AN EVALUATION OF EFL TEACHER ASSESSMENT AND EVALUATION:

A SUGGESTED MODEL

Sabahattin Yeşilçınar

Ph.D. DISSERTATION

DEPARTMENT OF FOREIGN LANGUAGE TEACHING

GAZİ UNIVERSITY

INSTITUTE OF EDUCATIONAL SCIENCES

SEPTEMBER, 2018

TELİF HAKKI VE TEZ FOTOKOPİ İZİN FORMU

YAZARIN

Adı: Sabahattin

Soyadı: YEŞİLÇINAR

Bölümü: İngilizce Öğretmenliği

İmza:

Teslim tarihi:

TEZİN

Türkçe Adı: Yabancı Dil Olarak İngilizce Öğretmeni Seçme ve Atamasının Değerlendirilmesi: Bir Model Önerisi

İngilizce Adı: An Evaluation of EFL Teacher Assessment and Evaluation: A Suggested Model

ETİK İLKELERE UYGUNLUK BEYANI

Tez yazma sürecinde bilimsel ve etik ilkelere uyduğumu, yararlandığım tüm kaynakları kaynak gösterme ilkelerine uygun olarak kaynakçada belirttiğimi ve bu bölümler dışındaki tüm ifadelerin şahsıma ait olduğunu beyan ederim.

Yazar Adı Soyadı: Sabahattin YEŞİLÇINAR

İmza: ………………………

JÜRİ ONAY SAYFASI

Sabahattin YEŞİLÇINAR tarafından hazırlanan "An Evaluation of EFL Teacher Assessment and Evaluation: A Suggested Model" adlı tez çalışması aşağıdaki jüri tarafından oy birliği/oy çokluğu ile Gazi Üniversitesi Eğitim Bilimleri Enstitüsü Yabancı Diller Eğitimi Anabilim Dalı'nda Doktora tezi olarak kabul edilmiştir.

Danışman: (Prof. Dr. Abdulvahit ÇAKIR)

(İngiliz Dili Eğitimi, Gazi Üniversitesi) …………………

Başkan: (Prof. Dr. Gülsev PAKKAN)

(İngiliz Dili ve Edebiyatı, Selçuk Üniversitesi) …………………

Üye: (Prof. Dr. Paşa Tevfik CEPHE)

(İngiliz Dili Eğitimi, Gazi Üniversitesi) …………………

Üye: (Doç. Dr. Gonca Yangın EKŞİ)

(İngiliz Dili Eğitimi, Gazi Üniversitesi) …………………
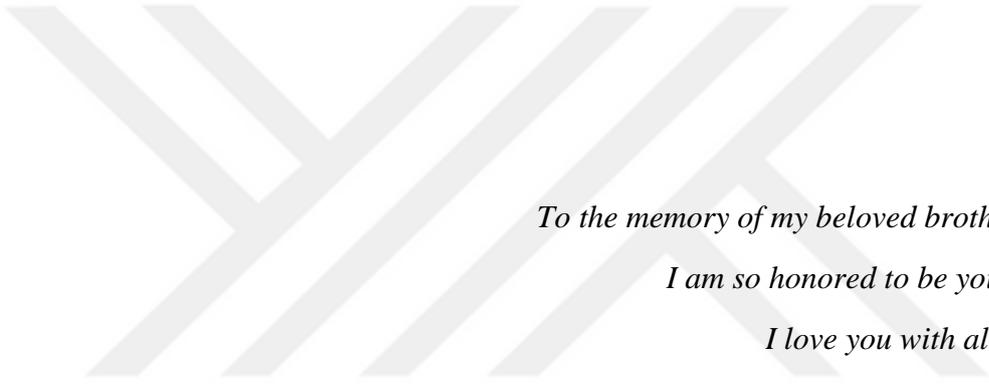
Üye: (Dr. Öğr. Üyesi Neslihan ÖZKAN)

(İngiliz Dili Eğitimi, Ufuk Üniversitesi) …………………

Tez Savunma Tarihi: .. /… / …

Bu tezin Eğitim Bilimleri Enstitüsü Yabancı Diller Eğitimi Anabilim Dalı' nda Doktora tezi olması için şartları yerine getirdiğini onaylıyorum.

Prof. Dr. Selma YEL

Eğitim Bilimleri Enstitüsü Müdürü …………………

To the memory of my beloved brother, Fehmi

I am so honored to be your brother

I love you with all my heart

ACKNOWLEDGEMENTS

YABANCI DİL OLARAK İNGİLİZCE ÖĞRETMENİ SEÇME VE ATAMA MODELİNİN DEĞERLENDİRİLMESİ: BİR MODEL ÖNERİSİ

(DOKTORA TEZİ)

Sabahattin YEŞİLÇINAR

GAZİ ÜNİVERSİTESİ

EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Eylül, 2018

ÖZ

Öğrencilerin başarısını etkileyen en önemli etkenlerden bir tanesi öğretmen niteliğidir. Nitelikli öğretmen seçebilmenin ön koşulu da güvenilir, geçerli ve adil bir ölçme ve değerlenme sisteminin olmasıdır. Çoğu ülkelerde önemli kararların alınmasında standartlaştırılmış sınavlar kullanılmaktadır. Bu sınavlardan etkilenenlerin düşüncelerine başvurmak sınavların sonuçsal geçerliliğini belirlemek açısından son derece önemlidir. Bu çalışmanın amacı, mevcut atama modelinin geçerliliğini sorgulayıp gerekli bilgi, beceri ve tutumlara sahip İngilizce öğretmeni seçip bu nitelikleri taşımayan adayları eleyecek bir atama modeli önermektir. Karma araştırma desenine sahip bu çalışmanın nitel boyutunu 78 paydaş oluşturmakta. Bu katılımcılar arasında öğretmen adayı (31), İngilizce öğretmeni (25), öğretmen yetiştiriciler (9), eğitim uzmanı (8) ve kanun koyucular (5) bulunmaktadır. Türkiye'de eğitim fakültesi dışında mezun olanlar formasyon almak suretiyle öğretmen olabildiği için, hem öğretmen adayları hem de hizmet içi öğretmen örnekleminde formasyonla atanmış katılımcılara yer verilmiştir. Nitel verilerin elde edilmesinde yarı yapılandırılmış görüşme, odak grup ve doküman inceleme yöntemi kullanılmıştır. Nicel veriler için 271 İngilizce öğretmeninden faydalanılmıştır. Nicel verilerin elde edilmesinde

araştırmacı tarafından geliştirilen güvenilir ve geçerliği kanıtlanmış İngilizce Öğretmenlerinin Atamaya Yönelik Tutum Ölçeği kullanılmıştır. Nicel veri analizinde SPSS kullanılmıştır. Kolmogorov-Smirnov normallik testine göre dağılımın normal olduğu görülmüştür. İki gruptan oluşan değişkenler için t-test; üç ve daha fazla değişkenden oluşanlar için ise ANOVA kullanılmıştır. Nitel veri için içerik analizi kullanılmıştır. Kod ve kategorilerin oluşumu öncelikle araştırmacılar tarafından manuel yapılmış; daha sonra veri kaybını önlemek için ise NVivo 12 programı kullanılmıştır. Nicel veriler İngilizce öğretmenlerinin atama modeline yönelik olumsuz tutum takındıklarını göstermiştir. Nitel veriler nicel verileri doğrulamakta ve mevcut atama modelinin ölçmek istediği yapıyı temsil etmediğini ve puanlama açısından da dil ve edebiyat mezunlarının lehine olduğu, dolayısıyla adil ve geçerli sayılamadığını göstermektedir. Bu bağlamda, öğretmen adayının bilgi, beceri ve tutumlarını da değerlendirebilecek yeni bir atama modeli geliştirilmiştir. Bu modelin nitelikli öğretmen seçimine ve dolayısıyla ülkenin eğitimine pozitif katkı sağlayacağı düşünülmektedir.

AN EVALUATION OF EFL TEACHER ASSESSMENT AND EVALUATION: A SUGGESTED MODEL

(Ph.D. Dissertation)

Sabahattin YEŞİLÇINAR

GAZI UNIVERSITY

INSTITUTE OF EDUCATIONAL SCIENCES

September, 2018

ABSTRACT

The quality of a teacher is determinative of student achievement, and the precondition for choosing a qualified teacher is only possible if there is a reliable, valid and fair teacher assessment and evaluation (TAE) model. Most countries use standardized exams for high-stakes decisions. Stakeholders' perception of testing is informative in terms of consequential validity of exams. In addition to examining the validity of the current TAE model, the purpose of this study is to develop an English language teacher assessment and evaluation model that can distinguish effective student teachers (who have required knowledge, skills, and attitudes) from ineffective ones. In this mixed-method research, the qualitative part consists of student teachers (n=31), in-service English language teachers (n=25), teacher trainers (n=9), education experts (n=8), and policymakers (n=5). The sample of student teachers and in-service teachers involve both graduates of ELT and other departments (such as English Language and Literature, American Culture and Literature, Department of Translation and Interpreting or Department of English Linguistics graduating from the Faculties of Science and Letters) since they have right to become English language teachers via pedagogic formation certificate. Semi-structured and focus group interview techniques were used for quantitative data. Regarding quantitative data, 271 in-service English language teachers were given the English Teachers' Attitudes

towards Recruitment System (ETARS), a valid and reliable scale developed by the researcher. Statistical Package for Social Sciences (SPSS) was used for quantitative data. According to the Kolmogorov-Smirnov normality test, the distribution was found to be normal. T-test (Independent Samples Test) was used for the demographic variables consisting of two categories, while one-way ANOVA tests were applied for the demographic variables consisting of 3 and more categories. Content analysis is used for qualitative data. The formation of codes and categories was made primarily by the researcher; then the NVivo 12 program was used to prevent data loss. The quantitative data showed that English teachers had a negative attitude towards the current TAE model. Qualitative findings confirmed quantitative data and indicated that the current TAE model is invalid and unfair as it has both construct underrepresented and construct-irrelevant variance (favoring graduates of English Language and Literature in terms of scoring). In this context, a new TAE model, which can evaluate student teachers' knowledge, skills, and attitudes, has been developed. This model is considered to be helpful in selecting qualified teachers and, therefore, increasing the quality of the education of the country.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CEFR	Common European Framework of Reference for Languages
CEO	Abbriation used for 'Competencies, 'Effective Teaching', and 'Openness to Personal and Profesional Development'
CLT	Critical Language Testing
EE	Education Expert
EFL	English as a Foreign Language
ELT	English Language Teaching
ELTAE	English Language Teacher Assessment and Evaluation
ELTE	English Language Teacher Education
EPOSTL	European Portfolio for Student Teachers of Languages
ESL	English as a Second Language
ETARS	English Teachers' Attitudes towards Recruitment System
IELTS	International English Language Testing System
ITT	Initial Teacher Training
KPSS	Kamu Personeli Seçme Sınavı (Public Personnel Selection Exam)
MoNE	Ministry of National Education
NCATE	The National Council for Accreditation of Teacher Education
OECD	Organisation for Economic Co-operation and Development
ÖABT	Öğretmenlik Alan Bilgisi Testi (Teaching Field Knowledge Test)
ÖSYM	Assessment, Selection and Placement Center (Ölçme, Seçme ve Yerleştirme Merkezi)

PASS	Abbriation used for 'Priority of System Requirements', 'Anxiety', 'Skill Delay', and 'Shortcomings'
PAST	Abbriation used for 'Practice', 'Assessment', 'Speaking', and 'Technology Literacy'
PISA	Programme for International Student Assessment
PM	Policymaker
ST	Student Teacher
T	In-service Teacher
TAE	Teacher Assessment and Evaluation
TT	Teacher trainer
TESOL	Teaching of English to Speakers of Other Languages
TEYL	Teaching English to Young Learners
TOEFL	Test of English as a Foreign Language
YÖK	Yüksek Öğretim Kurulu (Council of Higher Education)

CHAPTER I

INTRODUCTION

In this chapter, background to the study, statement of the problem, purpose of the study, research questions, significance of the study, assumptions, limitations of the study, and definitions of the key terms are explained.

1.1 Background to the Study

It is the goal of every country to offer a high-quality education to its students. In order to meet this goal, it is necessary to recruit people who are eager and able to serve as teachers. As there is a positive relationship between the issue of teacher recruitment and the issue of teacher quality (Rivkin, Hanushek & Kain, 2005), even those nations whose students have high scores on international exams seem to be interested in the issue of teacher quality. Selecting and recruiting teachers is not just about ensuring an adequate number of teachers, but about having them in the teaching profession who are as accomplished as possible (Allen, 2005 p.10). The process of the quality teachers has become one of the most important agendas for nations (Beauchamp, Clarke, Hulme & Murray, 2013; European Commission, 2013b; Ingersoll, 2007, p.1; OECD, 2013c). It may be due to the fact that, among all school resources, the most important determinant of student achievement is quality teachers (Berry, Daughtery & Weider, 2009; Chetty, Friedman & Rockoff, 2014; Darling-Hammond, 2010b; Ingle et al. 2011; Koedel, Mihaly & Rockoff, 2015; Rivkin, Hanushek & Kain, 2005; Stronge, 2010). In this concept, there is a strong relationship between the effectiveness of any educational system and its teacher assessment and evaluation model (Johnstone, 2004; OECD, 2005). Any disorder, failure, or malfunction

that may occur in assessment and evaluation will affect the functionality of an entire system that prevents the occurrence of effective teaching and learning environments desired (OECD, 2009).

Research provides clear and compelling evidence that a well-trained teacher is the most important factor in terms of influencing student learning (Chetty et al., 2014; Creemers & Kyriakides 2008; Darling-Hammond, 2010b; Goldhaber and Anthony 2007; Hanushek 2010; Ingle, Rutledge & Bishop, 2011; Liu & Zhao 2013; OECD/UNESCO, 2016; Stronge 2007, 2010). Therefore, enhancing the quality of teachers is of importance since poorly trained teachers do the most harm to the students (Mitchell & Barth, 1999). As Clement (2009) indicates, bad recruitment decision not only may decrease student achievement but also it may cause poor morale among colleagues. Likely, teachers are considered to be responsible for student outcomes (National Council on Teacher Quality, 2013). According to Pillsbury (2005), recruiting an effective teacher is thought to be the most important decision by many administrators who possess ultimate chance to influence both instructional practices and student learning via teacher selection process (Mertz, 2010). The mentioned studies reveal that the policymakers should pay special care to teacher assessment and evaluation for better quality education and greater student achievement.

The quality of an education system can be increased only if teachers are trained with the necessary knowledge, skills and attitudes. In other words, the key to students' success is a well-trained and an effective teacher. So, what makes a teacher effective? It has been a subject of importance for many educational scholars. An effective teacher is defined by Schussler, Stooksberry and Bercaw (2010) as a person who "employs both successful teaching, which realizes intended outcomes, and good teaching, which is morally worthwhile" (p. 351). According to Pettis (1997), a professionally competent teacher, at first, needs to be principled, knowledgeable and skillful. Secondly, the effective teacher's professional needs and interests must change over time and continue to emerge during throughout the teaching process. Finally, the professional development of the competent teacher needs to be ongoing and personal. In Cripps Clark and Walsh's (2002) classification, it is claimed that an effective English language teacher must have discipline knowledge, pedagogical knowledge, knowledge of context, and personal knowledge. For Brown (2001), in addition to a teacher's language proficiency and personality, language-teaching skills are a must for successful teaching. Kilickaya and Krajka (2013) are in line

with Stronge (2007) who asserts that verbal ability, content knowledge, affective aspect of human behavior and knowledge of teaching and learning are the most common ones among the most significant prerequisites to create effective learning and teaching and achieve student success (p. 94). Some consider personal teacher dispositions, such as caring, as a vital factor in teacher effectiveness (e.g., Averill, 2012; Noddings, 2013). Whether the student teachers acquire these qualities requires teacher assessment. So, how should teachers be assessed? In their study, Leshem and Bar-Hama (2008) focused on three options: by observable behavior and measurable components, taking into account the practicum, or by their formal knowledge. Another question they ask is about whether the teachers should be assessed via traditional forms such as standardized and close-ended tests or alternative ones (e.g. portfolios, observation lessons, self-assessments, peer assessment, etc.). After deciding which forms of assessment need to be applied and which areas of teachers' knowledge and/or performance need to be assessed, the next step is to determine whether teachers should be assessed only when they are recruited, or, on the contrary, they are to be evaluated at regular intervals throughout their teaching profession (Kilickaya & Krajka, 2013).

Consequently, identifying and choosing appropriate student teachers has been a continuing challenge for nearly one century (Staiger & Kane, 2015). That is, assessing and evaluating effective teachers and providing every child with a high-quality education should be one of the most important issues for all nations, because making the best selection and recruitment choices is paramount to better teaching and greater student success. This purpose cannot be achieved without qualified teachers, because an effective teacher reduces students' anxiety (Chen & Lin, 2009; Ghasemi & Hashemi, 2011; Çelik, Arıkan & Caner, 2013), treats students with warmth and kindness and provides an environment in which there is a positive student-teacher interaction (Telli, den Brok & Çakıroğlu, 2008). In this context, designing an effective TAE model is vital and finding practices that can overwhelm barriers to appropriate recruitment of qualified teachers is subject of educational study (Liu & Johnson, 2006; Donaldson, 2013).

1.2 Statement of the Problem

In Turkey, English is a compulsory subject to be taught at primary school starting from second grade although it is neither an official nor a second language. In contrast to ESL

setting, as English is common only in education, government, and private business in Turkey (Dogancay-Aktuna, 1998), the success of English language education is not at the desired level. This problem has forced the Ministry of National Education (MoNE) to initiate various reforms, specifically in 1997, 2006 and 2012 (see Kırkgöz, 2005, 2007 for more details). For example, with the introduction of Teaching English to Young Learners (TEYL) in 1997, the purpose was to develop students' communicative competence, raise their awareness of the nature of English as a foreign language (EFL), and developing their language - culture relation. Moreover, many attempts were applied in 2006 and 2012 in order to raise the quality of English teaching. However, despite these reforms, being fluent in English, especially being able to communicate in English, is still a questionable issue in Turkey, and this situation has put English language teachers into target, because "what teachers know and can do is the most important influence on what students learn" (National Commission on Teaching & America's Future, 1996, p. 10). So, having qualified teachers is crucial for helping students to enhance their success. For this reason, assessing and evaluating effective and well-qualified teachers are the greatest challenges that today's educational leaders face (Stronge & Hindman, 2006 p. vii).

Research displays that the quality of teachers mirrors the quality of teacher assessment and evaluation model (Jacob, 2016; Stronge & Hindman, 2006 p. 17). Therefore, one of the fundamental assumptions of this study is that an effective recruitment model results in choosing more effective teachers. The assessment tools used and the way assessment process actualized all determine the quality that an English teacher has. Turkey often changes and/or adopts the teacher assessment and evaluation models in order to recruit effective teachers. However, the use of only one method of assessment (multiple-choice questions) in teacher assessment and evaluation has led to a much more test-driven study habit. In other words, student teachers' listening, writing and speaking skills, and their teaching ability are not assessed. Currently, the selection and recruitment of teachers for both public and private schools is done according to the criteria determined by the MoNE. However, while private schools are autonomous in deciding how to assess student teachers with the aim of finding the best person for their schools, those who want to work in public schools are recruited via centralized assessment system. First, those who want to be teachers at public schools need to take the examination called KPSS (Kamu Personeli Secme Sınavı, in English: Public Personnel Selection Examination). Through KPSS, student teachers are subjected to multiple-choice exams including general culture, general

skill, and educational sciences. Second, for becoming an English language teacher, student teachers are asked to take another test prepared in English, namely ÖABT (Öğretmenlik Alan Bilgisi Testi, in English: Teaching Field Knowledge Test). In addition, a new component has been introduced very recently. Each candidate is interviewed and the interview score is determinant in deciding whether the candidate will be assigned. The interview is conducted in Turkish and it is applied to assess the student teachers' communication ability. However, the Ministry seems to terminate the Interview process later this year.

