quality of teachers' work and plays a positive role in promoting teachers'

self-improvement, improving the quality of education and teaching, and ensuring the purpose of training quality teachers (Dong et al., 2019). Singh & Sarkar (2015) consider the quality of teaching as a fusion of both teacher characteristics such as inputs (professional qualifications, experience, in service training etc.) as well as what the teacher 'does' and demonstrates in the class regarding the practices, the teacher's attitudes and the content knowledge which promote positive educational outcomes. The focus on the quality of teaching is the general trend of higher education studies (Mosser et al., 2018) and many pieces of research have been conducted considering teaching quality. In the recent decades this concept has been studied in many different fields of education such as language teaching (Uygun, 2013), mathematics (Lazarides & Buchholz, 2019), medical profession (Diaz et al., 2010), engineering (Calvo et al., 2010) and online and distance education (Yang, 2018).

A considerable amount of studies has been conducted so far on evaluation and improvement of teaching quality in different fields of education (Mosser et al., 2018). The concept of teaching quality assessment first came into being in Harvard University in 1920s and flourished in the later period in the popularization of the higher education (Li, 2012), has its main principal: the school serve the students and the goal of it is to realize the personal expectation (Xian et al., 2016). Many colleges generally regard the teaching evaluation as a way to improve the teaching quality (Xian et al, 2016). According to Ome et al. (2017), teacher training constitutes a promising policy area for improving the quality of educational system. Maslow & Kelley (2012) studied the ways that teachers and schools use information from teacher evaluations to advance teaching practice through formative and systemic feedback that support research by Goldstein (2004, 2006) which suggest that teacher evaluation designs that include peer evaluation

and feedback can encourage collaboration and promote learning for some motivated teachers.

Students' perception and evaluation of teaching quality in higher education has been also taken care of in the recent research. Üstünlüoğlu (2017) performed a case study considering teaching quality in higher education in Turkey and Slovakia. Her study aimed to investigate the perceptions of both students and lecturers on teaching in higher education. The results indicated a difference between students' and lecturers' perceptions highlighting a discrepancy over views on the pedagogical competence of lecturers. Spooren et al. (2007) developed a theory-based and thoroughly validated evaluation instrument. The results of their study underline the value of the use of a scaling technique in students' evaluation of teacher performance. Dunrong & Fan (2009) with the intention of improvement of the teaching quality conducted a study with aim to build up a scientific system of student evaluation of teaching, and draw on advanced experiences from abroad so as to improve the system of student evaluation of teaching and perfect the teaching quality assurance system.

1.1. Purpose And Importance Of The Research

Though, a large number of studies have been conducted on teaching quality, yet there are a number of unresulted issues in the field and the lack of such studies has resulted in a vague picture of research on this issue. Also, researchers' methodology and content selection are full of replication and parallel studies. Accordingly, these issues have made the field disorderly and obscurant. Hence, this study intends to develop and validate a scale of EFL teacher quality. The results of this study may have contributions to the theory and the practice of language teaching and particularly to the EFL teaching and can provide educators with a large landscape of the area of EFL teaching and teach-

ers' assessment and qualification. This study also helps EFL teachers to be more aware of their knowledge, skills and behaviors and their effect on EFL learners' language learning and this awareness can assist learners to experience more pleasurable language learning environments. Therefore, the underlying research question of this study were as follows:

What are the constructs and components of a teacher quality scale in an Iranian EFL context??

2. Methods

2.1 Participants

The current study encompassed 334 EFL teachers (166 females and 168 males) from Ilam, Khuzestan, Kermanshah, Kohgiluyeh and Boyer-Ahmad, and Hamedan, Iran. The teachers with educational levels of B.A, M.A and Ph. D in Literature and Teaching English as a Foreign Language who teach English to the male and female students were participants of this study. Participants of the study were selected based on the stratified sampling technique in which different smaller parts of the population, namely strata, were selected to participate in the project.