ÖABT is a high-stakes test which is taken by thousands of student teachers every year. For example, 11.535 English language student teachers took this examination in 2017 (ÖSYM, 2017). Selecting the best student teachers among them requires a well-fitted selecting system which can distinguish those who fit the job from those who do not. Accordingly, the introduction of ÖABT aims to assess and evaluate those who have required knowledge and skills in their field. However, the exam format obstructs teaching skills to be assessed. That is to say, what student teachers know is assessed through paper-pencil exams (in multiple-choice format) in the general culture, general skills, educational sciences, and field knowledge. Nonetheless, whether they can translate what they know into teaching is not assessed. In contrast to Turkey, when countries achieved high success in PISA 2006, 2009, 2015 (such as Finland, Japan, South Korea, and New Zealand) are examined in terms of the policy of teacher training, selection and recruitment, it is seen that various assessment tools (e.g. written exams and performance exams) are used. Because the authorities in these countries are aware of the fact that hiring a qualified teacher positively affects student outcomes (Cruickshank & Haefele, 2001; Rivkin et al., 2005; Stronge, Ward, Tucker & Hindman 2007).

Klassen, Durksen, Kim, Patterson, Rowett, Warwick, Warwick and Wolpert (2017, p. 97) stated that

> In any jurisdiction, selection is necessary for three reasons: a) to make decisions about 'selecting in' when the number of applicants outweighs the number of available places, b) to make decisions about 'selecting out' in order to identify those candidates who may be unsuitable, and c) to provide a profile of candidates' strengths and weaknesses for future development. At the foundation of selection research is the belief that individuals vary in personal attributes and experiences, and that these individual differences are related to future behaviors in training and professional contexts.

Unfortunately, in Turkey, the current teacher assessment and evaluation (TAE) model basically depends on the first criterion. From that point of view, developing a new fair and valid TAE model was seen as de rigueur.

1.3 Purpose of the Study

It is widely accepted that high-stakes tests have consequences at the both micro level of language teaching and learning inside the classroom and macro level of education and society as a whole. Literature shows that tests are used for many purposes, such as selection, employment, promotion, immigration, citizenship, monitoring the performance of students and schools, "implementing educational policies, reforming educational systems, and deciding on the distribution of funding" (Cheng, 2005; Eades, 2005; McNamara, 2012 as cited in Pan & Roever, 2016, p. 1). Due to the importance of tests in all these domains, investigating the consequences of test use has been considered as vital steps in validating a test (Bachman, 2005; Hitchcock, Onwuegbuzie & Khoshaim, 2015; Kane, 2006; Messick, 1989). However, studies on test washback and test impact have focused primarily on issues conducted within classroom settings and they have devoted little attention to macro level (social) consequences (e.g. Cheng, 2005; Green, 2007; Saville & Papp, 2009; Shohamy et al., 1996; Shohamy et al., 2009; Wall, 2005 as cited in Pan & Roever, 2016).

Research indicates that the process of teacher assessment and evaluation is paramount and all countries aim to recruit effective teachers (Kyriacou, 2009; Moore, 2004; Tsui, 2003). If this process is not managed well and correctly, it can cause serious damages. That is, recruiting more teachers will not solve the quality education and teacher crisis if the rationale behind the selecting and recruiting process is not considered as finding the best person who best fits the needs of the students, schools, and finally quality education. From this point of view, the author of this study insists that there is a need to adapt policies to facilitate the process of assessing and evaluating quality teachers with the required knowledge, skills and attitudes. Therefore, this study aims to develop a TAE model for EFL contexts in which the assignment process is done centrally, criticizing 'one-size-fits all' assessment.

While there is a consensus on the fact that teacher quality is a problem in Turkey, there is little concurrence and much disagreement about what the sources and reasons behind it. Various competencies that an effective English language teacher needs to have are listed by researcher (e.g., discipline knowledge, pedagogical knowledge, language proficiency, interpersonal communication ability, and personality, love for the target language (TL), language-teaching skills, motivation and enthusiasm etc.) (e.g., Brown, 2001; Casey & Childs, 2007; Cripps Clark & Walsh, 2002). The MoNE has revised competencies for teachers. Whether these domains are kept in mind during teacher assessment is another goal of the current study. The researcher has also argued the characteristics that an English language teacher should have in order to be effective or qualified by searching local (for example, MEB, YÖK) and global competencies (such as CAEP, CEFR, EPOSTL, Education Council, TESOL).

In a nutshell, the primary purpose of this dissertation is to suggest a more reliable, valid and fair TAE model for English language teachers and it has the following sub-purposes;

1- to examine the perceptions and attitudes of education stakeholders (e.g. student teachers, in-service English language teachers, teacher trainers, education experts, and policymakers) about the current TAE model in Turkey regarding consequential validity,

2- to develop an attitude scale for English language teachers so as to reach a wide population with the aim of generalizing the results, and

3- to reveal the impact of the current TAE model on student teachers, in-service English teachers and education.

1.4 Research Questions

Research questions of the study are as follows:

1. Which competencies should a qualified English language teacher have?

2. What are the perceptions of stakeholders of English language teacher assessment and evaluation (ELTAE)?

   2.a. What are the student teachers' perceptions of ELTAE?

   2.b. What are the in-service EFL teachers' perceptions of ELTAE?

   2.c. What are the teacher trainers' perceptions of ELTAE?

2.d. What are the education experts' perceptions of ELTAE?

2.e. What are the policymakers' perceptions of ELTAE?

3. What are the attitudes of in-service EFL teachers towards ELTAE?

4. What might an appropriate EFL teacher assessment and evaluation be?

1.5 Significance of the Study

With the globalization and the widespread use of English in the world, Turkey has updated its language education policy to develop the quality of education such as 1997, 2006 and 2012 reforms (see Kırkgöz, 2005, 2007 for more details). Although the goal for adopting reforms is to make students become more familiar with English, and to improve communicative language abilities of language learners, studies conducted on second/foreign language learning depict that the desired level of proficiency in English is not achieved by Turkish learners (Karahan, 2007), and that foreign language learners, particularly Turkish learners, are unwilling to participate orally (Jackson, 2002). These limitations let public question the quality of teachers and education, because there is a positive correlation between teacher quality and learning outcomes of students (Araujo, Carneiro, Cruz-Aguayo & Schady, 2016; Canales & Maldonado, 2018). In addition, although many factors play a role in shaping the quality of education, experts arrive at a consensus that the quality of any educational system depends on the efficiency of teachers (Eurydice, 2012 p. 9; Ingersoll & Merrill, 2010). In this concept, not only will the conclusions drawn from the study be meaningful for the teacher assessment and evaluation policy of Turkey, but also it will contribute to the quality of education, as there is a strong relationship between the effectiveness of an education system and its teacher selection and recruitment techniques (Johnstone, 2004; OECD, 2005).

To overcome the aforementioned problems, stakeholders in education have sought "the most effective strategies for educating and training the nation's teachers" (Allen, 2003 p. 3). In a similar manner, researchers have started to focus on qualities that contribute to making effective English teachers, investigating views of foreign language teachers and/or students in Turkey (Çelik, Arıkan & Caner, 2013), in Israel (Brosh, 1996), in Korea (Barnes & Lock, 2010; Park & Lee, 2006), in Thailand (Wichadee, 2010), in Taiwan (Chen & Lin, 2009). However, whether student teachers acquire these qualities requires

teacher assessment (Kılıçkaya & Krajka, 2013). Thereby, this study has several contributions to the literature as it intends to bridge the gap between theory and practice of teacher selection and recruitment in EFL setting, applying the voices of stakeholders including student teachers, in-service English teachers, teacher trainers, education experts, and policymakers.

There are studies conducted to reveal the opinions of school administrators regarding vital factors in assessing teacher recruitment and/or which factors principals think most important in teacher selection (Ballou, 1996; Boyd, Lankford, Loeb, Ronfeldt & Wyckoff, 2005; Cain-Caston, 1999; Donaldson, 2013; Liu, Liu, Stronge & Xu, 2016; Mertz, 2010; Ralph, Kesten, Lang & Smith, 1998). However, in those studies, school principals have autonomy in selecting and recruiting teachers which is totally different from Turkey. Additionally, these studies usually gave voice to only one party (e.g., school principals, or teachers etc.) while other stakeholders were not taken into consideration. Though, all stakeholders should be given an opportunity to participate actively in the process of testing, from construction to assessment systems (Shohamy, 1996). The current study is significant in giving voices to stakeholders during the process of developing a teacher assessment model.

In Turkey, there are studies comparing Turkey's teacher training programs to other countries such as Japan, South Korea, Finland, and New Zealand (Mete, 2013); Austria (Canbulat & Canbulat, 2015); Japan (Aldemir & Er, 2012); Denmark and Sweden (Solak, 2016). Literature is also rich with studies dealing with teachers' and student teachers' perceptions towards KPSS (Atav & Sönmez, 2013; Çimen & Yılmaz, 2011; Erdem & Soylu, 2013; Gökçe, 2013). Yet, no information was given about how the process of English language teacher recruitment was actualized. Although Kılıçkaya and Krajka (2013) provided a good review of how English language teachers were selected and recruited in Turkey and Poland, they did not inform the readers about the impact of the TAE model on participants and no information was available about the recent developments of the KPSS such as ÖABT and Interview, because none of them was a component at the time when they conducted the study. More importantly, to the best knowledge of the researcher, no study has been conducted to offer a teacher recruitment model for English language teachers in a centralized EFL context.

Literature shows that despite the fact that many studies are available concerning teacher training, selection, and recruitment, "[c]ompared with other areas of educational research, studies of hiring practices are few...(and) advice is scarce" (Jensen, 1989, pp. 229-230). It seems that there is a need for studies explaining the required English language teacher competencies, whether these competencies are assessed while assign and evaluating teachers in Turkey. While doing this, a comparison among various countries was provided in order to gain an understanding of how different nations manage their teacher supply. This consideration seems to be very important, not only for the future employability of English teachers in EFL contexts where the recruitment takes place centrally but for the welfare of the learners. However, it is common that policies are often imposed on the users "for political and social reasons, without attention being paid to the needs and wishes of those affected by the policy and those expected to carry it out" (Shohamy, 2006: 143), so the voices of education stakeholders will be carefully considered. In this sense, this study aims to offer a teacher assessment and evaluation model which would yield useful insights into assessing foreign languages teachers in EFL context, by examining the perceptions and attitudes that different stakeholder groups (e.g. student teachers, in-service teachers, teacher trainers, education experts, and policymakers) have about the process since their perceptions can be value in understanding consequential and social validity of the standardized testing movement (Hitchcock et al., 2015).

1.6 Assumptions

The current research includes both verbal and written consents of participants via semi-structured interviews, focus group interviews and an attitude scale. Therefore, all participants are expected to understand the questions of the interviews and the items of the attitude scale appropriately and it is assumed that the participants responded to questions and items objectively and honestly. In addition, the study assumed that the interpretation of the data would accurately reflect the view points of the respondents.

1.7 Limitations of the Study

This study has some limitations. One of the limitations of this study is the qualitative sample, particularly student teachers because it consists of student teachers from three

universities, namely Ankara University, Gazi University, and Hacettepe University, which are located in Ankara province. Thus, their reflections may not represent the whole student teachers in Turkey. In addition, this study is limited to accessible in-service English language teachers, teacher trainers, education experts, and policymakers. Thus, their statements may not represent those of stakeholders in Turkey. Another limitation is the quantitative sample of the study. It is limited to the attitudes of 271 in-service English language teachers.

1.8 Definitions of the Key Terms

Effective/Quality Teacher: According to some researchers (for example, Stronge & Hindman, 2006), attributes of teachers can be listed as verbal ability, content knowledge, education coursework, teacher certification, and teaching experience. However, literature shows that little consensus is available "on how to define a qualified teacher" (Ingersoll, 2001, p. 42). Therefore, in this research an effective/qualified EFL teacher is that who has required knowledge (e.g. literacy, numeracy, ÖABT, curriculum etc.), skills (ability to perform a lesson, technology literacy, a good language proficiency, assessment literacy, classroom management, communication etc.), and attitudes (attitude about teaching, self-efficacy, motivational disposition etc.).

Equitable/Fair: Literature provides various definitions for the description of f*airness* in testing, as there is no a sole universally adopted definition of it in the literature (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014, p. 49). However, its definition in the current study is in compliance with ETS Standards for Quality and Fairness which defines the definition of fairness as "the extent to which the inferences made on the basis of test scores are valid for different groups of test takers." (Educational Testing Service, 2014, p. 19).

Out-of-field teachers refer to those who teach subjects for which they have not attained specialized knowledge during their undergraduate or graduate school period (Ingersoll, 2007). In the present study, it was used for in-service English teachers who were not graduates of ELT programs.

Teacher Assessment and Evaluation/Teacher Selection and Recruitment: The process of choosing high-quality teachers from among the assembled applicants and assigning them

11

to the schools. According to Klassen et al. (2017, p. 97), there are three necessary reasons for selection in any jurisdiction, and one of them is for making "decisions about 'selecting in' when the number of applicants outweighs the number of available places", so in the present study *teacher assessment and evaluation* and *teacher selection and recruitment* are interchangeable.

Teacher Qualities: Attributes teachers possess that enable them to effectively impact student achievement. These include characteristics that teachers bring with them to the job, performance on a daily basis, and their ability to build relationships and make connections with students (Kennedy, 2008).

CHAPTER II

REVIEW OF THE LITERATURE

Conceptual frameworks enable researchers to conceptualize and conduct research studies, by adding consistency, clarity, relevance, and rationale to a literature review and helping "organize the researcher's thinking about the research problem, current practices, contextual features of the environment, participant characteristics, theoretical foundations, empirical findings, and appropriate methods for addressing the problem." (Antonenko, 2015, p. 67). A well-organized conceptual framework resembles a ''spotlight'' (Maxwell, 2013, p. 49) in elucidating the current knowledge on the problem of interest, helping notice the appropriate ways for problem solution, and identifying the connections between related ideas and approaches that may otherwise be overlooked. In addition, a good conceptual framework can be considered as a ''map'' (Miles & Huberman, p. 20), because it conspires the conceptual landscape of the problem and delineates probable avenues to explore it. In a nutshell, a conceptual framework supports the goal, an is a crucial tool for organizing the study, as it provides not only a theory-based but a data-driven argument for the significance of the issue.

2.1 Teacher Assessment and Evaluation

The process related to assessment and evaluation is highly emphasized in many international academic studies and reports (Al-Issa, Al-Bulushi & Al-Zadjali, 2017; Brutti & Sánchez, 2017; Darling-Hammond & Rothman, 2011; Guha, Hyler & Darling-Hammond, 2017; Herrmann, 2018). According to those reports, a selective system is of importance for selecting and recruiting qualified people to the education system as there is

a significant relationship between recruiting practices and teacher effectiveness (Rice, 2008). Stronge and Hindman (2006, p. 1) indicated, "if we want to maximize the impact of the hiring process, we need to actively look for, identify, and hire teacher applicants who exhibit what research indicates to be the qualities of effective teachers". On the contrary, a poor hiring decision can result in decreasing quality of education for years to come (Clement, 2005; Mason & Schroeder, 2010). It is known that not all students are taught by an effective teacher, so hiring better teachers is a must for educational leaders (Rose, English & Finney, 2010) in order to ensure that all students have access to high-quality teaching. Similar emphasis is also placed in the Teacher Strategy Paper published by the MoNE in 2017, whose primary goal is to ensure that every student in public school is being educated by quality teachers (MEB, 2017a).

The idea that teachers have impact on student academic achievement makes many nations revise and renew the way they evaluate their teachers with the aim of improving their education. Although many variables can be listed as promoters of student success in schools, the most significant factor seems to be teacher quality (Harris & McCaffrey, 2010), which affects students' both short- and long-run life outcomes (Chetty, Friedman & Rockoff, 2014; Jackson, 2014). In other words, having qualified teachers in every classroom has become a desire for parents, educators, and policymakers. To do this, developing a valid, fair and reliable teacher assessment and evaluation process should be the very first step that committed by any government in order to enhance the quality of education in their country (Thaman, 2007).

There is no single answer of what and how to evaluate. Yet, there is a growing consensus that every student teacher's pedagogical knowledge, subject-area knowledge, and professional teaching ability need to be evaluated before becoming a teacher (National Education Association, 2010) via the combination of multiple measures as no single measure is by itself capable of evaluating the efficacy of teachers (Cantrell & Kane, 2013). That is to say, if the goal of teacher evaluation is to select and recruit English language student teachers, what they know and whether they can translate their knowledge into practice should be assessed in order to ensure that the right person is chosen for teaching English. Additionally, a great deal of attention should be paid to the assessment process because "the choice of evaluation tool would likely influence the qualities and activities of teachers" (Harris, Ingle & Rutledge, 2014).

The available international literature on the context of assessment of teacher recruitment is totally different from that of Turkey. The latter offers no autonomy to schools or districts due to its highly centralized structure while, in the former, schools or district are mostly autonomous to hire their teachers, to develop their own teacher evaluation policies so that they could eliminate ineffective teachers (Adnot, Dee, Katz & Wyckoff, 2017). Thus, studies on teacher hiring in international setting focus largely on either what kind of features principals look for, or which processes applied while selecting and hiring a teacher. Researchers (e.g. Rutledge, Harris, Thompson and Ingle, 2008) identify portfolios, videos of instruction, and demonstration lessons as promising tools to ascertain a student teacher's teaching ability. However, in Turkey, the centralized recruitment process is carried out by the MoNE and student teachers who wish to work in public schools (from primary to upper secondary school) need to pass a norm-referenced paper-pencil examination organized by ÖSYM (Ölçme, Seçme ve Yerleştirme Merkezi, in English: Assessment, Selection and Placement Center). Although standardized test scores may provide a general idea of what student teachers know and are able to do, these tests are not great measures of student teachers' achievement since they do not provide detailed insight into how they can teach in practice (Darling-Hammond, 2014; Toch & Rothman, 2008).

Even if researchers have not reached a consensus on what 'quality' is and the best way to measure it (Cohen, 2010; Kennedy, 2010), teacher evaluation should assess teaching quality and it should cover multifaceted evidence of teacher practice (Darling-Hammond, 2014). Moreover, a fair and an accurate evaluation system requires "gathering multiple sources of evidence about teacher practice", considering the characteristics and practices that a qualified teacher should have (Danielson, 1996, 2007; Peterson, 2000 as cited in Isoré, 2009 p.20). Consequential validity takes into account the perspective of multiple stakeholders (in this study; student teachers, in-service English language teachers, teacher trainers ELT departments, education experts, and policymakers) to question the logic of a high-stakes educational intervention (in this case, KPSS) (Cook & Lineberry 2016). Shohamy (2001b) acknowledges that critical testing attempts to explain misuses of the tests so as to decrease their harmful effects and it also aids critical tactics to examine consequences and impacts of the test uses, and help individuals have voices in test development process. Qi (2005) believes that stakeholders pay attention to the results of high-stakes tests. The influence of test results on teaching and learning is generally defined as washback (Bailey, 1996: 259), so "[e]xam boards need to monitor the test's effects on

language materials and on classroom activity and to seek information on, and the views of, a full range of stakeholders" (Weir, 2005 p. 213). From this point of view, it is useful to mention consequential validity, critical language testing, high-stakes testing, and washback under this title.

2.1.1 Consequential Validity

Examining how education stakeholders perceive education policy and the use of testing to develop educational outcomes for all students is about test validity, because any knowledge of reality (e.g. test validity, fairness and impact) does not remain independently of how people experience it (Cohen, Manion & Morrison 2011 as cited in Kumar, Roberts, Bartle & Eley, 2018). Researchers have not reached a settlement whether consequential validity needs to be considered as a part of validity. According to Popham (1997), the consequences of testing are vital and both test developers and test users should investigate it, but it should not be thought part of validity. However, Cizek, Rosenbert and Koons (2008) argued that consequential validity may be the most disputable area of test validity. In consequential validity, we are interested in a test's impact on individuals, institutions and society, and in the use made of test results (Weir, 2005 p.1).

Messick (1989) asserts a unified theory for construct validity and consequential validity issues. In much broader terms, construct under-representation and construct-irrelevant variances are two major threats to validity (Mesick, 1995). The first one includes error which is based on measurement of the specific test construct, while the latter includes disconnected factors that influence test performance (Cheng, 2014). In other words, construct under-representation happens when it is not possible for test takers to display their performance in a given examination due to the failure of the test in capturing and presenting vital aspects of the construct that it is aimed to measure, while construct-irrelevant variances involves factors that are irrelevant to the construct being measured and occurs due to the effects of irrelevant factors on scores. For example, construct under-representation could occur when a test designed for university admission is used for job certification purposes (Cheng, 2014). Test taker's psychology, personality characteristics, background knowledge, general intellectual or cognitive ability, knowledge of subject area, and test-taking strategies all can be potential sources as construct-irrelevant variances (Bachman, 1995; Cheng, 2014; Davies, 2010).

16

Bachman (2005) monitors that "the extensive research on validity and validation has tended to ignore test use, on the one hand, while discussions of test use and consequences have tended to ignore validity, on the other." (p. 7). The use of new validity theory (e.g., classroom assessment, and dynamic assessment) in instructional assessment contexts has established grounds for the inclusion of consequences and uses within validation studies (Kane, 2006). Specifically, construct under-representation and construct-irrelevant variance have been recognized to contribute towards test validation (Haladyna & Downing, 2004). Thus it is crucial to establish the link between test validity and consequences of test use by applying perspectives of multiple stakeholders within language assessment so as to advocate the use of test scores in pedagogical practices (Cheng, 2014). Fox and Cheng (2007, p. 9) voice that an analysis of "how test-takers interpret test constructs and the interaction between these interpretations, test design, and accounts of classroom practice" should take place in validation evidence from test takers. Cheng (2014) warns that the same claim should administer to research on other testing stakeholders.

Today, the effects and consequences of educational testing draw considerable interest by educational stakeholders, including student teachers, in-service teachers and policymakers in many societies. Likewise, the consequential aspect of validity in teacher selection and recruitment now seems here to stay, so policymakers and test developers must take into account the intended and unintended interpretations and uses of test scores. Researchers can examine consequences of testing in various policy contexts to support their efforts. As Cheng (2014) states we need empirical research that looking for evidence from multiple stakeholder perspectives on what specific constructs are under-represented and on what the sources of construct-irrelevant variances are. Consequential validity has been the center of interest in studies conducted in both educational sciences and other areas. For example, Brewer, Knoeppel and Lindle (2015) conducted a study with more 5,400 stakeholders, including parents, business and industry persons, community leaders, and educators in the state of South Carolina, with the aim of exploring stakeholders' perceptions of the variety of tests used in classrooms and schools. According to results, the majority see the assessments impractical in the teaching and learning process. In Australia, using consequential validity as a theoretical lens, Kumar et al. (2018) used an online survey to examine views and experiences of 447 medical school students and applicants towards the Graduate Medical Schools Admission Test (GAMSAT), and the consequences of testing. Results indicated that there was a moderate agreement regarding the GAMSAT as a fair

selection method. Using GAMSAT as a selection tool had some unintended consequences, such as obstacles related to test affordability and language. Alavi and Masjedlou (2017) examined the validity of IELTS Academic Writing Task One (IAWTO) by comparing it with IELTS Academic Writing Task Two (IAWTT). The results of 53 participants' responses showed the insufficiency of IAWTO input, giving the construct underrepresented. Ryan (2018) conducted a mixed methods study to reveal the positive and negative consequences of 245 EXAVER candidates as a result of preparing for and taking a language test, based on their responds through a web-based questionnaire survey and a series of semi-structured interviews. Language test candidates were found to possess positive and negative opinions about tests, and a strong desire to share them with test developers.

2.1.2 Critical Language Testing

Reliability and validity refers to the main criteria for well-established tests in the traditional psychometric view. However, with the introduction of Messick's consequential validity, instead of traditional testing, critical theories of validity have been valued. That is, test makers have started to pay attention to the uses and consequences of tests that affect the society and education. According Shohamy (2017 p.443), the uses of tests as disciplinary tools evoked strong feelings in test takers. She adds that that testing policy, their impact, consequences and intentions have started to be questioned is the essence of Critical language testing (CLT).

CLT refers to the investigation of uses and consequences of tests in education and society (Shohamy 2001a). The importance of tests as a powerful tool in education and society has led to CLT emerge. This realization may cause unintended consequences that need to be investigated and assessed. The power of high-stakes tests causes test takers and educational systems to alter their instructional behaviors (Shohamy, 2017). She also underlines that the existing research on CLT tries to explain how the misuses of high-stakes tests and their impact moves beyond learning and teaching. Shohamy (2001a) presents various examples to show that what kind of issues that CLT can address (for greater elaboration, see Shohamy, 2001b):

  • CLT investigates the intentions of tests

• CLT declares that the tester has incomplete knowledge

• CLT does not assume the uses of 'the test' as the sole instrument for assessing knowledge; instead, the use of multiple procedures is welcomed

• CLT deals with questions about testers, test-takers, their agendas, contexts, testing methods, decisions reached based on the tests, their impacts, and many other questions.

From this point of view, CLT takes into account the impact of tests on stakeholders and the importance of cooperation among them, which is also one of the aims of the current study. CLT assumes tests as unneutral instruments, but devices of power and control, used by authorities undemocratically and unethically to form individuals' lives. Noam (1996) warns that the use of tests can be constructive or destructive, noting that "[h]ow we assess can support careers and make people successful, but it can also destroy people's careers, place unfair burden on individual's self-perception and unnecessary hurdles in the path of their achievement" (p. 9). According to Shohamy (2001a), testees should be provided right to question the test itself as their views might be informative to test developers.

Safari (2016) used Sampad Entrance Exam (SEE) and Nation-wide University Entrance Exam (NUEE) to examine the status of language testing in Iran. Semi-structured interviews were used to gather in-depth data from 32 stakeholders consisting of teachers, administrators, test-takers and parents. All participants agreed on democratic multiple assessment sources of the construct to outweigh adverse consequences of the high-stakes tests. Shih (2008) interviewed with students and their families to investigate test takers' opinions of the General English Proficiency Test (GEPT). The findings displayed that all families and approximately sixty percent of the students confirmed this policy. In addition, students suggested that the Language Training and Testing Center (LTTC) change the way listening comprehension questions were administered, adopt more authentic speaking test, reduce registration fees, choose well-known places for test centers, strive for making certificates internationally recognized, and administer the GEPT more frequently.

2.1.3 Washback

The effect of language testing on teaching and learning is known as 'backwash' (Hughes, 1989) or 'washback' (Alderson, 2004). It is also used as "test impact" by Bachman and

Palmer (1996). In addition, Wall (1997) states that "test impact refers to any of the effects that a test may have on individuals, policies, or practices within the classroom, the school, the educational system, and society as a whole" (p. 11). It can be negative (harmful) or positive (beneficial). While previous one refers to a narrow definition of language ability, the latter results in 'good' teaching practice. Hughes (1989) presents backwash as the most important issue for teachers. Until the early 1990s, it was believed that there was a relation between washback and "the quality of the design of examinations" (Tsagari, 2011 p. 431). However, current research has shown that this relationship is not always linear. Both national (Sevimli, 2007) and international (Wall & Horak, 2011) research has revealed that washback is a multi-layered and complex phenomenon, which can vary in 'form' and 'intensity' (Cheng, 2005).

Washback becomes a well-known concept due to the misuses of tests. The literature on washback abounds in Turkey (Akpınar & Çakıldere, 2013; Çakır, 2017; Hatipoğlu, 2016; Kılıçkaya, 2016; Külekçi, 2016; Özmen, 2011a, 2011b; Sayın & Aslan, 2016; Yıldırım, 2010). When the aforementioned studies are examined, it is seen that the mostly focused on attitudes and perceptions of students towards exams (for more detail, please see Toksoz & Kilickaya, 2017). For example, Özmen (2011b) conducted a mix-method design study with 164 student-teachers so as to analyze washback effect of "The Selection Examination for Professional Posts in Public Organizations" (SEPPPO) in Turkey. The results of is study revealed that SEPPPO exerted negative and harmful effects on student-teachers, educational faculties and families. Akpınar and Çakıldere (2013) focusesd on the washback effects of two high-stakes in Turkey, Public Personnel Language Examination (KPDS) and Inter-University Foreign Language Examination (UDS), with the purpose of investigating the impact of these tests on academicians' receptive and productive language skills. They asked 103 academic personnel to complete a 26-item questionnaire. A significant difference was found to be between reading and writing; reading and listening, but no significant difference was observed to be between reading and speaking.

2.1.4 High-Stakes Tests

The high-stakes tests are those whose results are applied to make significant decisions which affect students, teachers, administrators, communities, schools and districts (Madaus, 1988 as cited in Hatipoğlu, 2016). They are steadily increasing around the globe

since they are used for the students' selection and placement (Choi, 2008; Hatipoğlu, 2013). Their results are accepted as the sole criteria while deciding whether candidates can enter the desired programs (Yildirim, 2010). Therefore, literature proves the vital impact of high-stakes tests on not only teaching paradigms, but also on educational systems in several countries (Lambourdi, 2014; Shohamy 2001), leading to the "narrowing" of the curriculum (Cheng 2005), to shift teaching methods (Spratt 2005), course objectives (Cheng, Watanabe & Curtis, 2004), and individual learning styles (Shih 2009).

Menken (2017, pp. 386-7) critiques that policymakers and test developers rarely take into account the language policy byproduct of standardized tests although high-stakes testing has major implications for language education. High-stakes testing has both supporters (e.g. Haycock & Jerald, 2002; Scheurich & Skrla, 2001) and opponents (Dounay, 2000; (Meier, 2002; Wagner, 2003). Supporters believe that it forces the educational establishment to balance the gap between different student groups, while opponents assert that high-stakes tests decrease motivation, increase student anxiety, narrow the curriculum, cause teach to test. Polesel, Rice and Dulfer (2014) carried out a study with 8000 educators in order to study the impact of NAPLAN on Australian schools and students. The NAPLAN (National Assessment Program – Literacy and Numeracy) is a national testing scheme, in which all Australian students in Years 3, 5, 7 and 9 should participate. The findings confirmed the results of studies in different countries on the capacity of high-stakes regimes to deform instructional practices, narrow the curriculum and students' educational experiences. Considering the relationship between classroom instruction and the impact of high-stakes tests, high-stakes tests may lead to teachers be prejudice against such exams because they may result in the exclusion of poor students (Au 2007; Moon, Brighton, Jarvis & Hall, 2007). In Turkey, Dikilitaş and Mede (2018) investigated English language teachers' involvement in High-Stakes language testing. Although they were observed to have affirmative perceptions of their involvement in High-Stakes language tests, their perceptual stance was low because they were not asked to contribute to process due to top-down assessment policy.

2.2 Teacher Selection and Recruitment in Turkey

Radical and comprehensive measures to train primary school teachers were needed when the primary education became free of charge and compulsory in state schools with the Law

on Unification of Education, published in 1924. In line with John Dewey (1924) and Alfred Kühne's (1925) suggestions, "Köy Muallim Mektepleri" (Village Teaching Schools) were established in 1926 in Denizli and Kayseri. They were transformed into "Köy Enstitüleri (Village Institutes)" in 1940. With the Law no. 6234 published in 1954, the Village Institute was terminated with the unification of teacher training institutions under the name of "İlköğretmen Okulu". With the National Education Basic Law numbered 1739 issued in 1973, teachers were asked to have a higher education degree and the ilköğretmen okulları which met criteria determined by the government were turned into Educational Institutions in 1974. With the Higher Education Law No. 2547 issued in 1981, all higher education institutions were gathered under the roof of the YÖK (Council of Higher Education) and academies and educational institutes were reconstructed as universities and faculties of education, respectively. Until 1982, the MoNE was responsible for training teachers. After that time, universities took over responsibility of teacher education. Since 1989-1990 academic year, the duration of education of all higher education institutions that educate teachers has been increased to at least undergraduate level. The teacher training programs have been revised many times according to national and international developments and these changes are continuing today.

Before the 1950s, high school grade point averages were used for the admission to the universities. In 1950s, universities began to conduct their own essay-type examinations. However, both the issue of subjectivity while rating essay-type questions and the increase in the number of high school graduates made authorities to centralize the entrance examination. Thus, almost all higher education institutions have started to take into account the results of the examinations organized by ÖSYM since 1974 (see http://www.osym.gov.tr/TR,8530/tarihsel-gelisme.html). Currently, those who want to enter an ELT program must take two examinations: Basic Competences Test and Foreign Language Test. They are all based on multiple-choice questions. When calculating language scores, the effect of Foreign Language Test will be 40% while the effect of Basic Proficiency Test will be 40% (YÖK, 2018).

In the post-republic period, many institutions were established and various regulations were made in order to train student teachers for primary and secondary education. Till 1950s, teachers who were assigned to villages were the graduates of Village Institutes. In 1959, Tevfik İleri, the Minister of National Education, announced a new model, "essah

öğretmen". All these attempts were inadequate to meet the need of teachers. That's why, different sources were applied in order to close that gap such as Reserve Officer Teachers (1960), Substitute Teacher (1961), Peace Volunteers (1962, US experts), Teaching Formation (1970), Teacher Training by Letters (1974), Intensive Teacher Training Program (1975), Those who did their military service as teachers (1987), and appointing all university graduates as teachers in mass without any conditions (1996) (Akyüz, 2001, p. 352; Çılgın, 2009 p. 28; Öztürk, 2005 pp.195-199; YÖK, 1998).

The MoNE was responsible for teacher selection, education and recruitment until 1982. Within the framework of the National Education Basic Law numbered 1739 in 1974, primary school teachers were being educated in two-year training institutes, which were renamed as educational colleges in 1982. Likely, secondary school teachers were being educated in educational institutes and higher teacher schools. After that time, universities became responsible for teacher education, so educational colleges and faculties of education started to train primary and secondary school teachers, respectively. In the 1989-1990 academic year, the duration of two-year education colleges was increased to four years and these colleges were transformed into faculties of education in 1992–1993 academic year. Those changes which caused a serious teacher shortage were due to the lack of an effective coordination and appropriate planning activities between the MoNE and the universities (Aydın & Baskan, 2005, p.192).

In 1985, for the first time, student teachers took a central examination in order to be an English teacher (see Çılgın, 2009 for more details). However, in 1992, due to the teacher shortage, the examination requirement for teachers was removed again. In 1996, almost about 50 thousand persons became a teacher without having pedagogical formation and without taking an examination. Yet, in 2001, teacher candidates started to be re-examined under the name of KMS (Kurumlar İçin Merkezi Eleme Sınavı) which was renamed as KPSS in 2002. The Public Personnel Selection Examination (KPSS) which determines whether the student teachers deserve the appointment is sine qua non for student teachers to work in public schools and conducted in four phases. Through KPSS, student teachers are subject to multiple-choice exams in the general culture, general skills (120 questions) and educational sciences (80 questions). The first phase consists of general culture and general skills. In the second phase, their knowledge of educational sciences is assessed. In order to be an English language teacher, student teacher should have the third phase,

ÖABT (50 questions) which has two main parts: Field Knowledge Test including *language proficiency, linguistics*, and *literature* (40 questions) and Field Training Test consists *of methods and approaches*, and *teaching language skills* (10 questions). The questions of the first and second phases are in Turkish, while those of third one are in English. Also a new development is the introduction of interview which is conducted in Turkish with the aim of assessing candidates' eloquence. However, why the communication skill of English teachers is assessed in Turkish is a vital question that requires a satisfactory answer.

2.3 Teacher Selection and Recruitment in Purposively Selected Countries

This part includes information about teacher assessment and evaluation of three countries. Before deciding on the sample, PISA 2015 results were examined. PISA divides countries into three categories: those in which students have high results, those in which students are at the average, and those in which students have low results (see Appendix A). The countries were chosen accordingly so that they could represent all results. In addition, three criteria were considered while selecting these contexts for making a more reasonable comparison between them and Turkey. The first criterion is the nature of the English. That is, English should be a foreign language instead of second language in the selected countries. The second one is that the languages of the selected countries should not belong to the family of English language. Finally, the third criterion is that there should be a balance between the population of the selected country and that of Turkey.

2.3.1 Teacher Selection and Recruitment in South Korea

Teaching is a permanent and highly respected job in South Korea (NCEE, 2013; OECD, 2016). Once recruited, teachers are automatically tenured till they are retired (Kang & Hong, 2008). When there is teacher vacancy as a result of teachers' temporary leave for child birth or rest, school principals employ contract teachers who cannot become regular teachers without passing the employment examination. Primary and lower secondary levels (from age 6 to age 14) are compulsory. Central and local authorities share the governance of the education system. The Ministry of Education is responsible for the whole process of education from primary to higher education. Also, municipal and provincial offices are

responsible for primary and secondary education, while councils for university and university college education are responsible for higher education (OECD, 2016).

The national entrance exam for tertiary education is the College Scholastic Ability Test. In 2008, the University Entrance Liberalization Policy was introduced, so universities could determine their own admissions criteria, such as students' senior high school records, their College Scholastic Ability Test (CSAT) grades, practical skills tests, essay-type examinations and oral interviews (Kang & Hong, 2008; Kim, et al., 2010; OECD, 2016). Toward the end of their teacher training program, students are asked to teach full-time for four to six weeks and they are supervised and evaluated not only by university advisor but also by classroom teacher and principal (Wang, Coleman, Coley & Phelps, 2003).

Primary school teachers are trained at universities of education, but secondary school teachers are educated at the following institutions: teacher's college, teacher training program of college of education, or school of education (Kim, et al., 2010). Graduates of these institutions can obtain the initial teacher's certificate, which is a must for those who want to be an English teacher in national and public schools. After obtaining teacher's certificate, they need to pass the National Teacher Employment Test (NTET) which is mainly composed of a paper-and-pencil test covering pedagogical and subject matter knowledge. While the administration of NTET is carried out by the councils of school districts, the Korea Institute for Curriculum and Evaluation (KICE) assigns specialists (mostly university professors who has more than 3 years' teaching experience) to develop the test (Kwon & Ju, 2012). Those who receive positions are preservice teachers with the top TET scores. However, private schools themselves determine their teacher recruitment methods and procedures. In addition to the Ministry of Education, the 16 metropolitan and provincial offices of education are responsible for the recruiting of national and public school teachers (Ingersoll, 2007; Kim & Han, 2002). Although the procedure of recruiting teachers is the same for both elementary and secondary schools, student teachers of these school levels take the examination separately.

The NTET assesses student teachers' knowledge, skills, and dispositions via three phases. The phases measure applicants' mathematics knowledge, pedagogical knowledge, and mathematics-specific pedagogical knowledge. While the first one consists of multiple-choice and short-answer formats, candidates are asked to write an essay in the second phase. The final exam is composed of teaching performance and interview. Unlike Turkey,

regional authorities employ teachers, so their scores are applicable in province they apply and the recruitment process has high standards since they were well assessed before admitted to a teacher training program (Kang & Hong, 2008).

The Korean history of teacher training and recruitment is not different from that of Turkey. In the case of an undersupply, out of-field teachers who did not undergo the regular processes of preparation, certification, and selection were recruited. Kim et al. (2010) provide detailed information how the process of teacher training and recruitment has been shaped since 1950s in their work. Temporary teacher training centers were established in order to meet the teacher shortage in 1950s. Transforming teacher's high schools into two-year universities of education and establishing graduate schools of education were important developments of the 1960s. The first one aimed to raise the quality of primary school teachers, whereas absorbing the demands of secondary school teachers was the reason behind the establishment of graduate schools. The 1970s may be claimed to be the birth of national and private teacher's colleges. In the 1980s, the duration of universities of education was raised to 4 years. However, the outstanding decision was to reduce the number of teacher training programs due to the problem of teacher surplus that the country first experienced. In addition, master teacher system was introduced with the aim of ensuring more professional teacher education practices. Until 1990, elementary school certificate was automatically given to graduates of universities of education and those who completed courses offered at temporary teacher education facilities and they were recruited as new teachers without an examination. Likely, regarding secondary school teachers, graduates of national colleges of education did not have to take an examination in order to be recruited. In the 1990s, the teacher recruitment examination system was renewed so that all student teachers had to sit for the exam based on open competition so as to be assigned to a national/public or a private institution. However, as mentioned above, private schools have started to determine their own teacher recruitment methods and procedures.