2.2 Instrument

Review of the recent research resulted in the collection of a number of definitions, constructs, and scales from which our construct table of teacher quality was developed. According to these definitions and constructs, we generated items which were factors that could assess and monitor teachers' quality. In order for the content validity of the scale, "Peer-reviewed" method was chosen and a group of 6 experts were asked humbly to investigate carefully the items to present evidence of the EFL TQS content validity. They rated the appropriateness of the items influencing teacher quality on a three-point

scale (1=suitable, 2=marginally suitable, 3=very suitable) and classified them into possible categories. Then, their ratings were analyzed to ensure which items to remain in the scale. Based on to the reviewers' comments, the researchers rephrased the items with ratings under 3. In order to apply the next validation (i.e. exploratory & confirmatory factor analyses) of the scale, several assertions for the explanation of the items were provided in the form of a Likert-type questionnaire. Afterwards, Cronbach's Alpha was calculated to evaluate the reliability of the scale. Its reliability value was 99.4% which indicated a high level of internal consistency.

2.3 Construct Description of TQS

In order to answer the question of this study, through studying the related literature, we developed a model based on existing definitions of teaching quality and also offered models and scales of teacher quality measurement in the literature. The developed model is a scale of teacher quality assessment for EFL teachers. Studying the related literature led us to the identification of the main sources that influence EFL teacher quality. These main dimensions were incorporated into the TQS. In addition, these aspects and their subscales were identified and confirmed in the content validity phase. These aspects encompassed in the TQS are enumerated as follows:

TPACK which is a framework that introduces the relationships and also the complexities between all three basic components of knowledge (technology, pedagogy, and content) (Mishra & Koehler, 2006). There is an intuitive understanding of teaching content with appropriate pedagogical methods and technologies at the intersection of these three types of knowledge. Seven components are included in the TPACK framework. Schmidt-Crawford et al. (2009) defined them as:

1. *Technology knowledge (TK)*: Technology knowledge does refer to the knowledge about different types of technologies, ranging from low-tech technologies such as pencil and paper to digital technologies such as the Internet, digital video, interactive whiteboards, and different software programs.

2. *Content knowledge (CK)*: Content knowledge is defined as the “knowledge about actual subject matter that is to be learned or taught” (Mishra & Koehler, 2006, p. 1026). Teachers must have nobility about the teaching content and the nature of knowledge for different content areas.

3. *Pedagogical knowledge (PK)*: Pedagogical knowledge refers to the familiarity with the methods and processes of teaching which includes knowledge in classroom management, assessment, lesson plan development, and student learning. In other word it is the knowledge of how to teach that is applicable across a range of teaching areas.

4. *Pedagogical content knowledge (PCK)*: Pedagogical content knowledge refers to the content knowledge that deals with the teaching process. Pedagogical content knowledge blends both content and pedagogy with the goal being to develop better teaching practices in the content areas so *PCK* is different for various content areas.

5. *Technological content knowledge (TCK)*: Technological content knowledge has to do with the knowledge of how technology can create new representations for specific content. *TCK* suggests that teachers understand that, by using a specific technology, they can change the way learners practice and understand concepts in a specific content area.

6. *Technological pedagogical knowledge (TPK)*: Technological pedagogical knowledge is being knowledgeable about different practical technologies in teaching, and the understanding of how using technology may change the way teachers teach.

7. *Technological pedagogical content knowledge (TPACK)*: Technological pedagogical content knowledge refers to the body of knowledge required by a teacher for integrating various technologies into their teaching in any content area. Teachers have an intuitive understanding of the complex interplay between the three basic components of knowledge (CK, PK, TK) by teaching content using appropriate pedagogical methods and technologies.

Big-Five personality trait model was originally based on a combination of the lexical and the statistical approach (Larsen et al., 2017). This has been widely influenced and accepted hierarchical model of personality structure (Djigic et al., 2014). According to this model, personality may be described with following five basic dimensions which represent broad domains of personality: neuroticism, extraversion, openness, agreeableness and conscientiousness. Neuroticism differentiates between persons in regard to emotional stability-emotional instability. It is a disposition of a person to experience negative emotions such as: sadness, anxiety, fear, wrath, guilt. Extraversion stands for sociability and activity. Persons with high scores are talkative and friendly, cheerful, active, optimistic and outgoing, full of energy. Introverts are reserved, closed, more independent and sensitive in his nature. Openness is related to intellectual curiosity, preference of diversity, a need for a change and tendency towards experimenting, inclination to new ideas and non-conventional values. A person with higher score in this part tends to be more open-minded, to question authorities and dogmas, is liberal and open to novelty. Agreeableness stands for altruism, trust, and compassion for others. Persons

with low score seem to to be selfish, cynical, suspicious about other's people's intentions, egocentric and competitive, while high score shows a tendency to be cooperative, altruistic and empathetic. Conscientiousness represents the ability of self-control in a sense of a disciplined inclination towards goals and duties, strict holding on one's own principles. So, this dimension is connected to academic and professional success.