2.3.2 Teacher Selection and Recruitment in Spain

With a new law called the Organic Law for the Improvement of Educational Quality (Ley Orgánica para la Mejora de la Calidad Educativa, LOMCE, 2013), Spain has aimed to develop the quality of education. Among its major precautions "are greater autonomy for schools, new preventive diagnostic testing in primary education, more vocational pathways

starting in the final years of lower secondary education, and exit exams in lower and upper secondary education" (OECD, 2014, p. 4). Although the issue of abolishing the Prueba de Acceso a la Universidad (PAU), University Entrance Test, and giving autonomy to universities for designing their own selection tests is a matter of debate, in order to enter university, students must take the nationally conducted PAU, which consists of a compulsory general stage and a specific voluntary phase, and which is the same for all candidates, irrespective of the courses they want to take (Pérez-Granados, 2018; Ramírez Carpeño & Mekochi, 2015).

Primary and lower secondary education are compulsory and free, and begins at the age six and ends at the age sixteen; however, upper secondary school is not compulsory (Ramírez Carpeño & Mekochi, 2015). According to the LOMCE 2/2006 of May 3rd, teachers of primary education should have a teaching degree that enables them to teach all subjects, but subjects such as music, physical education and foreign language need to be taught by specialized teachers. Teachers must complete at least a bachelor's degree (a master's degree for secondary school teachers) (OECD, 2014 p. 4). Despite its efforts to promote education improvement, one of the biggest problems in Spanish education system is the selection and training of teachers (Pérez-Granados, 2018). Private and co-funded schools determine their teacher recruitment procedure. The only necessity for a student teacher in order to be a teacher in private school is to possess teacher training certification. On the other hand, for becoming a teacher in public schools, teachers must go through a nationwide competitive examination, which has two phases. There is a practical exam in the first part which intends to evaluate student teachers' specific knowledge of teaching, learning and techniques whereas the second part measures pedagogical knowledge of the student teacher by asking them to "present a teaching program based on the first exam" and to "do an oral presentation of a teaching unit" (Ramírez Carpeño & Mekochi, 2015 p. 117). They have also a practicum period that monitors their aptitude in teaching. Once the student teachers pass the exam, they can determine the place they want to work in according to their marks. Teachers have the status of career civil servant in Spain.

2.3.3 Teacher Selection and Recruitment in Thailand

Basic education is made free of charge and covers pre-primary education, primary (six years), lower secondary (three years) and upper secondary (three years) education. Pre-

primary education is not compulsory, but has become free of charge since 2009. Compulsory education covers primary (grades P1-6) and lower secondary levels, starting at age 6 to age 14) with 12 years of free basic education guaranteed by the constitution (Ingersoll, 2007). The Ministry of Education governs all education levels from pre-primary to higher education and it is the most important education provider, but ten other public bodies oversee their own institutions (OECD/UNESCO, 2016).

Thailand introduced many changes through 1999 National Education Act (NEA), such as decentralization of administrative structures, establishment of equity and learner centered learning, to improve the quality of the education system. All teacher training programs require an entrance examination. Prior to 2005, a four-year baccalaureate-degree program was enough for teaching jobs. Since then, obtaining a teaching license which requires completion of a five-year bachelor's degree in teacher education has become a must for all teachers signifying professional training (Teacher and Educational Personnel Act, 2003 cited in Ingersoll, 2007). Other graduates have to complete a one-year post-baccalaureate diploma in teacher training for a teaching license.

Being extensively considered unappealing and cumbersome, Thailand's recruitment process seems similar to Turkey's in a lot of ways, such as no correlation between the number of teacher job positions and general shortage of teachers, mostly determined by government policy and budget, so the needs of schools are not taken into account (Mattavarat, Viseshsiri & Siribanpitak, 2017). The lack of qualified teachers and teacher training programs is often considered to be the reason behind the inadequate education in Thailand, (OECD, 2013b). However, the country's two open universities, in which "any person wishing to enroll can apply without having to take an entrance examination" (Tongliemnak, 2010) may cause an oversupply of low qualified teachers. Although Thailand has a general oversupply of teachers, there is evidence of teacher shortages in key subjects such as foreign languages (OECD/UNESCO, 2016), which make it possible to witness P.E teachers who teach English in the public elementary schools (Dhanasobhon, 2006). This may explain why many teachers are unqualified and thus often have limited skills (Khamkhien, 2010) and why out-of-field teacher is common in Thai context.

Teacher recruitment in Thailand is determined on a national basis, controlled by the Office of the Teacher Civil Service and Educational Personnel Commission (OTEPC) and implemented by staff in each Educational Service Area (ESA). The allocation of new

teachers is based on the results of a competitive employment exam. The student teacher who has the highest result on that exam gets the first choice of the positions available. After making their choice teachers must also pass a selection test, which resembles to the employment exam (OBEC, 2015).

2.4 Teacher Quality

There has been an increasing interest in policies that promote teacher quality in recent years (Adnot, Dee, Katz & Wyckoff, 2017; Beauchamp, Clarke, Hulme & Murray, 2013; European Commission, 2013b; Mourshed, Chijioke & Barber, 2010; OECD, 2013c), because educational researchers agree that teacher quality may be the most significant factor which contributes to student success (Hanushek & Rivkin, 2006; Harris & McCaffrey, 2010; Staiger & Rockoff, 2010). One of the greatest challenges that today's educational leaders face is to select and recruit highly effective teachers because "[i]f we want students to succeed to their maximum potential, having a quality teacher working with every student is paramount" (Stronge & Hindman, 2006 p.1). Blömeke, Olsen and Suhl (2016) investigated the relationship between teacher quality and student achievement applying multi-level structural equation modeling to grade four student and teacher data from TIMSS 2011. The sample of their study included 205,515 students from 47 countries nested in 10,059 classrooms. Findings confirmed the existence of a significant relationship between teacher quality and student achievement.

It is now seen as de rigueur that an effective teacher can influence student achievement. The question is how to define teacher quality or what kind of characteristics qualified teachers have. No Child Left Behind has defined qualified teachers as those who are certified in the area in which he or she is teaching. According to some researchers, they are fully certified and have proven subject area competence (Beyer & Johnson, 2005; Mosely 2006 as cited in Kersten, 2008). Although there is still discrepancy within the educational community about the desired qualities (Ingle et al., 2011; Papa & Baxter 2008; Roellke & King Rice, 2008), research shows "that many kinds of teacher knowledge and experiences may contribute to teacher effects, including teachers' general academic and verbal ability; subject matter knowledge; knowledge about teaching and learning; teaching experience; and the set of qualifications measured by teacher certification" (Darling-Hammond & Wei, 2012, p. 614). There are, however, both convergences and divergences in these qualities,

making effective English teachers vary across different countries. For instance, Brosh (1996) examined views of foreign language teachers and students in Israel to reveal desired qualities of foreign language teachers and according to the perceptions of the participants, good knowledge of the TL, taking into account students' interest and motivation and availability to students were found to be desirable characteristics. In Taiwan, the perceptions of 198 students were investigated via a 50-item questionnaire by Chen and Lin (2009). They concluded that a qualified EFL teacher should have a good instructional competence, personality, and teacher student relationship. In Korea, Barnes and Lock (2010) conducted a study on that issue by applying perceptions of 105 first-year university students. Students were asked to write the features of effective EFL teachers in their native language. Korean students' perceptions were categorized as rapport, lesson delivery, fairness, knowledge and credibility, and organization and preparation. In Thailand, 400 university students participated Wichadee's (2010) study and completed a questionnaire in order to express their perceptions about the characteristics of effective English teachers. English proficiency, pedagogical knowledge, organization and communication skills, and socio-affective skills were most-favored characteristics stated by the students. In Turkish context, a survey was conducted by Çelik, Arıkan, and Caner (2013) to 998 undergraduate students with the purpose of finding out Turkish students' perceived qualities of a successful EFL teacher. The results showed that the most-desired characteristics that the students identified are pedagogy-specific knowledge, personality traits, professional skills and classroom behaviors. In Iran, Ramazani (2014) looked for 384 university students' beliefs about the characteristics of qualified EFL teachers by using a self-report questionnaire. The results provide a detailed profile of students' beliefs with respect to teachers' pedagogical knowledge and their socio affective skills. More recently, in Vietnam, Nghia (2015) conducted a mixed-method study with 339 students with the aim of determining the most-desired qualities of English teachers in commercial English language centers that students perceived to develop their learning. The researcher used interviews and surveys to collect qualitative and quantitative data. Although twelve qualities were reported to influence their learning, Students valued English competence, teaching methods and socio-affective skills as the most important ones.

In conclusion, in EFL context, students can only interact with their English teachers in order to enhance their language skills; thus, providing opportunity for students being taught by an effective foreign language teacher is paramount. Although there are different

views for what qualities an effective EFL teacher should have, there seems to be a consensus on the underlying principles of qualified EFL teachers, such as having required

1- knowledge (e.g. literacy, numeracy, subject knowledge, knowledge of curriculum etc.)
2- skills (ability to perform a lesson, to use technology appropriatelly, to speak TL fluently, assessment literacy, classroom management, communication etc.)
3- attitudes (attitude about teaching, self-efficacy, motivational disposition etc.)

2.5 Teacher Competencies

Educational policies worldwide were influenced by two trends, namely the appearance of the global economy and the expansion of the Internet during the last quarter of the 20th century, whose outcomes emanated both on macro and micro level. Considering the macro-level, nations found themselves in a changing and competitive environment which forced them to adopt neo-liberal economic policies and to revise their educational systems (Güven, 2008). For the micro-level, thanks to technological development, the world has become a global village so people's communication, working, and learning habits have also changed. In parallel with these progresses, national and international agendas have started to focus on teacher education and teacher competencies with the aim of designing professional standards, strengthening teacher education and certification requirements, and transforming roles for teachers (Darling-Hammond, 2010a) since it is pointed out in various reports (e.g., OECD 2005; 2009) and research (e.g., Aaronson, Barrow & Sander, 2007; Chamberlain, 2013; Goldhaber, Grout & Huntington-Klein, 2014; Koedel, 2008; Goldhaber & Hansen, 2013; Jackson, 2012; Morgan & Hodge, 2014; Hanushek & Rivkin, 2010; Mentiş Taş, 2004; Rowan, Correnti & Miller, 2002; Sanders, Wright & Horn, 1997) that quality and effectiveness of teacher are crucial for enhancing teaching; raising workforce quality and status, hence sustaining growth; improving student achievement; affecting long-term student outcomes.

The main step in raising the quality of teachers is to determine the required teacher competencies. Competencies are defined as "knowledge, skills, attitudes, values, motivations and beliefs people need in order to be successful in a job." (Gupta, 2011 p. 24). Similarly, teacher competency is a concept used for knowledge, skills and attitudes necessary for the teaching profession (Van Der Schaaf, Stokking & Verloop, 2003).

Teacher competencies are expressed in different concepts such as teacher quality and standards in literature. Although there are differences in meaning among these concepts, they are often used interchangeably and it implies whether the teacher is qualified or not. Considering the evaluation process, standards are a must (Smith, 2005) and they provide foundation for teacher evaluation (Heneman III, Milanowski, Kimball, and Odden, 2006). Thus, standards could be seen as a tool of teacher evaluation; that is, they could be employed as criteria for judging and evaluating teachers (Bailey, 2006; Daresh, 2001 p. 281; Quirke, 2007, p. 2).

According to the Association of Teachers of English to Speakers of Other Languages (TESOL) (2010), five domains are vital for preparing English teachers: language, culture, instruction, assessment, and professionalism. Furthermore, there are a total of eleven standards within the five domains (please see Figure 1). *Language* and *culture* are considered as foundational knowledge domain in which teachers must be competent.
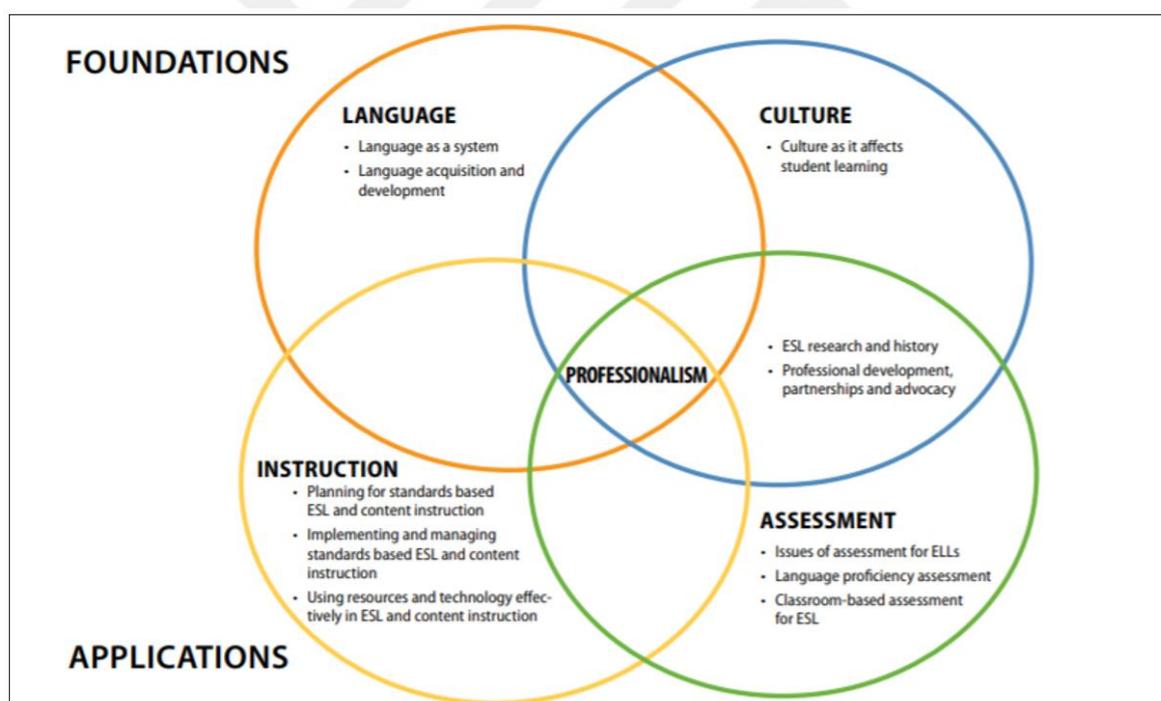
Figure 1. TESOL Standards. Taken from TESOL P–12 Teacher Education Program Standards (TESOL, 2010, p. 19)

The first one requires that English language teachers need to be proficient in the TL (Freeman & Freeman 2004) and theories about L1 and L2 acquisition (Gass & Selinker, 2008; Lightbown & Spada, 2006; VanPatten & Williams, 2007), while the latter deals with the characteristics of L2, stating the importance of English language teachers' awareness of

the construction of cultural identity, the occurrence of intercultural communication and acculturation processes, and whether cultural related educational practices debilitate or facilitate success in L2 acquisition (Cook, 2001; Díaz-Rico, 2008; Doughty & Long, 2003). Teaching competency is applied in the instruction and assessment domains. The previous one is about educational competencies. In other words, not only do English language teachers need to know how to select and implement the most effective types of curricula (Nation & Macalister, 2010; Richards, 2007) but they should be able to use appropriate instructional materials for the needs of their learners (McGrath, 2006). The assessment domain addresses teacher's assessment and evaluation literacy. Thus, the four outer domains form the content and skill areas of TESOL. How they intersect in each teacher's performance represents teacher's professionalism which requires the ability to keep up-to-date. Each of the domains is further divided into standards. There are a total of 11 standards and the standards themselves are divided into performance indicators. These indicators are to help one identify evidence of teacher's performance.

Four domains from which standards should be created were identified by the national teacher education accreditation agency in the United States, Council for the Accreditation of Educator Preparation (CAEP, formerly NCATE). These are Content Knowledge, Pedagogical Knowledge, Learning Environments, and Professional Knowledge. CAEP's "Content Knowledge" could include TESOL's Language and Culture domains while Instruction and Assessment would fit under "Pedagogical Knowledge," with Professionalism included in "Professional Knowledge." "Learning Environments" might include standards from any of the five TESOL domains (Kuhlman & Knezevle, 2013).

Considering the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR) (Council of Europe, 2011), it does not include standards per se. However, the way standards organize the Framework of Reference is essentially carried out by domains which include Structure, Knowledge and Understanding, Strategies and Skills, and Values. The Structure domain is about how a teacher preparation program is organized. The Knowledge and Understanding domain is similar to TESOL's Language and Culture domains, and CAEP's Content Knowledge whereas the Strategies and Skills domain is similar to TESOL's Instruction domain, and CAEP's Pedagogical Knowledge. Finally, it might be asserted that the Values domain included in TESOL's Professionalism domain.

Ministers of Education met in the Education Council (European Union 2007, 2008, 2009) with the intent of enhancing the whole teacher education process. In other saying, they put some issues on the agenda such as selection and recruitment of teachers, providing new teachers with initial teacher education of the highest quality, supporting beginning teachers systematically, and the relevance and quality of career-long opportunities for professional development. Even though it is not a complete list of the competencies that teachers require, Ministers have concurred that, as a minimum, teachers should have the subject(s) knowledge they teach, the required pedagogical skills to teach them, including teaching to heterogeneous classes, using ICT effectively. Additionally, according to Ministers, teachers need to have certain key professional values and attitudes such as engagement in research and innovation, autonomous learning, reflective practice, collaboration with colleagues and parents, and an involvement in the development of the whole school (European Commission, 2013a).

As mentioned above, the way domains and standards are labelled might vary from one country to another, but the rationale behind them is congruent. Competencies that teachers need to possess are constantly being discussed in Turkey as well as in the whole world. Turkish Ministry of National Education initiated a project to restructure education in 1999 and defined general and field-specific teacher competencies. In 2006, the national teacher standards were revised. As a result of this modification, field-specific teacher competency was prepared and put into effect with the Authority Approvals dated 04/06/2008 and 25/07/2008 (MEB, 2017b). However, the innovations in our education system and the new developments in the field of education created a need of updating generic teacher competencies once more. During the updating process, a large number of stakeholders (e.g. academicians, teachers, YÖK, ÖSYM, Vocational Qualification Institution, Board of Education, and other units of the MoNE) were consulted and the education and teaching related basic policy texts of international organizations (such as the Council of Europe, World Bank, ILO, OECD, UNESCO and UNICEF) and the competence documents of many different countries (e.g., USA, Australia Finland, France, Hong Kong) were examined. General qualifications and knowledge of field education have been added to the generic teacher competencies, rather than specifying a separate field-specific teacher competence so that a single, integral text has been created to cover the competencies of each teacher in his/her field (MEB, 2017b). Consequently, the new generic teacher competencies consist of three interdependent and complementary fields (namely,

professional knowledge, professional skills, attitudes and values) together with 11 sub-competencies (please see Figure 2).

A	Professional Knowledge	B	Professional Skills	C	Attitudes and Values
A1.	Content Knowledge	B1.	Planning of Education and Teaching	C1.	National, Moral and Universal Values
	She/he has an advanced and critical perspective on theoretical, methodological and factual knowledge in his/her subject field.		She/he plans education and teaching processes effectively.		She/he observes national, moral and universal values.
A2.	Pedagogical Content Knowledge	B2.	Creating Learning Environments	C2.	Approach to Students
	She/he has a good knowledge of the curriculum and pedagogical content knowledge of her/his subject area.		She/he prepares appropriate teaching materials and builds an healty and safe learning environments, where effective learning can be achieved for all students.		She/he has an attitude that supports the development of students.
A3.	Knowledge on Legislation	B3.	Managing the Teaching and Learning Process	C3.	Communication and Cooperation
	As an individual and teacher, she/he conducts her/himself according to the legislation related to her/his duties, rights and responsibilities.		She/he manages the teaching and learning process effectively.		She/he establishes an effective communication and cooperation with students, colleagues, families, and other educational stakeholders.
		B4.	Assessment and Evaluation	C4.	Personal and Professional Development
			She/he uses the methods, techniques and tools of assessment and evaluation that fit for purpose.		By carrying out self appraisal she/heparticipates in personal and professional development activities.

Figure 2. General Competencies for Teaching Profession. Taken from MEB (2017b). Directorate General for Teacher Training and Development. General Competencies for Teaching Profession (p. 14).

Defining teacher competencies is a vital issue, because they can be utilized in many areas, such as practicum, teacher employment, continuing professional development, and many other areas (please see Figure 3.). This dissertation accepts teacher competencies as a part of theoretical framework and benefits from it when defining the characteristics of effective teachers that should be considered during the teacher evaluation/recruitment process. Moreover, the present research considers teacher competencies under three categories, which is similar to the classification of the MoNE. That is, an effective/quality EFL teacher (a) should have the required knowledge of subject, curriculum, technology and assessment

literacy (knowledge-related competencies), (b) needs to perform lesson well, use technology appropriately, use assessment tools according to learners' needs, and speak English fluently (skill-related competencies), and (c) has positive attitude about teaching, loves his/her students, and is motivated to develop himself/herself (attitude-related competencies).
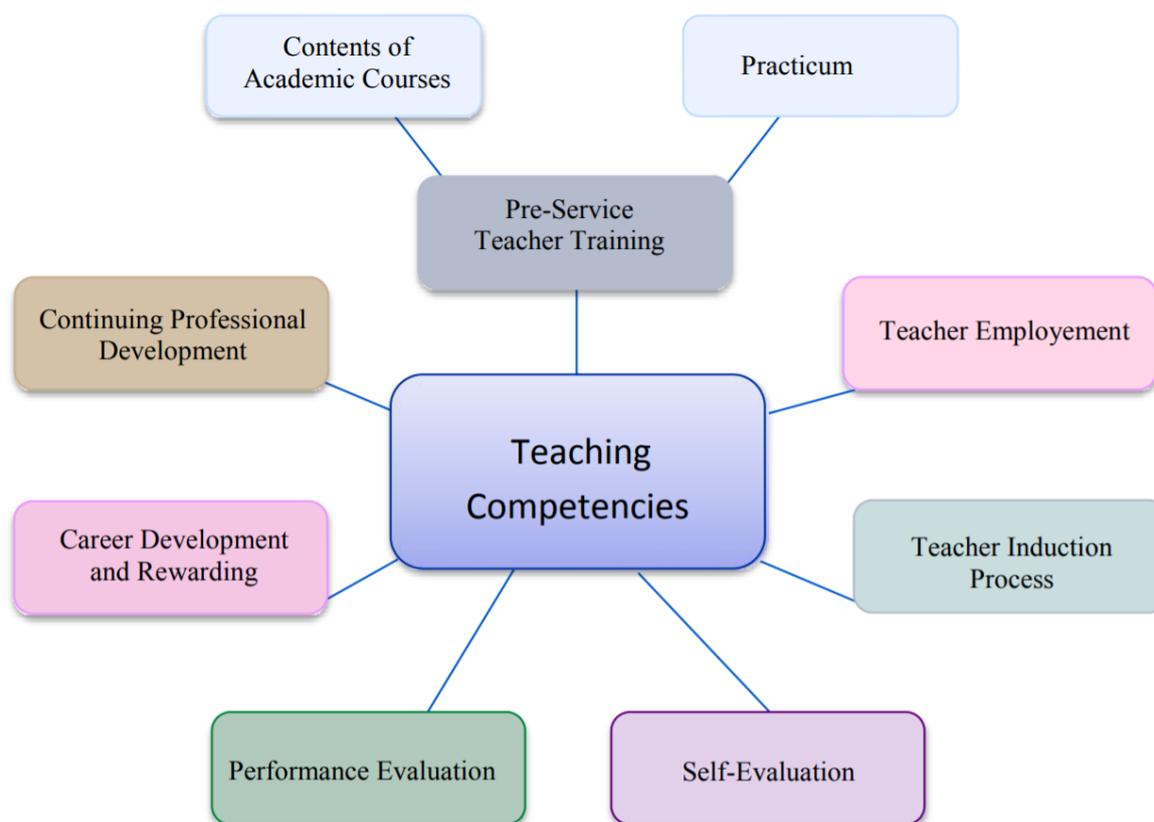
Figure 3.Areas of Using the Teacher Competencies. Taken from MEB (2017b). Directorate General for Teacher Training and Development. General Competencies for Teaching Profession (p. 15).

2.6 English as a Foreign Language in Turkey

The influence of EFL is dominant in all aspects of life, from economics and politics to education in Turkey. Being in Kachru's (1992) expanding circle, Turkey's sole official language is Turkish. Indeed, Turkish people learn English for instrumental reasons of education (Sebüktekin, 1981). As Doğançay-Aktuna (1998) states, Turks learn and teach English for the aim of establishing and maintaining international communication links and this communication mostly occurs with non-native English-using groups, as is the case with many EFL contexts.

Considering the development of foreign language education in Turkey, Kirkgöz (2007) presents three phases: 1) historical recognition of English and its spread in the country, 2) the 1997 reforms, and 3) changes introduced in ELT in 2005, with the aim of adapting it to EU standards. Despite the fact that the introduction of English in Turkey is said to be on the Tanzimat Period, which is known as the sign of Westernization in education (Kirkgöz, 2005), the period in which these effects were felt in the real sense seems to have started in the 1950s as a result of the impact of increasing American economic and military power (Demircan, 1988; Doğançay-Aktuna, 1998). More precisely, Doğançay-Aktuna (1998) asserts that two phases are crucial regarding the spread of English in Turkey: 1950s and 1980s.

Once the Republic of Turkey was established, modernization and Westernization movements made Turkey much closer to the Europe and especially to the French language. However, in the first phase (1950s), as pointed out by Doğançay-Aktuna (1998), English started to compete with French, and it became more preferable in various domains such as diplomacy, education system, arts and literature due to improving trade relations and technological developments. Anatolian school (Maarif Koleji) was the first state-funded English-medium secondary school and opened in 1955 for providing secondary education. Having a four-year period of education, such kind of school differed from other education institutes in providing mainly English education in the first year. In order to be educated at those schools, students required to pass a centralized test. Likely, the instruction for teaching mathematics and science was English in these schools. However, in 1992, English was not given as the language of instruction due to lack of competent teachers in terms of teaching content courses in English. The popularity of Anatolian schools spread country-wide so that the number of these schools raised day by day: In 1955, five schools were in operation; 12 in 1974 (Kırkgöz, 2005); 22 by the 1987-1988 school year (Demircan, 1988); 415 by 2004-2005 school year (Kirkgoz, 2007). Considering higher education, in 1956, Middle East Technical University, which is based on the American model, is the first established one among the English-medium universities. In 1981, English-Medium Education in the universities became a national-wide issue so that, in 1990, private universities could establish EME thanks to the Higher Education Law, which contributed to the expansion of the country's provision of English-medium higher education (Kırkgöz, 2005). According to Akunal (1992), it is career or academic purposes that guide students' attitudes toward EME.

Following the second phase (1980s), Turkey has tried to open its economy to global competition and to obtain advanced technologies. In this sense, the relationship with the international community in general and with the United States in particular has become more important, and the priority position of English in foreign language education has also been strengthened. In other saying, these developments have caused Turkey to be influenced by the globalization forces via English (Friedman, 1994; Robins, 1996). Due to growing popularity of English, English-medium schools increased in the number. For example, while the number of English-medium secondary schools were 193 (103 private, 90 state-owned) by 1987–1988 school year, the number raised to 1065 (650 private secondary schools; 415 Anatolian high schools) by the 2004–2005 school year (MEB, 2006 as cited in Kırkgöz, 2007). In addition to these schools, the MoNE put a new law into force which paving the way for opening 'Super English Language High Schools'. They were similar with Anatolian high schools in offering one-year English language education. However, entering these schools is possible for those students whose score is at least 4.0 out of 5.0.

Taking into consideration the Higher Education level, in response to the increasing demand for the medium of English language at tertiary level, Bilkent, the first private university, was established in 1983, in Ankara with the Higher Education Law which let private universities offer English-medium education. In universities providing English-medium instruction, there is a one-year intensive English education for those students who do not reach the threshold level. In Turkish medium universities, English language is a compulsory subject, and taught generally during first-year. The aim is to keep their general English knowledge alive so that they could follow and understand subject specific English publications. Currently, there are 129 state universities, 72 non-profit foundation universities, 5 Independent Post-Secondary Vocational Schools in Turkey (YÖK, 2018a). The language of instruction is English at some universities such as Middle Eastern Technical University, Boğaziçi University, and Bilkent University. Furthermore, universities offering certain programs in English or bilingual programs are available. Entrance to these types of programs requires a one-year prep class for those who have insufficient foreign language proficiency. Most non-profit foundation universities are also English-medium. However, according to YÖK, the English-medium universities ought to perform at least three major requirements: having sufficient teachers being able to teach in English, establishing a foreign language center so that it could offer English for Academic

Purposes (EAP) courses for each department, and the required course materials and other resources. There are also some universities which are partly English-medium with 30% to 40% of the subject courses taught in English (ÖSYM, 2001 as cited in Kırkgöz, 2005).

Prior to 1997, the time for students to learn English started at secondary school, in grade 6. In parallel with the many countries in Europe (Finland, France, Italy etc.) and Asia (China, India etc.), English became a compulsory school subject at the primary level with the 1997 reform in Turkey, and the duration of primary education was raised from 5 to 8 years so that students could start to learn English in grades 4 and 5, and the concept of the communicative approach into ELT (Kirkgöz 2005). Contrary to the traditional teacher-centered view, learning has become student-centered. Students are indicated to take an active role while the role of teachers is specified as facilitator. The introduction of TEYL let young learners learn English in grades 4 and 5

It was aimed to raise their foreign language awareness, promote a constructive attitude towards English, increase their interest and motivation, providing games so that young learners enjoy lessons while learning English, provide them with contextualized and meaningful learning activities, and help them develop suitable strategies (Kocaoluk 6 Kocaoluk, 2001) while enhancing the pupils' basic communicative skills are objectives for grades 6 to 8. In order to meet the need of TEYL teachers, relevant courses were integrated to the undergraduate level and ELT Departments started to educate new teachers in order to make student teachers be familiar with TEYLs

Nowadays, the only compulsory language to be taught at all levels of education is English. As part of the new reform in education, the 4+4+4 model was accepted in 2012. In this model English language learning starts at the age 7, Grade 2 as of the 2013–2014 Academic Year and children are taught in English for 2 hours per week since Grade 2, increased to 4 hours each week in their last four years of school. The curriculum of some schools includes German and French as elective subjects. In addition to being a compulsory school subject, English is "also a 'must' for entry and advancement in competitive jobs in urban Turkey" (Dogançay-Aktuna, 1998 p.33). According to some Turkish scholars (Alptekin & Tatar, 2011; Kırkgöz, 2009), Turkey's resolution of English as a compulsory foreign language in its education policies indicates Turkey's eagerness for complying with standards of the European Union (EU) and for being a member of European union in the future.

2.6.1 Foreign Language Policy in Turkey

Global demands have raised a tendency, for Turkey in particular and for many countries in general, to make necessary adjustments in their FL teaching policies (Ford, 2014 as cited in Gürsoy, Korkmaz & Damar, 2017). Since this education policy reforms usually occur as the introduction of English into the primary curriculum around the world (Garton, 2014; Nguyen, 2011), various countries have officially started to reduce the age of second/foreign language teaching and/or learning to 7 years old or below (Baldauf, Kaplan, Kamwangamulu & Byrant, 2011; Enever & Moon, 2009; Enever, 2011; Eurydice, 2012; Mourão & Lourenço, 2015).

Tanzimat Period is recognized as the date when English was introduced to the Turkish educational system (Kirkgöz, 2007) and since then English language teaching has been influenced by political and socioeconomic factors which forced English language teaching be exposed to many changes. Following World War II, English has become popular in Turkey due to its attempts towards modernization so as to compete internationally. As a result of these factors, together with globalization, English language policy has entered into various stages of change at all levels of Turkish education. In 1983, for the purpose of regulating foreign language teaching policy, the Foreign Language Teaching and Learning Act [Yabancı Dil Eğitim ve Öğretim Kanunu] was introduced to the Turkish education system. Within the context of this Act, the MoNE has started to be responsible for the administration of the English language curriculum at public and private schools in Turkey. Following the foundation of the Republic of Turkey, the foreign language reforms could be examined under three phases, namely the Ministry of Education Development Project (1994-1997), 2006 Education reforms, and 2012 Education reforms (Kırkgöz, 2016).

During the 1994-1997 periods, the MoNE and the YÖK initiated the Ministry of Education Development Project which was financed by the World Bank (YÖK, 1998) with the aim of redesigning the teacher-training departments and enhancing the quality of teacher training. Correspondingly, 'Eight-Year Uninterrupted Compulsory Primary Education Law' numbered 4306, on August 16, 1997 was enacted at the Turkish Grand National Assembly and entered into force after published in the Official Gazette dated 18 August 1997 and numbered 23084 (MEB, 1997). In this sense, in line with the policy change in 1997, English has been started to be taught as a compulsory school subject at the primary level, and duration of primary education was raised from 5 to 8 years (Gürsoy, Korkmaz &

Damar, 2013; Kirkgöz, 2007). Rearranging weekly classroom hours for English, the MoNE has determined two classroom hours for 4th and 5th Grades while four classroom hours for 6th to 8th Grades. One of the innovations of the 1997 education reform is to place the idea of 'communicative approach' into curriculum due to the introduction of English to the 4th Grade. Within the scope of this innovation, the teacher's role has been recognized as facilitator and students become more active in the learning process. Thanks to the In-service English Language Teacher Training and Development Unit (INSET), public primary and secondary English language teachers have been supported with seminars and in-service training workshops in order to make them be familiar with curriculum innovations. Likely, the collaboration between the MoNE and the English language Teachers' Association in Turkey (INGED), the British Council (BC), and the United States Information Agency (USIA) aims to help facilitate the implementation process of the curriculum reform (Kirkgöz, 2009).

Compared to the previous one, 2006 education reforms seem to have more comprehensive regulations such as multiple intelligence art and craft activities, and content and language integrated learning. Meanwhile, authentic assessments such as portfolio were offered to be applied in addition to traditional 'paper and pencil' tests (MEB, 2006). Considering the primary school, steps were taken to overcome deficiencies in the existing system of 1997 educational reforms and the revised English Curriculum for Primary Schools prepared for 4th to 8th Grades has been in operation since 2006 (MEB, 2006). Moreover, the course books for 4th and 5th Grades were updated in line with the revised foreign language education curriculum. At the higher education level, there seems to be an attempt to train well-qualified teachers in accordance with the needs of the country. In order to perform this goal, it is aimed to enhance student teachers' critical thinking, problem solving, and lifelong learning skills as well as raising the number of general culture courses, offering opportunities for selective courses, and introducing new courses into the curriculum (Kirkgöz, 2016).

For the upper secondary school (formerly known as high school), the curriculum of Anatolian high schools was dramatically affected by the foreign language education policy in 2002. The MoNE decided that Turkish would be the sole language of instruction (previously mathematics and science were taught in English) in Anatolian high schools for two reasons: lack of qualified teachers teaching mathematics and science in English, and

the claim that the centralized university entrance examination provided disadvantage to graduates of those schools since it was conducted in Turkish (Doğançay-Aktuna & Kızıltepe 2005). Additionally, for achieving language teaching standardization in all types of schools, the one-year intensive English preparation class in Anatolian, private and Super English Language High Schools was abolished. Currently, upper secondary schools offer ten courses of English per week for the first year whereas four hours for the rest years.

The 2012 education reform has raised the duration of compulsory education from 8 to 12 year as it is formulated as 4+4+4. The first 4 is used for primary education while the second one is called lower secondary education, and finally, upper secondary education is the last one (OECD, 2013a). One of the innovations that the 2012 education reform provides for foreign language education curriculum is the starting ages for foreign language education. In other words, the age for learning English in public schools is lowered to 6 years old (the 2nd grade) (MEB, 2013). For Grades 2 and 3, the importance of speaking and listening is underlined while other skills (reading and writing) begin to be implemented later since students at these ages are not ready for reading and writing.

The new program is based on the underlying principles of CEFR (Council of Europe, 2001), initiated by the European Council, expressing the objectives set in the teaching of foreign languages. Besides, issues such as appropriate tasks for students, learner autonomy, self-assessment, and communicative competence are included in the program (Kirkgöz, 2016). It is aimed to start teaching English at 2nd Grade from A1 (breakthrough) to B1 (threshold) at the end of upper secondary school.

2.6.2 English Language Teacher Education in Turkey

The importance of teacher education is indicated as a powerful predictor of teacher quality and student achievement (see, for example, Darling-Hammond & Youngs, 2002; Goldhaber & Brewer, 2000; Hanushek, 2012; Laczko-Kerr & Berliner, 2002; Rivkin, Hanushek & Kain, 2005). It is teacher education programs that should prepare student teachers to their future roles (Allen, 2009), and that carry responsibility of educating future teachers who will shape the future of the country (Keçik & Aydın, 2011) since it is "the first entry point to the teacher professional career" (Musset, 2010, p.16). In this sense, various models of teacher education have been applied since 1923, the establishment of the

Republic of Turkey. The history of teacher education of Turkey dates back to 1924 with the 'Law of Unification of Instruction' (Çakıroğlu & Çakıroğlu, 2003). However, the two major changes occurred nearly a half century later: The Basic Law of National Education in 1973, shaping the general framework of the Turkish national education system, and the reorganization of teacher education through the YÖK in 1981, which had been controlled by the MoNE till then (Yavuz & Topkaya, 2013).

Historically, after the establishment of the Republic of Turkey, the first foreign language department is the French language department established in Gazi Institute, 1941, followed by English language department in 1944, and finally, in 1947, German language department was opened (Demirel, 1991). A three-year training program was determined in the 1967-1968 Academic Year but the duration was increased to 4 years in the 1978-1979 Academic Year. Then, Gazi Institute was renamed as Gazi High Teacher School whose foundation led to the establishment of Gazi University and the Faculty of Education in 1982 (Akyüz, 2009). For the first time, the title of foreign language teaching departments was used with the foundation of Gazi and Çapa Institutes (Güneş, 2009). Although the foreign language teaching departments increased in number in 1960s, the establishment of four-year foreign language programs occurred only after 1970; however, this increase was about their quantity not quality and did not contribute to training qualified student teachers (Demircan, 2001).

Establishing standardization in certifying teachers of all subject areas, the Higher Education Reform enabled the Faculties of Education to carry out the pre-service teacher education in universities. Yet, as seen in the reports of the YÖK (YÖK, 1998; 2004), it is impossible to fulfil the demands for qualified teachers because of some gaps in the Faculties of Education after 1982. According to Şallı-Çopur (2008), one of these shortcomings is the fact that the academic background of their teaching staff determined the structure of the departments, focusing on the theoretical aspects of the subject areas and ignoring training competent teachers in their fields. This results from the decision of YÖK for transferring instructors of the Faculties of Letters, Arts and/or Science to the Faculties of Education because both subject matter knowledge and pedagogic knowledge and its application are of vital importance in the pre-service teacher education while the theoretical aspects of the subject areas are overemphasized in others.

The introduction of the YÖK in 1981 has constituted a top-down oligarchic system of higher education (Mahalingappa & Polat, 2013) creating a 'unified' model of higher education and making universities to be responsible for the whole academies and teacher training institutions (Güven, 2008). Therefore, this kind of higher education system has been implementing a high level of control over all elements related to teacher education. In a nut shell, the YÖK has started to control all staff, logistics, and transactions of Turkish universities since 1981. From then on, the curricula of teacher education have faced some levels of change. However, as alleged by Yavuz and Topkaya (2013), rapid changes appeared in initial teacher education in the 1990s. To make teacher education more effective, the duration of the primary teacher education was increased from two to four years in 1991. Following this reform, the period of compulsory education was extended from five to eight years in 1997, and students began to learn English language at 4th Grade. This change called for a revamp of English language teacher education (ELTE) programs for graduating more qualified English language teachers. Yet, the other side of the coin shows that 'aims' of the YÖK and 'changes' proposed by the YÖK are not always in harmony. For instance, the revision of the ELTE programs created a gap in teacher education programs in terms of qualified field-specific instructors. Instead of employing people who can do the job, the YÖK (2007) transferred instructors of the Faculties of Science and Letters (Güven, 2008) to Faculties of Education, causing the Turkish higher education system to experience a long-run counterproductive tradition which did not take into account the needs of the program while selecting courses but the expertise of the instructors, and as a consequence of this tradition, courses were taught by unqualified faculty members (Mahalingappa & Polat, 2013).

After the 1997 Education reform when the duration for the compulsory education was determined to be eight years, in response to a growing teacher demand (YÖK, 1998), YÖK collaborated with MoNE to reshape the English language teacher education programs with the aim of equipping student teachers with the basic teacher competencies (Yıldırım & Ok, 2002). Thereupon, with the financial support of the World Bank, YÖK initiated the Pre-service Teacher Education Project so as to enhance the quality of future teachers via developing the curriculum of the pre-service teacher education programs (YÖK, 1999). With this new project, in addition to redefining teacher competencies and redesigning teacher qualification courses, 11 compulsory courses, which were crucial for certifying

graduates as language teachers, were included in the new foreign language teacher education program (Şallı-Çopur, 2008, pp. 3-4).