Different professional Skills consisting Leadership Skills, Critical & Reflective Thinking Skills, Communicative & Verbal Skills, Creativity Skills and Mindfulness; which is comprised of four general domains including (1) planning and organization; (2) teaching mindfulness; (3) guiding mindfulness practices; and (4) management of the learning environment (Broderick et al., 2018).

Researchers used the following categories for teacher quality characteristics identified by Bolyard and Moyer-Packenham (2008); Darling-Hammond (2000); Ren (2009); Boyd et al (2011): (a) teacher behaviors, practices, and beliefs; (b) Certification status; (c) Experience; (d) Preparation; and (e) Ethical principles. The category *teacher behaviors, practices, and beliefs* included what the teacher does in the classroom, for example, questioning strategies, instructional equity, classroom management, beliefs about students' learning, such as beliefs about the way students learn content and beliefs about who can and cannot learn. *Certification* describes teachers' certification status (including whether they are emergency, provisionally, or fully certified) and whether a teacher is certified in the field in which they are teaching. *Experience* is defined as the total number of years an educator has been teaching and/or the number of years a teacher has taught a particular grade level or field of study. The morality of teachers which is considered as an important aspect of teacher quality, in the form of teacher *professional*

Ghasemi, J. (2023) **Development and Validation of an EFL Teacher Teaching Quality Scale**
International Journal of Quality in Education

ethics (Ren, 2009). Also Boyd et al. (2011) and Darling-Hammond (2000) identified *preparation* as a key characteristic of quality teachers.

2.4 Procedure of Data Collection

After finalizing the questionnaires, they were distributed among EFL teachers and students in Ilam, Hamedan, Kermanshah, Khuzestan and Kohgiluyeh, Iran. Under the influence of health care protocols and limited contact and communication we had to distribute and collect the questionnaires only through online forms. Participants were in contacted through different ways with the researchers and it was explained how to complete the questionnaires. Finally, it took two months to distribute and collect the questionnaires. From distributed questionnaires, only completely filled questionnaires were downloadable and they formed the basis of data analysis using Statistical Package for Social Science (SPSS) and jamovi.

2.5 Data Analysis

In order to analyze the collected data sets, descriptive and inferential statistics were utilized. Descriptive statistics were performed to calculate frequency and percentage of each section of the questionnaires. Our data analysis included two phases: exploratory factor analysis and confirmatory factor analysis. We ran SPSS version 26, Amos version 20 and jamovi softwares for analyzing our data. The first phase, the exploratory factor analysis, was used to determine the appropriate number of variables in a study through using some statistical. The second phase, confirmatory factor analysis, through the investigation of all the associations among main scales and their sub-scales, tries to confirm or validate the model using goodness of fit indices. The statistics used to examine the model fit for exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) were chi-square statistic, Comparative Fit Index (CFI), Tucker Lewis index

(TLI) and the Root Mean Square Error of Approximation (RMSEA), as well as Standardized Root Mean Square Residual (SRMR), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC). In general, Chi-squared statistic of less than 3, with CFI and TLI greater than .90 and RMSEA and SRMR of Less than .06 and 0.8 respectively are considered as an acceptable model fit (Hu & Bentler, 1999). The EFA descriptive analysis of EFL teachers' quality was carried out with SPSS version 26, whereas CFA and model evaluation was conducted using jamovi software version 1.2.27 and Amos version 20.

3. RESULTS AND DISCUSSION

3.1. Results of Exploratory Factor Analysis

In this study, at first, EFA based on Principal Component Analysis (PCA) with Promax rotation was performed on 62 items using SPSS version 26. Items loaded heavily on more than one factor. If an item's highest factor loading was greater than an a priori determined cutoff value, the item was retained until reaching the desirable result. Also setting a cutoff at 0.40, this level resulted in the removal of 5 items in our sample loaded. Finally, we came up with 57 items. To address the issue of suitability of the data, the strength of the inter-correlations among the items, we checked the Kaiser-Meyer-Olkin Measure of sampling adequacy (KMO). This value should be over 0.6. Bartlett's Test of Sphericity value should also be considered that is the Sig value should be .05 or smaller. As it is clear from table 4.8 in this study our sample is appropriate for factor analysis because the KMO value is 0.709 (it should be above 0.6) and Bartlett's Test of Sphericity (chi-square = 7660.331, df = 990, p = .000) were factorable (p < 0.05).

Table 1. KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.709
Bartlett's Test of Sphericity	Approx. Chi-Square	10372.136
	df	1653
	Sig.	.000

As is shown in table 2, four factors were identified as underlying latent constructs from 57 items based on parallel analysis, accounted for 51.91% of the total variance in the data. These factors included the 4 dimensions used to measure teacher quality.

Table 2. Total Variance Explained

#	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings
	Total	% of Variance	Cumulative%	Total	% of Variance	Cumulative %	Total
1	14.042	24.210	24.210	14.04	24.210	24.210	9.828
2	9.381	16.175	40.384	9.381	16.175	40.384	9.140
3	3.505	6.042	46.426	3.505	6.042	46.426	8.229
4	3.186	5.493	51.919	3.186	5.493	51.919	7.726

3.2. Results of Confirmatory Factor Analysis

In the next step, the Confirmatory Factor Analysis (CFA) was conducted using jamovi software version 1.2.27. The maximum likelihood method was used to estimate the parameter. The results of the first CFA analyses indicated a relatively adequate good model fit. The results indicated an overall good model fit; Chi-squared=1/789, $p < .001$,

CFI = 0.942, TLI = 0.906, RMSEA= 0.058, SRMR =0.045, AIC =16600 and BIC =17040.

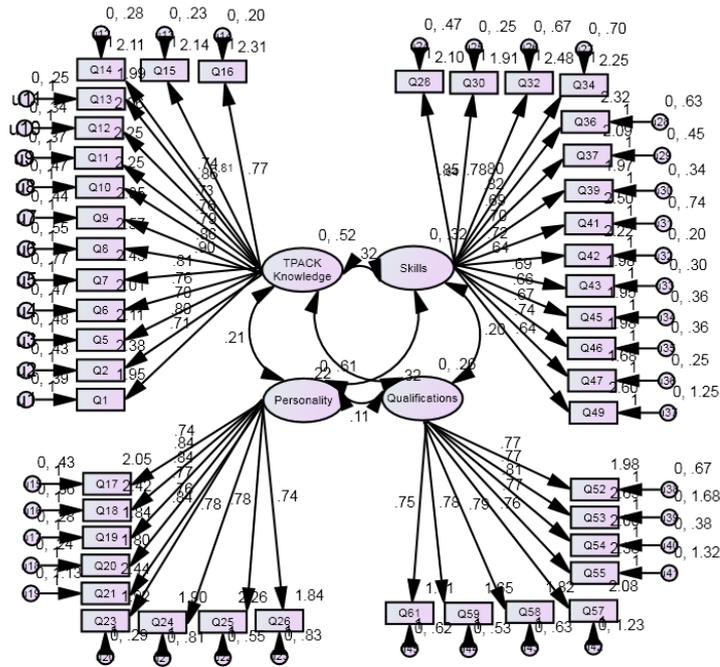


Figure 1 Fitted CFA model

Table 3. Component Correlation Matrix

C	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1.000													
2	.312	1.000												
3	-.167	.260	1.000											
4	-.115	.240	.485	1.000										
5	.013	.199	.264	.275	1.000									
6	.063	.134	-.046	.145	.163	1.000								
7	-.001	.049	.151	.022	-.160	-.030	1.000							
8	.365	.285	.090	.019	.113	.175	.008	1.000						
9	.442	.292	.042	.197	.029	.061	.073	.190	1.000					

10	.251	.168	.050	.080	.119	.298	-.053	.067	.256	1.000				
11	.279	.119	-.259	-.272	-.033	.135	.144	.328	.281	.202	1.000			
12	-.072	-.112	.062	-.123	.014	-.239	-.023	-.219	-.136	-.166	-.018	1.000		
13	.003	-.171	-.054	-.080	-.179	-.153	.116	-.117	-.170	-.120	-.186	.070	1.000	
14	.285	.336	.125	.153	.241	-.032	-.182	.108	.106	.250	-.014	.115	-.092	1.000

Table 4. Model Fit Measures

CFI	TLI	SRMR	RMSEA	RMSEA 90% CI		AIC	BIC
				Lower	Upper		
0.942	0.906	0.045	0.058	0.073	0.082	16600	17040