The new reform put emphasis on instructors' professional development, and it put the training of faculty members in the US and Europe on the agenda, focusing on revising teaching techniques and methodologies (YÖK, 2000). However, Turks do not change their attitudes and behaviors rapidly (Sands & Stevens, 2004). In response to changing demands of education in Turkey, the current curricular framework has gone under some reforms in order to improve the quality of teacher preparation programs (Kirkgoz, 2007). Based on these reforms, in addition to redesigning foundation courses, new lessons were introduced, namely classroom management, instructional planning, and computer and instructional technology (Buyukkantarcioğlu 2004; Çakıroğlu & Çakıroğlu, 2003; YÖK, 2007; Şimşek & Yıldırım, 2001). Furthermore, the previously ignored courses such as subject-based teaching experiences and methodology started to gain more importance.

Very recently YÖK has revised the teaching undergraduate program will come into effect introduced in 2018-2019 academic year (see Figure 4.). The title, description, weekly hours and credits of the courses have changed. Courses are collected under three categories, namely vocational knowledge, field education and general culture. Especially in teaching programs related to basic education, emphasis has been given to field education and teaching instead of transferring only field knowledge. Programs related to this field have been structured around basic skills, attitudes, value and orientation. The new program has four new important features:

- The course "Morals and ethics in education" has become compulsory,
- Foreign language lessons have been associated to everyday life,
- Courses related to field teaching have gone up to fourfold, and
- Total course credits have reduced. (please see YÖK, 2018b for more details)

I. Semester

	Courses	T	U	K	ECTS
VK	Introduction to Education Science	2	0	2	3
VK	Educational Sociology	2	0	2	3
GC	Principles of Atatürk and History of Revolution I	2	0	2	3
GC	Foreign Language I	2	0	2	3
GC	Turkish Language I	3	0	3	5
GC	Information Technologies	3	0	3	5
FE	Reading Skills I	2	0	2	2
FE	Writing Skills I	2	0	2	2
FE	Listening and Phonetics I	2	0	2	2
FE	Oral Communication Skills I	2	0	2	2
	Total	22	0	22	30

II. Semester

	Courses	T	U	K	ECTS
VK	Educational Psychology	2	0	2	3
VK	Philosophy of Education	2	0	2	3
GC	Principles of Atatürk and History of Revolution II	2	0	2	3
GC	Foreign Language II	2	0	2	3
GC	Turkish Language II	3	0	3	5
FE	Reading Skills II	2	0	2	2
FE	Writing Skills II	2	0	2	2
FE	Listening and Phonetics II	2	0	2	2
FE	Oral Communication Skills II	2	0	2	2
FE	Structure of English	2	0	2	2
	Total	21	0	21	30

III. Semester

	Courses	T	U	K	ECTS
VK	Instructional Technologies	2	0	2	3
VK	Techniques and Principles in Teaching	2	0	2	3
VK	Elective Course I	2	0	2	4
GC	Elective Course I	2	0	2	3
FE	Elective Course I	2	0	2	4
FE	Approaches of English Language Teaching & Learning	2	0	2	3
FE	English Literature I	2	0	2	4
FE	Linguistics I	2	0	2	3
FE	Critical Reading and Writing	2	0	2	3
	Total	18	0	18	30

IV. Semester

	Courses	T	U	K	ECTS
VK	Turkish Educational History	2	0	2	3
VK	Scientific Research Methods	2	0	2	3
VK	Elective Course II	2	0	2	4
GC	Elective Course II	2	0	2	3
FE	Elective Course II	2	0	2	4
FE	Programs of English Teaching	2	0	2	3
FE	English Literature II	2	0	2	4
FE	Linguistics II	2	0	2	3
FE	Language Acquisition	2	0	2	3
	Total	18	0	18	30

V. Semester

	Courses	T	U	K	ECTS
VK	Classroom Management	2	0	2	3
VK	Morals and Ethics in Education	2	0	2	3
VK	Elective Course III	2	0	2	4
GC	Elective Course III	2	0	2	3
FE	Elective Course III	2	0	2	4
FE	Teaching Foreign Language to Young Learners I	2	0	2	5
FE	Teaching of English Language Skills I	2	0	2	5
FE	Literature and Language Teaching I	2	0	2	3
	Total	18	0	18	30

VI. Semester

	Courses	T	U	K	ECTS
VK	Testing and Evaluation in Education	2	0	2	3
VK	Turkish Education System and School Management	2	0	2	3
VK	Elective Course IV	2	0	2	4
GC	Elective Course IV	2	0	2	3
FE	Elective Course IV	2	0	2	4
FE	Teaching Foreign Language to Young Learners II	2	0	2	5
FE	Teaching of English Language Skills II	2	0	2	5
FE	Literature and Language Teaching II	2	0	2	3
	Total	18	0	18	30

VII. Semester

	Courses	T	U	K	ECTS
VK	Teaching Practice I	2	6	5	10
VK	Special Education and Inclusion	2	0	2	3
VK	Elective Course V	2	0	2	4
GC	Community Service Practices	1	2	2	3
FE	Elective Course V	2	0	2	4
FE	Preparing Course Content in ELT	3	0	3	3
FE	Translation	3	0	3	3
	Total	15	8	19	30

VIII. Semester

	Courses	T	U	K	ECTS
VK	Teaching Practice II	2	6	5	15
VK	Guidance in Schools	2	0	2	3
VK	Elective Course VI	2	0	2	4
FE	Elective Course VI	2	0	2	4
FE	Preparing exams in ELT	3	0	3	4
	Total	11	6	14	30

	Courses	T	U	K	ECTS	HOUR	%
VK	Vocational Knowledge	44	12	50	91	56	34
GC	General Culture	26	2	27	42	28	18
FE	Field Education	71	0	71	107	71	48
	Total	141	14	148	240	155	100

*Figure 4.*English Language Teaching Undergraduate Program. Adapted from YÖK (2018b, p. 2).

Becoming an English language teacher in Turkey requires studying in ELTE programs which take four years (in addition to four years, the program can be extended one year for intensive English preparatory). However, students graduating from other faculties (e.g., the Faculties of Science and Letters) and various departments (such as English Language and Literature, American Culture and Literature, Department of Translation and Interpreting or Department of English Linguistics) have an opportunity to become English teachers via pedagogic formation certificate. Generally, assessing courses consists of mid-term and final exam which are mostly in open-ended format. There are both compulsory and elective courses in ELTE programs in which the selected course books are commonly prepared in Britain or the U.S.A. In the first year of this new program, it is aimed to enhance student teachers' specific language competencies and to help them become competent in English. The second year offers various courses which are professionally oriented so as to equip them with not only pedagogical (teacher education) and general knowledge but also with linguistic competence (field knowledge). In the third year, as it is clear in Figure 4, the student teaches are asked to be aware of the basic concepts about classroom management; physical, social and psychological dimensions of the class; class rules and discipline; teaching English to young learners; testing and evaluation in education; Turkish education system and school management; literature and language teaching; and the importance of measurement and evaluation in education. Finally, in last year, the student teachers are provided with observation (Teaching Practice I) and experience (Teaching Practice II) of actual teaching practice. In addition to observation and experience, student teachers take some cources such as special education and inclusion, community service practices, prepare exams in ELT, prepare course content in ELT, and translation.

Having different categorization, the competencies of Turkey's current ELTE programs resemble those of TESOL. Unlike TESOL's five-domain categorization, the YÖK (2010) provides three domains: (1) Language Teaching Subjects; (2) General Culture; and (3) Pedagogical Formation. According to Çelik (2011), the YÖK needs to give due importance to the professionalism and culture domains of TESOL, which are underrated in the YÖK standards. Besides, this program depends on CEFR and on the European Portfolio for Student Teachers of Languages (EPOSTL). Although ELTE programs have flexibility in curriculum design, they are obliged to pursue these basic requirements. In other saying, even if the YÖK give universities authorization to decide the course books that will be used, how courses will be selected, and how the examinations will be administered to

English student teachers, all ELTE programs must follow the basic requirements determined by the YÖK.

CHAPTER III

METHODOLOGY

Benefiting from both quantitative and qualitative methods of research, this study is exploratory sequential mixed methods design and aims at suggesting a teacher assessment and evaluation (TAE) model for EFL contexts, in which teacher recruitment is centralized and based heavily on multiple choice paper-pencil exam, with the aim of making the teacher evaluation process more reliable, fair and valid, by applying the perceptions of education stakeholders through qualitative and quantitative data collection tools. To this respect, this chapter begins by providing an overview of the methodology utilized to conduct this study. First, the sequential exploratory mixed methods research design is described in Part 3.1, the 'Research Design'. Next Part 3.2 'Universe and Samples' illustrates the selection process of the samples. This is followed by Part 3.3 'Data Collection Techniques', which presents various stages of the research such as pilot studies, the instruments developed and used as data collection tools and finally the procedure followed in this research study. The chapter concludes with 'Data Analysis', Part 3.4, which presents how quantitative and qualitative data were analyzed.

3.1 Research Design

The purpose of this study is to develop a fair, reliable and valid teacher assessment and evaluation model for EFL contexts, in which recruitment is centralized for English language teachers and based heavily on multiple choice paper-pencil exam, so the model can contribute to the Ministry of National Education in recruiting effective or qualified English teachers. In this sense, firstly, documents and literature were applied. then, the

perceptions of education stakeholders were applied via semi-structured interviews and focus groups interviews. Although a new model was developed by the help of content analysis, in order to generalize the findings, an attitude scale was developed by the researcher since the literature did not provide any attitude scales or surveys regarding English teachers' attitudes towards teacher evaluation or recruitment model in EFL contexts, in which the teacher recruitment is actualized centrally. Using such instruments is of importance, because they are disseminated to large numbers of participants "to ascertain whether qualitative research findings held over large groups of respondents." (Hitchcock & Nastasi, 2011; Hitchcock et al., 2005, 2006; Nastasi, Hitchcock, Burkholder, Sarkar, & Varjas, 2007 as cited in Hitchcock, et al, 2015 p.5). To this end, benefiting from strengths of both quantitative and qualitative methods of research, this study is exploratory sequential mixed methods design (Creswell, 2012; Creswell & Plano Clark, 2011).

A number of advantages of mixed methods, also called "the third methodology or research (along with qualitative and quantitative research)" (Johnson, Onwuegbuzie & Turner, 2007, p.129), has been underlined in literature. For example, Johnson et al (2007) articulate that combination of qualitative and quantitative methods provides a better understanding of the research phenomenon than either of each alone. Some researchers (for example, Johnson & Onwuegbuzie, 2004) advocate that the aim of combining quantitative and qualitative methods is to benefit from the strengths of each tradition of research. In addition, five rationales are listed by Johnson et al. (2004) for the necessity of conducting mixed methods research:

1. triangulation of data,
2. complementarity of data,
3. initiation arising out of the need for more clarification,
4. development; and
5. expansion of research by using multiple strands.

According to Tashakkori and Teddlie (2003), mixed methods can respond to research questions that other methodologies cannot answer. Another advantage of mixed methods is triangulation which "entails the use of multiple, independent methods of obtaining data in a single investigation in order to arrive at the same research findings" (Mackey & Gass, 2005 p. 181). In this research, the data were collected through a variety of means in an integration of some qualitative and quantitative methods. The aforementioned explanations

make it clear that mixed methods research provides opportunities to researchers to make stronger inferences via the combination of both qualitative and quantitative methods. That is, the burden on the shoulders of researcher may get lighter when approaching research question using strengths of mixed methods research approach.

Greene, Caracelli, and Graham (1989) presented triangulation, complementarity, development, initiation, and expansion as the rationale for conducting mixed methods research. Later, Bryman (2006) expressed that there might be various reasons for using mixed methods and listed sixteen reasons, including triangulation, offsetting weakness, completeness, process, answering different research questions, explanation, unexpected results, instrument development, sampling, credibility, context, illustration, utility, confirming and discovering, diversity of views, and enhancement. Collins, Onwuegbuzie, and Sutton (2006) reported four rationales in order to utilize mixed methods, namely participant enrichment, instrument fidelity, treatment integrity, and significance enhancement. Participant enrichment underlines the importance of the rationale of optimizing the sample. To achieve it, the researcher mix both qualitative and quantitative data. Increasing the number of participants is one way to optimize a sample. Instrument fidelity referred to "steps taken by the researcher to maximize the appropriateness and/or utility of the instruments used in the study" and "the investigator could assess the validity of information (i.e. qualitative or quantitative) yielded by the instrument(s) as a means of putting the findings in a more appropriate context." (Collins, et al, 2006, p.76). Treatment integrity refers to the degree to which intervention is implemented as intended. Finally, significance enhancement represents the researcher's mixing quantitative and qualitative techniques in order to enhance interpretations of data.

Researchers have identified six most commonly used mixed method research designs: (a) the convergent parallel design, (b) the explanatory sequential design, (c) the exploratory sequential design, (d) the embedded design, (e) the transformative design, and (f) the multiphase design (Creswell, 2012 pp.540-7). The six mixed methods designs are categorized based on four criteria: (a) what priority and weight the researcher give to the quantitative and qualitative data collection, (b) the sequence of collecting quantitative and qualitative data, (c) how to analyze the data, and (d) where in the study the researcher mix the data (Creswell, 2012, p.540). In an exploratory sequential design, firstly, the qualitative data is collected and analyzed, and categories are used for developing a possible

quantitative tool to further explain the research problem (Creswell & Plano Clark, 2011; Onwuegbuzie, Bustamante & Nelson, 2010; Teddlie & Tashakkori, 2009). Researcher applies an exploratory sequential mixed methods design aims to collect quantitative data for the explanation of relationship occurred in qualitative data and develop an instrument (Creswell, 2012, p.543)

3.1.1 Qualitative Research Design

Qualitative research methodologies are utilized to scrutinize why or how a phenomenon occurs, obtain rich information, develop a theory, whereas quantitative methodologies are used for causality and generalizability of effect (Fetters, Curry & Creswell, 2013). Mixed methods is a research methodology that benefits from the strengths of both qualitative and quantitative research by combining them within a single study in the full process of research, including philosophical assumptions, research design, data collection methods, data analysis methods and the interpretations of results (Creswell, 2009, p. 3; Creswell & Plano Clark, 2011; Johnson & Onwuegbuzie, 2004; Leech & Onwuegbuzie, 2009; Tashakkori & Teddlie, 2003, 2010; Teddlie & Tashakkori, 2009).

The qualitative design used in this thesis is based on grounded theory which originates in the work of Glaser and Strauss (1967). It is called grounded theory because the occurring theory is grounded in empirical data and derived from inductive analysis (Strauss & Corbin, 1990). Thus, it is not "the formulation of some discovered aspect of a preexisting reality 'out there'" (Strübing, 2002, p. 321). In the broadest sense, grounded theory is the generation of theories from data and simply the discovery of emerging patterns in data (Walsh, Holton, Bailyn, Fernandez, Levina and Glaser, 2015). The advocates of this theory are of the opinion that an approach dealing with the concerns of the participants of the research is more successful in illustrating the world and priorities of participants, when compared those start with the prejudice towards the world and its meaning (De Burca & McLoughlin, 1996).

Apart from statistical analysis, applying public perception is kind of validation because understanding public perceptions becomes important when adopting a lens of consequential validity (Hitchcock, et al, 2015 p.5). Applying their perceptions to learning how they give the meaning to their realities may provide invaluable understanding. In

addition to this, this process might display how people differ on the same subject (for example, some members of a population might endorse recruitment model while others no) or a person's both negative and positives views of the same subject (a person might support ÖABT but find it inadequate). With this in mind, most qualitative studies contain the use of some combination of multiple data collection tools, including interviews, focus groups, observations, and archival analyses. It is beyond the scope of this article to report other choices. Rather, in addition to documents, this dissertation took into account the likely advantages of conducting a series of semi-structured interviews and focus groups with different stakeholders, including student teachers, in-service teachers, teacher trainers, and policy makers, about the process of teacher recruitment.

3.1.2 Quantitative Research Design

The quantitative methods have several strengths and it based on experiments and surveys to gather measurable data for statistical processes (Creswell, 2003). One of the major advantages is that quantitative methods enable researchers to generalize the results to larger populations. As for this research, applying merely a quantitative design may give us some idea about whether the results can be generalized to wider population. Quantitative research methodology involves both experimental design which is run in a laboratory setting (Brown & Lord, 1999) and non-experimental design methods, such as survey/scale, and other related methods that are used to gather numerical data. Scales can be used for a variety of useful information. In the current study, the scale was administered to a sample of English language teacher in Turkey for revealing their attitudes. In order to learn the way that English language teachers make sense of the way they are assigned, the researcher developed an attitude scale of recruitment model, because the use of such instruments is vital for understanding public perceptions (Hitchcock, et al, 2015).

3.2 Participants

The universe of the study consists of education stakeholders, including English language student teachers, in-service English language teachers, teacher trainers who work in ELT departments, education experts, and policymakers in Turkey from various institutions such as YÖK, ÖSYM, MoNE, and universities. The qualitative sample consists of 78

educational stakeholders, whereas the quantitative sample includes 271 in-service English language teachers.

3.2.1 Universe and Sampling in the Qualitative Phase

The universe of the qualitative phase of the study involves English language student teachers, in-service English teachers, teacher trainers, education experts and policymakers in Turkey. Particularly, the participants in the individual interviews were purposefully selected as a part purposive sampling, which "is a practical and efficient tool when used properly, and can be just as effective as, and even more efficient than, random sampling" (Tongco, 2007, p. 155). The student teachers participated in the study were the last year students. They were either students in ELT Department or those who were studying in other departments of Faculty of Science and Letters (e.g. English Language and Literature, American Culture and Literature, Department of Translation and Interpreting, Department of English Linguistics) and those who were in need of a pedagogic formation certificate for becoming an English teacher at public schools. The sample was taken from three universities in Turkey, namely Gazi University, Hacettepe University and Ankara University (see Table 1). Two imperative criteria were considered while choosing those who were studying at other departments: a) those who started pedagogic formation training, b) those who would start pedagogic formation training as soon as the school finished.

Like student teachers, the sample of in-service English language teachers included both those who graduated from ELT departments and the graduates of other departments, such as English Language and Literature, American Culture and Literature, Department of Translation and Interpreting, and Department of English Linguistics. They were graduates of different universities (see Table 2). Two criteria were considered while choosing the sample of English language teachers who would participate the qualitative part of the study: They were a) those who were working in public schools (the teachers of private schools were not chosen since private schools determine their own assessment criteria), and b) those who had ÖABT and/or interview experience.

For the sample of teacher trainer, the most salient criterion was their experience. Thus, at least, having at least a ten-year experience as a teacher trainer was the main criterion while

selecting them. The interviewed teacher trainers were either associate professor or full professor from different universities (see Table 3.

The sample of education experts involved members from different institutions. The sample included education experts from YÖK, ÖSYM, MoNE, and universities. They were those who participated workshops on ÖABT, and contributed to the content and process of it (see Table 4).

Finally, five policymakers were interviewed. They were executive persons and closely involved in the education policy of the country (see Table 5).

3.2.1.1 Demographic Features of Student Teachers

This part includes information about some demographic features (gender, affiliation and department)   of student teachers. Below is the table providing information of the student teachers who participated in this study from various universities. The universities in the tables were listed according to participant size. When the table blow is examined, it is seen that student teachers (male 11, female 20) were from three universities, namely Ankara, Gazi and Hacettepe. According to the table, student teachers of ELT department were all from Gazi University. Ten student teachers came from Ankara University. They were studying either at American culture (N=3) and literature, English language and literature (N=3) or department of English linguistics (N=4). Participants from Hacettepe Univerity (N=6) were students at department of translation and interpreting (N=4) and English language and literature (N=2).

Table 1.