Table 6. Factor Covariance

		Estimate	SE	95% Confidence Interval		Z	p	Stand. Estimate
				Lower	Upper			
TPACK Knowledge	TPACK Knowledge	1.0000 ^a						
	Personality	0.3063	0.0773	0.155	0.4578	3.961	< .001	0.3063
	Skills	0.6407	0.0587	0.526	0.7558	10.914	< .001	0.6407
	Qualifications	0.0910	0.1022	-0.109	0.2914	0.890	0.374	0.0910
Personality	Personality	1.0000 ^a						
	Skills	0.7120	0.0527	0.609	0.8153	13.511	< .001	0.7120
	Qualifications	-0.6894	0.0634	-0.814	-0.5651	-10.871	< .001	-0.6894
Skills	Skills	1.0000 ^a						
	Qualifications	-0.2583	0.1015	-0.457	-0.0594	-2.545	0.011	-0.2583
Qualifications	Qualifications	1.000 ^a						

^a fixed parameter

3.6. Results of Reliability Analysis of TQS

Using jamovi for the reliability analysis of our items, which is indicated in the table 6, we have the reliability calculated for the questionnaire with a satisfactory Cronbach's Alpha of .923 and McDonald's Omega of 0.934. As we already know, a Cronbach's Alpha of more than 0.70 and as a general guideline, for McDonald's Omega, threshold value of 0.70 is for research purpose and 0.90 for clinical or important decisions. Therefore they are acceptable and considered as reliable.

Table 6. Scale Reliability Statistics

	sd	Cronbach's α	McDonald's ω
scale	0.423	0.923	0.934

This paper described the development and validation of a reliable instrument for measuring teacher quality in an EFL context. The scale was designed to be used with classroom teachers in an EFL contexts. While some aspects of the scale are naturally specific to the EFL contexts, we believe that this system could be effectively used with slight modifications with other educational contexts as well. It should be stated that the behaviors and attitudes are not mutually exclusive entities, and most of the key features contain both behavioral and attitudinal elements. However, some key features are more behavioral and concrete than others.

To fulfill our goal, we constructed a model consisting of exploratory and confirmatory analyses. This model was actually employed to test the construct validity of proposed four factors, i.e. knowledge, personality, skills and qualifications. As was mentioned earlier, the hypothetical model was developed based on a comprehensive review

of the related literature pertinent to teacher quality and was then tested on a sample of 334 EFL teachers while using EFA, CFA and Model Evaluation estimates. Despite the overall data-to-model fit observed in this paper resulting in the removal of 5 items. Although all of the four initially proposed components in the instrument were substantiated by the gathered data, 5 of the items did not statistically load during exploratory data analysis phase reducing the inventory to 57 items. All the remaining 57 items tapping into knowledge, personality, skills and qualifications were found to have significant statistical relationships with their corresponding factors. The calculated model-fit estimates also verified this CFA model as a valid measure of change. Specifically, Items 22, 38,44, 48 and 56 did not load. Although further research is needed for investigating why these 5 items were discarded in the exploratory analyses, some of them can be accounted for by the context of the present study within which our teacher participants teach.

Results suggested that ELT teacher Quality Scale had a reasonable factor structure and a desirable convergent validity. The results of the EFA showed that the instrument had a rescannable degree of factorial validity based on the good model fit and factor pattern loadings. Generally, the results showed large factor pattern loadings. The result of CFA also showed good factorial validity. Each item was highly correlated with its corresponding factor and not with the other factors. All the estimated factor loadings were greater than 0.40. Accordingly, the results of EFA and CFA confirmed the model of the ELT Teacher Quality in the Iranian EFL context. So, one of the strengths of the teacher quality instrument developed in the present study is that it reflects an Iranian educational context. This study also helps to better understand the performance of high school teachers in terms of effective teaching practices in the classroom and also tries to

add to the base of knowledge that could be used by teacher training institutions. The findings can also assist in the development of curricula that will improve the effectiveness of teaching practices of teachers to increase the quality of teaching and learning in secondary schools. The results can also be used as a basis for measuring the impact of career and professional development programmes.

The construction of a valid and reliable scale requires systematic research, in which both the literature and empirical data play an important role. This type of preliminary research does not yet seem to be popular (Spooren, 2007). This is the strength of our scale and ensures those who are concerned about the reliability, validity and thus the usefulness of this teacher teaching quality scale. The process of validating an assessment instrument is a never-ending task. So, future research will need to be performed in order to establish the validity of the teaching quality assessment questionnaire we constructed and studied.

Recommendations

- ❖ The study was conducted in five provinces namely, Ilam, Kohgiluyeh, Khuzestan, Kermanshah and Hamedan, Iran. Therefore, the sample size may limit the generalizability of our results. With regard to the pandemic and under the influence of health care protocols and limited contact and communication we had to distribute and collect the questionnaires just through online forms.
- ❖ Also considering the large number of questions in the questionnaire, some teachers may tend to move through the questionnaire too quickly and carelessly in order to complete the questionnaires faster.
- ❖ To compensate the limitations of this study, similar studies can be conducted in more provinces, or investigate how other factors like age, gender, and experience

Ghasemi, J. (2023) **Development and Validation of an EFL Teacher Teaching Quality Scale**
International Journal of Quality in Education

level influence teachers' quality. Also it would be a valuable topic for future researchers to investigate the potential relationship between the variables of the current study with students' achievement and leaning.

References

Abell, S. (2007). Research on science teachers' knowledge. In S. Abell & N. Lederman (Eds.), *Handbook of research on science education* (pp. 1105–1149). Mahwah, NJ: Lawrence Erlbaum Associates.

Boyd, D., Grossman, P., Ing, M., Lankford, H., Loeb, S., & Wyckoff, J. (2011). The Influence of School Administrators on Teacher Retention Decisions. *American Educational Research Journal*, 48(2), 303–333.
<https://doi.org/10.3102/0002831210380788>

Broderick, P., Frank, J., Berrena, E., Schussler, D., Kohler, K., Mitra, J., Khan, L., Levitan, J., Mahfouz, J., Shields, L. & Greenberg, M. (2018). Evaluating the Quality of Mindfulness Instruction Delivered in School Settings: Development and Validation of a Teacher Quality Observational Rating Scale. *Mindfulness*, 10.1007/s12671-018-0944-x.

Carroll, J. (1963). A model of school learning. *Teachers College Record*, 64(8), 723–733.

- Ghasemi, J. (2023) **Development and Validation of an EFL Teacher Teaching Quality Scale** *International Journal of Quality in Education*
- Chu, F. (1990). Quantitative evaluation of university teaching quality — an application of fuzzy set and approximate reasoning. *Fuzzy Sets and Systems*, 37(1), 1-11. doi: 10.1016/0165-0114(90)90058-e
- Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives*, 8(1). Retrieved from <http://epaa.asu.edu/epaa/v8n1/>.
- Diaz, V., Carek, P., Dickerson, L., & Steyer, T. (2010). Teaching Quality Improvement in a Primary Care Residency. *The Joint Commission Journal On Quality and Patient Safety*, 36(10), 454-AP5. doi: 10.1016/s1553-7250(10)36067-3.
- Djigic, G., Stojiljković, S. & Dosković, M. (2014). Basic Personality Dimensions and Teachers' Self-efficacy. *Procedia - Social and Behavioral Sciences*, 112. 593-602. 10.1016/j.sbspro.2014.01.1206.
- Dong, Q., Wang, S., Han, F. & Zhang, R. (2019). Innovative Research and Practice of Teachers' Teaching Quality Evaluation under the Guidance of 'Innovation and Entrepreneurship'. *Procedia Computer Science*, 154. 770-776. 10.1016/j.procs.2019.06.123.
- Dunrong, B. & Fan, M. (2009). On Student Evaluation of Teaching and Improvement of the Teaching Quality Assurance System at Higher Education Institutions. *Chinese Education & Society*, 42. 100-115. 10.2753/CED1061-1932420212.
- Fairbrother, P. (1996). Recognition and assessment of teaching quality. *Nurse education today*, 16 1, 69-74.