Features of Student Teachers

Interviewee	Gender	Department	University
ST01	Male	ELT	Gazi
ST02	Male	ELT	Gazi
ST03	Female	ELT	Gazi
ST04	Female	ELT	Gazi
ST05	Female	ELT	Gazi
ST06	Male	ELT	Gazi
ST07	Female	ELT	Gazi
ST08	Male	ELT	Gazi
ST09	Female	ELT	Gazi
ST10	Female	ELT	Gazi
ST11	Female	ELT	Gazi
ST12	Female	ELT	Gazi
ST13	Male	ELT	Gazi
ST14	Female	ELT	Gazi
ST15	Female	ELT	Gazi
ST16	Male	American Culture and Literature	Ankara
ST17	Female	American Culture and Literature	Ankara
ST18	Female	American Culture and Literature	Ankara
ST19	Female	English Language and Literature	Ankara
ST20	Male	English Language and Literature	Ankara
ST21	Female	English Language and Literature	Ankara
ST22	Male	English Linguistics	Ankara
ST23	Male	English Linguistics	Ankara
ST24	Female	English Linguistics	Ankara
ST25	Female	English Linguistics	Ankara
ST26	Female	Department of Translation and Interpreting	Hacettepe
ST27	Female	Department of Translation and Interpreting	Hacettepe
ST28	Male	Department of Translation and Interpreting	Hacettepe
ST29	Male	Department of Translation and Interpreting	Hacettepe
ST30	Male	English Language and Literature	Hacettepe
ST31	Male	English Language and Literature	Hacettepe

3.2.1.2 Demographic Features of In-Service Teachers

In this part, in-service teachers' demographic features (gender, department and graduation) are presented. The table provides information of the in-service teachers who participated in this study. The table was organized according to interview date.

Table 2.

Features of In-Service Teachers

Interviewee	Gender	Department	Graduation
T01	Male	English Linguistics	Ankara
T02	Female	ELT	Necmettin Erbakan
T03	Female	American Culture and Literature	Ankara
T04	Male	ELT	Çanakkale
T05	Male	ELT	Necmettin Erbakan
T06	Female	ELT	Necmettin Erbakan
T07	Female	ELT	Boğaziçi
T08	Female	English Language and Literature	Süleyman Demirel
T09	Male	English Linguistics	Ankara
T10	Male	American Culture and Literature	Ankara
T11	Male	ELT	Necmettin Erbakan
T12	Female	ELT	Gazi
T13	Female	English Language and Literature	Süleyman Demirel
T14	Female	ELT	Erciyes
T15	Female	American Culture and Literature	Hacettepe
T16	Female	ELT	Hacettepe
T17	Female	ELT	Marmara
T18	Male	English Language and Literature	Atatürk
T19	Male	English Language and Literature	Karabük
T20	Female	ELT	Gazi
T21	Male	ELT	Russia
T22	Female	ELT	Pamukkale
T23	Male	ELT	Akdeniz
T24	Female	English Language and Literature	Atatürk
T25	Female	English Language and Literature	Hacettepe

As table 2 presents, 25 in-service teachers (male 10, female 15) joined the qualitative part. While 14 of them graduated from English language teaching department, 11 of them came from other departments, and had joined a pedagogic formation course to get teaching certification. They represent various universities, such as Ankara, Akdeniz, Atatürk, Boğaziçi, Çanakkale, Erciyes, Gazi, Hacettepe, Karabük, Marmara, Necmettin Erbakan, Pamukkale and Süleyman Demirel, and one participant graduated from a university in Russia.

3.2.1.3 Demographic Features of Teacher Trainers

Demographic features of teacher trainers (gender, title and affiliation) are presented in this part. The initial letters of the universities were taken into account while organizing the table.

Table 3.

Features of teacher trainers

Interviewee	Gender	Title	Affiliation
TT01	Female	Professor	Çukurova
TT02	Male	Associate Professor	Çukurova
TT03	Male	Professor	Gazi
TT04	Female	Associate Professor	Gazi
TT05	Female	Associate Professor	Gazi
TT06	Female	Associate Professor	Hacettepe
TT07	Female	Associate Professor	Hacettepe
TT08	Male	Professor	METU
TT09	Female	Professor	Sakarya

As it is seen in table, the sample consists of three males and six females. While four of them are full professors, five participants are associate professors. They work in Çukurova, Gazi, Hacettepe, METU and Sakarya universites.

3.2.1.4 Demographic Features of Education Experts

Below is the table providing demographic features (gender, title and affiliation) of education experts who participated in this study from various institutions. The institutions in the tables are listed according to their initial letters.

Table 4.

Features of Education Experts

Interviewee	Gender	Title	Affiliation
EE01	Male	General Director	MEB
EE02	Male	General Director	MEB
EE03	Female	Professor	University
EE04	Male	Associate Professor	University
EE05	Male	Professor	University
EE06	Male	Professor	YÖK
EE07	Male	Professor	YÖK
EE08	Male	Professor	YÖK

Table 4 shows that seven males and one female are in the sample of education experts. Two experts are general director in the Ministry of National Education. Three experts work at universities as full professors, while three experts represent the Council of Higher Education.

3.2.1.5 Demographic Features of Policymakers

This part introduces gender, title and affiliation of policymakers. The initial letters of the institutions were considered while organizing the table.

Table 5.

Features of Policymakers

Interviewee	Gender	Title	Affiliation
PM01	Male	General Director	MEB
PM02	Male	General Director	MEB
PM03	Male	General Director	MEB
PM04	Male	Professor	YÖK
PM05	Male	Professor	YÖK

Each member of this sample is selected carefully and purposively. Each policy maker is authorized and he has voice in the education policy of the country. The sample consists of three general directors in the Ministry of National Education and two professors working as members of the Council of Higher Education.

3.2.2 Universe and Sampling in the Quantitative Phase

In the quantitative phase, 271 English language teachers participated in the study. As the participants were selected according to their convenience, English teachers working in public schools (from primary to upper secondary school) in Ankara were asked to join the study. However, since this scale was in an online format and sent to participants via e-mails, we asked each participant whether s/he could forward the mail to one or more of her/his friends to ensure the adequate number. That is to say, the quantitative part included both convenience sampling which is used for "choosing settings, groups, and/or individuals that are conveniently available and willing to participate in the study" and snowball sampling in which "participants are asked to recruit individuals to join the study" (Onwuegbuzie & Collins, 2007, p. 286). The researcher took into account two criteria while choosing the sample of in-service English language teachers who would participate the quantitative part of the study: They were those a) who were working in public schools, and b) who had ÖABT experience.

3.2.2.1 Demographic Features of Participants in the Quantitative Phase

This part involves English language teachers' some demographic variables, such as gender, grade taught, graduated department, reason for choosing department, desired profession. The percentages and frequencies of these variables are presented in this section.

Table 6.

Frequency and Percentage Distribution According to Gender

Gender	N	%
Male	83	30.6
Female	188	69.4
Total	271	99.6

According to Table 6, the quantitative sample consists of 271 participants. They are 188 females (69,4) and 83 males (30,6).

Table 7.

Frequency and Percentage Distribution According to Grade Taught

Grades	N	%
Primary	53	19.6
Lower secondary	124	45.7
Upper secondary	94	34.7
Total	271	100

As it is clear in Table 7., it was aimed to investigate the relationship between grade taught and English teachers' attitude towards the current recruitment model. Grade taught was divided into three categories. It is clear that 53 (19.6%) teachers teach English to primary school, 124 (45.7%) teachers teach English to lower secondary school, and 94 (34.7%) teachers teach English to upper secondary school.

Table 8.

Frequency and Percentage Distribution According to Graduated Department

Graduated department	N	%
ELT	218	80.4
Formation	53	19.6
Total	271	100

In this table, the relationship between graduated department and English teachers' attitude towards the current recruitment model is examined. Those who graduated from ELT programs are 218 (80.4) while those who had formation are 53 (19.6).

Table 9.

Frequency and Percentage Distribution According to Reason for Department Choice

Reason for department choice	N	%
Like	172	63.5
Guidance	43	15.9
English teacher	32	11.8
Other	24	8.9
Total	271	100

The reason why teachers chose the graduated department was examined and categorized. 172 (63.5) teacher chose the graduated department because they like it. Those who were guided are 43 (15.9) whereas 32 teachers (11.8) chose the graduated department because of their English teacher. Those who indicated other reasons are 24 (8.9).

Table 10.

Frequency and Percentage Distribution According to Desired Profession

Desired profession	N	%
Teacher	146	53.9
Academician	73	26.9
Other	52	19.2
Total	271	100

Table 10 shows the profession that English language teachers desired to be. The majority (146 teachers with 53.9 percentage) stated that being a teacher was desired profession. Those who wanted to be academicians were 73 (26.9). There are also some teachers (52 with 19.2 percentage) who considered other professions as desired one. As it might be clearly noticed in Table 10, most of the participants see teaching as a desired profession.

3.3 Data Collection Techniques

The researcher administered several data collection techniques to collect data in order to examine the phenomenon in question. Therefore, various data collection tools were employed to answer research questions.

3.3.1 Qualitative Data Collection Techniques

The qualitative data consisted of semi-structured interviews, focus groups and documents. The rationale of using more than one data collection tools and techniques is to increase the trustworthiness of the qualitative data (Graneheim & Lundman, 2004).

3.3.1.1 Semi-structured interview as a Data Collection Tool

Interviews is generally preferred to gather deeper exploration from experts dealing with complex topics because time and focus are spent with a single participant (Hitchcock, et al, 2015). As parallel with the aim of the study, interview can be structured, unstructured or

semi-structured. In the current study, semi-structured interviews were used. Being a more flexible and common version of interviews (Holloway & Wheeler, 2010), a semi-structured interview "allows depth to be achieved by providing the opportunity on the part of the interviewer to probe and expand the interviewee's responses" (Rubin & Rubin, 2005 p. 88). The researcher formed a semi-structured interview forms consisting of two parts. In the first part of the form, there were questions about participants' demographic information (gender, institution and region they work, professional experience, university and department where they graduated, weekly course load… etc.). In the second part, there are questions and sub-questions that aim to reveal their perceptions about the components of the current recruitment model in general and the Teaching Field Knowledge Test, specifically (e.g., its usefulness, its functionality, whether it should be revised… etc.). The interviews were conducted in Turkish in order to raise the quality and quantity of the data (Mackay and Gass, 2005 p. 174). Before data collection, the interview protocol was reviewed for accuracy by a group of PhD candidates and piloted. All interviews, except two (discussed in 3.3.1.1.2 Interview with the In-Service Teachers) were recorded. The telephone interview approach was chosen for participants due to some reason (explained in 3.3.1.1.1 Interview with the Student Teachers).

3.3.1.1.1 Interview with the Student Teachers

Thirty-one student teachers were interviewed in the study. The researcher was familiar with the student teachers in ELT department of Gazi University, because he was also a research assistant at the same university. So, the interviews with those participants took place mostly in the researcher's office. However, the interview process with other participants followed two phases. In the first phase, the researcher made contact with them to introduce himself and give a brief information about the aim and content of the study. The goal of this step is to make interviewees be familiar with the research environment for ensuring credibility (Richards, 2003). The telephone interview approach was applied for six of them due to their working conditions. The duration of interviews was between 19 to 35 minutes. These interviews were audiotaped and transcribed. The main questions of semi-structured interviews with the student teachers are available in the Appendix B.

3.3.1.1.2 Interview with the In-Service Teachers

The interview process of the in-service teachers did not differ from that of student teachers. The telephone interview approach was not used as all teachers agreed face-to-face interview. Though two teachers expressed discomfort in the case of being recorded and rejected audio recording accordingly. The researcher took notes when those two participants responded to the questions. The duration of interviews was between 23 to 45 minutes. The main questions of semi-structured interviews with the in-service teachers are available in the Appendix C.

3.3.1.1.3 Interview with the Teacher Trainers

The sample of teacher trainers was selected from different universities (e.g. Çukurova, Gazi, Hacettepe, METU, and Sakarya) purposively. They were role models for student teachers, so their opinions were necessary. The reason behind choosing different universities was the assumption that different institution culture would add variety to data. The duration of interviews was between 25 to 40 minutes. The main questions of semi-structured interviews with the teacher trainers are available in the Appendix D.

3.3.1.1.4 Interview with the Education Experts

In May, 2017, there was a workshop on ÖABT which was performed by the collaboration of ÖSYM, YÖK, MEB and Universities in Ilgaz, Kastamonu, Turkey. The researcher was invited to that organization as a field expert of ÖABT. All participants in the workshop were field experts. The researcher seized the opportunity, and interviewed with some of them. The duration of interviews was between 20 to 40 minutes. The main questions of semi-structured interviews with the education experts are available in the Appendix E.

3.3.1.1.5 Interview with the Policymakers

In democratic countries, there seems to be a positive correlation between the quality of policy advice and the quality of the information available to those most involved in the policy process (Howlett & Newman, 2010 as cited in Head, Ferguson, Cherney and Boreham, 2014). From this point of view, the perspective of policy makers was assumed to

provide valuable data about the rationale behind the current recruitment model. Interviews took place in the office of the relevant policy maker. The duration of interviews was between 30 to 43 minutes. The main questions of semi-structured interviews with the education experts are available in the Appendix F.

3.3.1.2 Focus Group as a Data Collection Tool

Focus groups were conducted with groups of approximately six to twelve people at a time. Although they are considered to be more efficient than interviews, fewer questions can be asked when compared to interviews. The need for focus groups was because of the fact that the researcher aimed to generate lively discussions among participants so as to understand perceptions of different stakeholder towards the assessment process of recruitment policy and observe whether stakeholders arrived at a consensus on a controversial topic over time (see Onwuegbuzie, Dickinson, Leech, and Zoran, 2010 for focus group practices and frameworks for assessing group consensus). Denzin and Lincoln (2011) noted that focus-group interviews made it possible to examine discussions of the participants all together. Therefore, the researcher conducted two focus groups; one with student teachers, the other with teachers. The focus group with student teachers included 12 participants (six graduates of ELT and six graduates of other departments). The in-service teacher focus group consisted of five ELT and four of non-ELT graduates. Focus groups were selected with the assumption that they might particularly help us understand the perceptions of ELT and non-ELT graduates of TAE model, how they addressed a conflicting issue and whether a change occurred in their perceptions. Interviewees were from different places, so a well-known cafe located in Kızılay was determined the rendezvous. Plenty of time was spent to introduce participants to each other and create a warm atmosphere because the majority did know each other. Therefore, the average duration of focus groups took 3 hours.

3.3.1.3 Document Review as a Data Collection Tool

Documents help researchers reach the information easily, and may be seen as sources that need to be taken into consideration in qualitative research because

> document reading can also be part of an observational study or an interview based project. Documents can provide background information prior to designing the research project, for

example prior to conducting interviews. They may corroborate observational and interview data, or they may refute them, in which case the researcher is 'armed' with evidence that can be used to clarify, or perhaps, to challenge what is being told, a role that the observational data may also play. (Yanow, 2007, p. 411)

According to Atkinson and Coffey (1997), they are 'social facts', which are generated, shared, and utilized in socially organized ways (p. 47). Documents can take a variety of forms, including advertisements, minutes of meetings, manuals, diaries and journals, organizational or institutional reports, survey data, and various public records, which can provide valuable information and insights (Bowen, 2009). Therefore, the government's development plans and the national education councils were examined in order to see the education policy of Turkey from 1923 to today. In addition, for determining a framework of teacher competencies that an effective English language teacher should have in Turkey, both local competencies (for example, MEB, YÖK) and global competencies (for example, CAEP, CEFR, EPOSTL, Education Council, TESOL) were reviewed.

3.3.2 Quantitative Data Collection Techniques

3.3.2.1 ETARS as Quantitative Data Collection Techniques

The quantitative part consisted of an attitude scale developed by the researcher: *English Teachers' Attitudes towards Recruitment System (ETARS)* (Appendix G & H). Literature lacks an attitude scale for teacher recruitment system in a centralized context. Thus, the development of ETARS is thought to be crucial. It aims to examine English teachers' attitude towards their recruitment to public school. It is a Likert-type scale with two parts. First part includes demographic questions (such as gender, grades s/he teaches, graduated department, reason for choosing the job, and desired profession). Second part consists of three components/dimensions/factors together with 23 items. The first dimension (D1), *positive attitude towards teacher assessment and evaluation* (TAE), is about the positive attitude of English language teachers towards the selectiveness and fairness of the teacher selection and recruitment model, while third one (D3) (*negative attitude towards TAE*) displays their negative attitudes of the aforementioned issues. The second factor (D2) (*being interested in TAE*) aims to reveal whether English language teachers are curious about the way they are assessed and evaluated while being assigned to public schools.

67

3.4 Reliability and Validity

Reliability and validity are the two most important criteria in scientific research. Although the concept of validity and reliability is widely used in a quantitative research, it is uncommon in qualitative research (Mackey and Gass, 2005). Instead, there exist some other criteria in qualitative research (please see Table 11). This study will use both quantitative and qualitative data collection tools. This is because quantitative research focuses on to what extent a phenomenon exists. However, qualitative research is more concerned with the meaning and interpretation of those events.

Table 11.

Criteria for Judging Research Quality from Quantitative and Qualitative Perspective

Traditional Criteria for Judging Quantitative Research	Alternative Criteria for Judging Qualitative Research
Validity-Reliability	**Trustworthiness**
Internal Validity	Credibility
External Validity/ Generalizability	Transferability/ Applicability
Reliability	Dependability
Objectivity	Confirmability

Adapted from Daymon & Holloway (2003, p. 101), Guba (1981, p. 80) and Trochim & Donelly (2007, pp. 162-3)

As it is clear in Table 11, the advocates of qualitative paradigm prefer the term *trustworthiness* instead of 'validity and reliability'. In addition, they utilize the qualitative terms *transferability/applicability, dependability,* and *confirmability* in place of 'external validity/ generalizability', 'reliability', and 'objectivity', respectively.

3.4.1 Reliability and Validity in Qualitative Research Design

As mentioned above, the proponents of this paradigm use trustworthiness instead of reliability and validity which includes credibility, transferability, confirmability and dependability. The term *credibility* refers that "the results of qualitative research are credible or believable from the perspective of the participant in the research." (Trochim & Donelly, 2007 p. 162). The role of researcher in qualitative research (both data collector

and analyst) might increase the likelihood of researcher bias (Miles & Huberman, 1994). In other words, the personal beliefs and interests of qualitative researchers might prevent the voice of participants (Mason, 2002). One way to reduce that kind of researcher bias is member checking which is used to validate trustworthiness of qualitative results (Doyle, 2007; Guba, 1980), because the participants play active role regarding the credibility of a study (Kumar, 2011). In order to ensure credible results in this study, the researcher used both member checking (after the preliminary categories of interviews occurred, results were presented to the majority to check whether results represent their views) and triangulation (Patton, 2014) (through the combination of scale, semi-structured interviews, focus groups and document reviews).

Transferability, concerning the aspect of applicability (Lincoln & Guba, 1985) is another important issue in qualitative research. For ensuring transferability, it is the responsibility of a researcher to inform readers about participants and the research process via a 'thick description' in order to enable the readers to determine whether findings can be transferable to their own setting (Korstjens & Moser, 2018 p.122). Accordingly, the whole process of this study was introduced thoroughly so that readers could be aware of all details.

According to Lincoln and Guba (1985) *dependability* involves the aspect of consistency while *confirmability* is about the aspect of neutrality. The first one refers to the researcher's responsibility to make the analysis process be in accordance with the accepted standards for a particular design, whereas the second requires securing the inter-subjectivity of the data (Lincoln & Guba, 1985 as cited in Korstjens & Moser, 2018). The comment should not be based on the researcher's viewpoints. Instead, it should be grounded in the data. The focus is on the comment process embedded in the data analysis process. According Korstjens and Moser (2018), an 'audit trail' is the strategy that is required to ensure dependability and confirmability. Therefore, an audit trail which requires a transparent description of the research steps was applied to secure both dependability and confirmability in this study. In this sense, the researcher enables readers to have an idea about decisions taken during the research process, sampling, how findings emerged and how data were managed.

3.4.2 Reliability and Validity in Quantitative Research Design

Reliability and validity are sine qua none of scientific research. According to Nunan (1992), reliability is about "the consistency of the results obtained from a piece of research" while validity can be defined as "…the extent to which a piece of research actually investigates what the researcher purports to investigate." (Nunan, 1992, p. 14) or "…the best available approximation to the truth of a given proposition, inference, or conclusion." (Trochim and Donelly, 2007, p. 20).