- Ghasemi, J. (2023) **Development and Validation of an EFL Teacher Teaching Quality Scale** *International Journal of Quality in Education*
- Goldstein, J. (2004). Making sense of distributed leadership. *Educational Evaluation & Policy Analysis*, 26(2), 173–197.
- Goldstein, J., & Noguera, P. (2006). A thoughtful approach to teacher evaluation. *Educational Leadership*, 63(6), 31–37.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1e55
- Jasman, A., Blass, E. & Shelley, S. (2013). Becoming an Academic for the Twenty-First Century: What Will Count as Teaching Quality in Higher Education. *Policy Futures in Education*, 11. 660. 10.2304/pfie.2013.11.6.660.
- Larsen, R. J., Buss, D. M., Wismeijer, A., Song, J., & van den Berg, S. M. (2017). *Personality psychology: Domains of knowledge about human nature*. (2nd ed.) McGraw Hill Education.
- Lazarides, R., & Buchholz, J. (2019). Student-perceived teaching quality: How is it related to different achievement emotions in mathematics classrooms?. *Learning and Instruction*, 61, 45-59. doi: 10.1016/j.learninstruc.2019.01.001
- Li, X. F. (2012). *The ideal and reality of students' evaluation: The study based on the fourth generation evaluation theory[D]*. Beijing: Beijing Normal University.
- Maslow, V. & Kelley, C. (2012). Does Evaluation Advance Teaching Practice? The Effects of Performance Evaluation on Teaching Quality and System Change in

Ghasemi, J. (2023) **Development and Validation of an EFL Teacher Teaching Quality Scale**
International Journal of Quality in Education

Large Diverse High Schools. *Journal of School Leadership*, 22. 600-632.
10.1177/105268461202200307.

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for integrating technology in teachers' knowledge. *Teachers College Record*, 108(6), 1017–1054.

Mosser, G., Frisch, K., Skarda, P. & Gertner, E. (2009). Addressing the Challenges in Teaching Quality Improvement. *The American journal of medicine*, 122. 487-91. 10.1016/j.amjmed.2009.01.013.

Moyer-Packenham, P. S., Bolyard, J. J., Kitsantas, A., & Oh, H. (2008). The assessment of mathematics and science teacher quality. *Peabody Journal of Education*, 83(4), 562–591.

Ome, A., Menendez, A. & Le, H. (2017). Improving teaching quality through training: Evidence from the Caucasus. *Economics of Education Review*, 61. 10.1016/j.econedurev.2017.09.003.

Pintrich, P. R. & Schunk, D. H. (2002). *Motivation in education: theory, research, and applications*. Merrill, Prentice-Hall International.

Ren, W. (2009). Reinforcement on Teachers' Morality Construction --- An Eternal Subject in Educational Development. *Asian Culture and History*, 1. 10.5539/achv1n2p180.

Ghasemi, J. (2023) **Development and Validation of an EFL Teacher Teaching Quality Scale** *International Journal of Quality in Education*

Roman, I. (2014). Qualitative Methods for Determining Students' Satisfaction with Teaching Quality. *Procedia - Social and Behavioral Sciences*, 149 (2014) 825 – 830

Schmidt-Crawford, D., Baran, E., Thompson, A., Mishra, P., Koehler, M. & Seob, S. (2009). Technological Pedagogical Content Knowledge (TPACK): the development and validation of an assessment instrument for preservice Teachers. *Journal of Research on Technology in Education*, 42. 123-149. 10.1080/15391523.2009.10782544.

Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14.

Singh, R. & Sarkar, S. (2015). Does teaching quality matter? Students learning outcome related to teaching quality in public and private primary schools in India. *International Journal of Educational Development*, 41 (2015) 153–163

Spooren, P., Mortelmans, D. & Denekens, J. (2007). Student evaluation of teaching quality in higher education: development of an instrument based on 10 Likert-scales. *Assessment & Evaluation in Higher Education*, 32:6, 667-679, DOI: 10.1080/02602930601117191

Üstünlüoğlu, E. (2017). Teaching quality matters in higher education: a case study from Turkey and Slovakia. *Teachers and Teaching*, 23, 367 - 382.

Uygun, S. (2013). English Language Teaching Quality Criteria in Turkish Secondary Education (based on EAQUALS' Staff Charter). *Procedia - Social and Behavioral Sciences*, 70. 1354-1359. 10.1016/j.sbspro.2013.01.197.

Ghasemi, J. (2023) **Development and Validation of an EFL Teacher Teaching Quality Scale**
International Journal of Quality in Education

Xian, S., Xia, H., Yin, Y., Zhai, Z. & Shang, Y. (2016). Principal component clustering approach to teaching quality discriminant analysis. *Cogent Education*, 3:1, DOI: 10.1080/2331186X.2016.1194553

Yang, S. (2018). Construction research on index system of teaching quality of distance education. *Journal of Discrete Mathematical Sciences and Cryptography*, 21:6, 1431-1436, DOI: 10.1080/09720529.2018.1527489