In quantitative study, instruments such as surveys/scales and questionnaires are usually utilized as data collection tool (Leedy & Ormrod, 2010) because of some advantages such as ease of administration, time and cost efficient, clarity of directions, scoring simplicity, accessibility to larger population, reliability, and validity (Gay, Mills & Airasian, 2011; Fraenkel, Wallen & Hyun, 2011).

The instrument utilized in this study was developed by the researcher. After being applied to the participants, the 32-item scale was analyzed and factor load values of some items (1, 2, 7, 12, 16, 25, 26, 28, and 31) were observed to be smaller than 0.45, which is the lower limit. These items were removed from the scale and the factor analysis was repeated.In order to obtain better results from the factor analysis, the promax technique was applied to the scale items as a rotation method and it was decided to classify the scale under three factors. As can be seen from Table 12, the KMO value was found to be .880 and the appropriateness of the sample size was confirmed by KMO and Barlett statistic (KMO = .880; $\chi2 = 3154.059$, p 0.000). The KMO value relates to the suitability of the sample. A value higher than .60 indicates that it is suitable for factor analysis (Tabachnick & Fidell, 2013). In addition to the KMO value, the high Barlett test scores indicate the feasibility of factor analysis and the high correlations between items (Şeker & Gençdoğan, 2006 as cited in Günüç, 2009). When the Table 12 is examined, the total variance of the three components of the scale is found to be 56.610, and it is also noted that the first component, the second component and the third component explained 29.909 %; 20.180%; 6.520 % of the scale, respectively.

Table 12.

Findings of KMO Barlett Test Values and Factor Analysis

Kaiser-Meyer-Olkin (KMO)		.880
Bartlett's Test of Sphericity	Kay-Kare (χ^2)	3154.059
	Sd	253
	p	.000

	Initial Eigenvalues			**Total Variance Explained** Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings
Item	Total	Variance (%)	Cumulative (%)	Total	Variance (%)	Cumulative (%)	Total
1	6.879	29.909	29.909	6.879	29.909	29.909	6.530
2	4.641	20.180	50.089	4.641	20.180	50.089	4.680
3	1.500	6.520	56.610	1.500	6.520	56.610	3.571
4	1.099	4.778	61.388				
5	.921	4.006	65.394				
6	.829	3.606	69.000				
7	.787	3.424	72.424				
8	.695	3.022	75.446				
9	.647	2.814	78.260				
10	.629	2.736	80.996				
11	.548	2.383	83.379				
12	.519	2.254	85.634				
13	.467	2.029	87.663				
14	.398	1.729	89.392				
15	.358	1.556	90.948				
16	.335	1.455	92.402				
17	.312	1.359	93.761				
18	.304	1.320	95.081				
20	.272	1.183	96.264				
21	.253	1.098	97.362				
22	.234	1.016	98.378				
23	.217	.943	99.321				

Table 13. presents the load values of three components.

Table 13.

Rotated Factor Load Values

Item	Component 1	2	3
m21	.858		
m6	.790		
m7	.788		
m10	.777		
m12	.776		
m15	.743		
m9	.743		
m4	.730		
m18	.727		
m2	.593		
m19		.801	
m22		.796	
m16		.791	
m8		.747	
m11		.736	
m23		.726	
m13		.705	
m20		.646	
m5			.814
m14			.666
m3			.623
m17			.584
m1			.513

When the Table 13 is examined, it seen that the load value of the first component is between .593 and .858; the load value of the second component is between .646 and .801; the load value of the third component is between .513 and .814

The Table 14 displays the scale of English teachers' attitudes towards recruitment model.

Table 14.

The Scale of English Teachers' attitudes Towards Recruitment Systems

Positive Attitude towards TAE	2- I think that the current TAE model selects student teachers with sufficient field knowledge
	4- I think that the current TAE model selects student teachers who love teaching profession
	6- I think that the current TAE model selects highly motivated student teachers
	7- I think that the current TAE selects student teachers who can communicate with stakeholders (e.g. students, colleagues, parents)
	9- I believe that the selection criteria of the current TAE model are fair
	10- I believe that the current TAE model can select student teachers who can teach English at every level (e.g. primary school, lower secondary school, upper secondary school)
	12- I think that the current TAE model selects student teachers who can constantly improve themselves
	15- I think that the current TAE model selects student teachers with required technological knowledge
	18- I think that the current TAE model selects student teachers with required professional knowledge
	21- I think that the current TAE model selects English student teachers according to the needs of the students
Being interested in TAE	8- I follow the recent developments in Turkey related to the current TAE model
	11- I follow the recent developments in the world related to the current TAE model
	13- I am interested in discussions related to the current TAE model
	16- I read written sources (e.g. books, magazines, news) about the current TAE model
	19- I follow electronic sources (e.g. blogs, forums) about the current TAE model
	20- I discuss with colleagues to increase my knowledge of the current TAE model
	22- Everything about the current TAE model attracts my attention
	23- I try to get information about the current TAE model in my spare time
Negative Attitude towards TAE	1- I believe that the current TAE model selects also the student teachers who cannot teach
	3- I believe that the current TAE model needs to be changed
	5- I become happy if the current TAE model is removed
	14- I become happy if the current TAE model is changed
	17- I think that the current TAE model is unfair

According to Table 14, the Scale of English Teachers' attitudes Towards Recruitment Systems consists of three components. The first component is *positive attitude towards TAE* which includes 10 items. The second factor is *being interested in TAE* which involves 8 items, and the final component is *negative attitude towards TAE* which has 5 items. Before naming the componets, the researcher asked four experts to label the componets. One of them was full professor and teaching a PhD scale development course. The second

expert was associate profesore in measurement and evaluation department. The last two were PhD candidates in in measurement and evaluation department. All of them reached a consensus on the name of components. The first component was labeled as Positive Attitude towards TAE and it contained items 2, 4, 6, 7, 9, 10, 12, 15, 18, 21. The second component was Being Interested in TAE and it contained items 8, 11, 13, 16, 19, 20, 22, 23. The third component was labeled as Negative Attitude towards TAE and it contained items 1, 3, 5, 14, 17. Table 14 shows the distribution of the items of all the components.

The internal consistency of the scale instrument was assessed through Cronbach's alpha, which is usually preferred by researchers for measuring internal consistency in Likert-type scales (Huck, 2012, p. 74). Regarding Cronbach's alpha coefficient which is presented between 0.00 and +1.00, if the coefficient is close to +1.00, it means that it has high internal reliability (Huck, 2012, p. 72). The reliability coefficients of the study are illustrated below (please see Table 15).

Table 15 displays Cronbach's alpha (α) reliability coefficients of components. First component consists of 10 items and its reliability is .918. Its item load value is between .593 and .858. Second component has 8 items, with .883 reliability coefficient. Item load value of second component is between .646 and .801. Finally, third component includes 5 items and its reliability coefficient is .703. Item load value of third component is between .513 and .814. Cronbach alpha (α) internal consistency coefficient for the scale is .758.

Table 15.

Information on the Structure of the Scale

Dimension	N. of Items	Items	Name of the component	α	Item load value
1	10	m2, m4, m,6, m7, m9, m10, m12, m15, m18, 21	Positive attitude towards TAE	.918	between .593 and .858
2	8	m8, m11, m13, m16, m19, m20, m22, m23,	Being interested in TAE	.883	between .646 and .801
3	5	m1, m3, m5, m14, m17	Negative attitude towards TAE	.703	between .513 and .814

As it is clearly seen in Table 15, the reliability coefficients of dimensions are between .703 and .918. Therefore, the scale is considered to be reliable since .70 and preferably higher is an acceptable value for reliability (Büyüköztürk, 2011; Fraenkel et al., 2011, p. 157).

Table 16 presents Descriptive statistics of the item total score of the scale.

Table 16.

Descriptive Statistics of the Item Total Score

	Mean	Std. Deviation	Minimum	Maximum
Total score	69.121	10.306	42	99

As it is seen, the mean was found as 69.121 (=69.121). The lowest total score in the sample is 42 and the highest total score is 99.

3.5 Data Analysis

The current study involves both quantitative and qualitative data sources. Content analysis was utilized for qualitative data. Both manual coding and Nvivo were used in qualitative data analysis while the statistical analysis software called Statistical Package for Social Sciences (SPSS) was used for quantitative data.

3.5.1 Analysis of the Qualitative Data

Before deciding on the use of a software package (NVivo) for the full study, the qualitative data were initially manually analyzed as piloting process (Welsh, 2002) so that the researcher could observe, control and manage the data. This process was not for any data manipulation. Instead, it was conducted for the intention of enabling the researcher to take precautions for data loss. It is not computer but the researcher who analyze qualitative data, computer just design data (Merriam, 1998 p.167).

Mackey and Gass (2012) underline the importance of a systematic data coding process for discovering, developing and forming well-grounded interpretations. While analyzing the data gathered from the semi-structured interviews and focus-group interviews, Dörnyei's

(2007) following stages were considered; a) transcribing, (b) pre-coding/coding, (c) growing ideas, and (d) interpreting the data and drawing conclusions (p. 246). Content analysis was used, which means that all categories were determined while analyzing the transcripts. First, all data obtained from the semi-structured interviews and focus-group interviews were transcribed, since the transcription is the first step of analyzing and deducing themes (Rubin & Rubin, 2012). While transcribing the data, the researcher omitted some words (e.g. hmm, uh, huh etc.) by adopting 'edited transcription' (Hansen, 2003 p. 136) in order to ease the analysis process. Second, the researcher read the data for getting a general idea and for checking pre-coding. Then, the data were examined and categorized. Finally, the same transcriptions were sent to an independent researcher in the field of Measurement and Evaluation and the researcher was requested for labeling the transcriptions. Miles and Huberman's (1994) formula was applied for interrater reliability which was found as 96%. The interrater reliability was found to be satisfactory since it exceeded the threshold value (90%) determined by Miles and Huberman (1994).

As mentioned above, after the preliminary manual coding, NVivo 12 was chosen for the full research, agreeing with Welsh's (2002) opinion that a combination of old (manual) and new (computer) methods gives rise to best results. NVivo enables a user to import any documents directly from a pdf, word or voice source and code them easily on screen (data are "coded" at "nodes" in NVivo). Also, as the program makes strips visible on screen, the researcher can see all codes and categories or themes at a glance. Consequently, it is generally accepted that, if used appropriately, such programs are effective in developing data analysis process (Bringer, Johnston & Brackenridge, 2004; Johnston, 2006). According to Richards (2002), researchers should not let software package manage their data but instead, consider the software as a helper for their project.

3.5.2 Analysis of the Quantitative Data

Data obtained from the English Teachers' Attitudes towards Recruitment System (ETARS) were analyzed via the statistical analysis software called Statistical Package for Social Sciences (SPSS). Descriptive statistics are presented including measures of frequency, central tendency, and variability. The researcher has made use of these measures to determine types of statistical analyses that have been applicable for the data (Mackey & Gass, 2005). The statistical technique can be in the form of either 'parametric' or 'non-

parametric' tests according to the distribution of the data. The normality distribution of the relationships between demographic variables and dimensions D1, D2, D3 together with the item total score of ETARS was analyzed. The 'Z test', in which the statistical significance of the skewness index is tested, is one of the tests that can be used for this purpose. The fact that the p value calculated in the test result is less than $\alpha = .05$ proves that the scores come from the normal distribution (Field, 2009; Howitt & Cramer, 2011).

Kolmogorov-Smirnov normality tests were applied, and the data was noted to show normal distribution. According to Tabachnick (2013), a ratio for skewness and kurtosis values between -1.5 +1.5 is the indicator of normal distribution in a data set. T-test (Independent Samples Test) and ANOVA tests were applied for the demographic variables two and three or more categories respectively. For this reason, t-test was used when the population means of only two groups was to be compared; however, ANOVA was preferred when means of 3 and more groups are to be compared.

3.6 Procedure

At the hearth of this study was the idea that education leaders must pay attention to teacher selection and recruitment model because an effective teacher evaluation model contributes not only teachers but also school and education improvement (Darling-Hammond, 2010b; Liu & Johnson, 2006; Stiggins & Duke, 1988). In order to select and recruit the most suitable student teacher from graduates, the MoNE has aimed to develop a selection system, based on multi-evaluation resources such as undergraduate success, portfolio, evaluation of practicum (teaching practices), interviews etc. (Teacher Strategy Paper 2017-2023). However, developing a new model cannot be considered without taking into account consequential and social validity of testing. Learning how education stakeholders perceive testing can be helpful not only "in identifying aspects of a testing procedure that trouble public stakeholders, the process also may yield insights into the psychometric notions of consequential and social validity." (Hitchcock et al., 2015 p.2). To gain a richer understanding of how education stakeholders perceive the current teacher evaluation system and how they view an appropriate model for selecting and recruiting EFL teachers in parallel with education needs of the country, a sequential exploratory mixed methods research design was conducted to gather data. That is, the research begins with qualitative phase and is followed by quantitative one.

Phase one of this study included qualitative data from interviews of stakeholder who participated in the present study and expressed their feelings and perceptions of evaluation. First, researcher spent 8 months for examining the literature in order to probe the teacher evaluation of different countries. In the same breath, the national education councils and development plans of Turkey were investigated so as to draw a picture of Turkey's recruitment policy from the establishment of the republic to this day. Interview questions were prepared for different stakeholders. Then, these questions together with interview process were piloted to adjust the interview guide (Harding, 2013), which enables researchers to delve into the participants' views and experiences, and generate rich data (Ryan & Cronin, 2009; Smith, Flowers & Larkin, 2009). Next, the interviews were conducted with student teachers, in-service teachers, teacher trainers, education experts and policymakers for the main study. Except experts and policymakers, all stakeholders were those who are related to English language teaching. Interviews were held in June, October and November 2017. The researcher had a talk with each participant either face-to-face or on the phone before the interview was conducted for three purposes:

(1) to inform the interviewee about the scope of the study,

(2) to benefit from interviewee's free time appropriately,

(3) to create a sincere atmosphere.

The majority was interviewed face-to-face; however, some participants had lack of time, so they were reached by telephone. Both on a one-to-one and focus group interviews were conducted. All interviews (except two) were audio recorded and then transcribed by researcher. The data were analyzed via content analysis. Although researcher analyzed data manually at the beginning of the study, he used NVivo 12 for the rest of qualitative data.

In the second phase of the study, it was aimed to reach more participants for generalization issues. Due to the absence of a scale for gauging EFL teachers' attitudes towards the way they are recruited to public schools in a centralized system, researcher developed an attitude scale for EFL context in which teachers are assigned centrally via standardized high-stakes tests. For scale development, researcher joined a PhD scale development course for one term. The instructor was a full professor who developed a few scales. During the course, researcher created an item pool that consists of 81 items. In accordance with the instructor's suggestions, some items were omitted from the pool and 68 items were used for piloting and administered to 319 English language teachers.

ETARS was administered to EFL teachers through an online service. Participation in the scale was voluntary. To ensure that only teachers who took ÖABT completed the scale, the demographic information part was carefully prepared, asking the date they were assigned to public school. ÖABT was introduced in 2013, so if a participant indicated any date before 2013, their responses were eliminated and not considered during SPSS analysis. The analysis of piloting provided three components, consisting of 32 items for the main study. The new 32-item scale was administered to 271 English language teachers and reanalyzed. At the end, 9 items were excluded from the scale. Finally, ETARS involved 23 Likert-type scale response items, examining teachers' attitudes about the way they are recruited. The quantitative results were used for the last question which was for generalizing the findings.

3.7 Piloting Process

It is vital that data collection tools and procedures need to be piloted to identify possible problems that may arise before actual work is undertaken and to take the necessary precautions to address them (Dörnyei, 2007; Tashakkori & Teddlie, 2003). Piloting as "dress rehearsal" (Nunan & Bailey, 2009, p. 145) is an important way of determining the applicability and usefulness of data collection methods and making necessary changes on these methods before they are implemented with participants (Mackey & Gass, 2005, p. 43). From this point of view, both the scale and the semi-structured interview processes were piloted. In this way, it was aimed to prevent problems that may arise due to the lack of understanding of the questions or the problems that may occur due to voice recordings.

3.7.1 Piloting Interviews

Interviews are usually administered when in-depth information is required. They help researchers discern the phenomenon from the interviewee's account (Merriam, 2016). Interview questions are central to the interviewing process, and piloting interviews is vital so as to observe whether questions work. For example, during her pilot study, Castillo-Montoya (2016) observed that interview questions and protocols could be enhanced because it enables researchers to realize and identify the flaws and make necessary modifications to the main study (Kvale, 2007). It includes all necessary precautions related to interviewing protocols, from checking the suitability of questions to other issues (for

example, practicality of audio recorder, appropriateness of the place in which they are interviewed). So, the interview was piloted with six participants (three student teachers and three in-service teachers). Their responses were recorded by an audio recorder. Each interviewee was asked to share his/her feelings about the question and the process. They were also asked to express whether they got bored or whether they felt under pressure for any reason. Their feelings and perceptions were informative in redesigning the process. All audio recordings were transcribed to begin the coding and extracting phases. Then, codes were gathered under categories. As a result of the piloting, one question was reformed as it had ambiguity problem and three new interview questions were added.

3.7.2 Piloting Scale

Having found no scales related to EFL teachers' attitude towards the way they are evaluated, the researcher developed with the help of an expert in the field of scale development. Any particular survey/scale requires piloting after being drafted. Researcher outlines steps in which these instruments can be assessed. Some of these steps are as followings (Willis, 2005 and Fowler, 2009 as cited in Hitchcock, 2015 p. 9);

> (a) whether items seem adequate to assess the intended topic; (b) whether any words/items are problematic in that they are confusing, offensive, or evince some level of ignorance that will undermine trust; (c) whether the order of items appears to be problematic in any way; (d) whether directions are clear and accurate and timing estimates are accurate; (e) whether any technological applications (e.g. web-based surveys) are apparent; (f) whether the visual layout is easy to follow and properly sub-divides items into logical sections; and (g) whether response options match up well with items; and so on.

An 81-item pool was created; however, due to the suggestion of the instructor (an expert in the field of scale development), so the number of items decreased to 68 for piloting. Before the piloting process, the 68-item form was sent ten in-service English language teachers who work in various school levels. They were requested to read the items carefully and respond on each item. All items were checked according to their reflections, and the necessary changes were done. Then, the piloting was conducted with 319 English language teachers. A three-component scale, consisting of 32 items occurred after analysis. The new 32-item scale was administered to 271 English language teachers and reanalyzed. At the end, 9 items were excluded from the scale. Finally, ETARS involved 23 Likert-type scale response items, examining teachers' attitudes about the way they are recruited.

CHAPTER IV

FINDINGS AND DISCUSSION

This study aimed at suggesting a more valid, reliable and fair teacher assessment and evaluation (TAE) model for EFL contexts where the recruitment of English language teachers to public schools is centralized, by using a sequential exploratory mixed methods research design. The perceptions of stakeholders may reveal problems related to aspects of a testing procedure that trouble them (Hitchcock, et al., 2015) and act an important part of the evidence base that is essential to evaluate the validity argument (Messick, 1989, 1995) for the evaluation model. Thus, this dissertation approached the subject of stakeholders' perceptions of teacher assessment and evaluation as validity evidence for recruitment model. Semi-structure interviews and focus group techniques were conducted with student teachers, in-service English language teachers, teacher trainers who work in ELT programs, education experts, and policymakers in order to display how they perceive the current recruitment model. An attitude scale was developed by the researcher by the help of these perceptions and literature, and this scale was utilized to gather quantitative data from in-service English teachers by revealing their attitudes towards the current recruitment model. Accordingly, this part includes both quantitative and qualitative findings together with their discussions. To make the study more user friendly, the qualitative and quantitative findings will be presented under separate titles.

4.1 Qualitative Findings

The results are introduced with reference to the research questions of the study. The data were analyzed through content analysis. Both manual coding and NVivo 12 were utilized.

4.1.1 Findings Related to the Competencies That a Qualified EFL Teacher Should Have

For the first qualitative research question, 'Which competencies should a qualified English language teacher have?', it was aimed to display the competencies that were emphasized by education stakeholders (student teachers, in-service English language teachers, teacher trainers, education experts, and policymakers). The researchers decided to present the opinions of all stakeholders under one title after analyzing the data carefully, because almost all of them mentioned the similar competencies that an English language teacher need to have in order to teach English effectively. During the coding process, it became clear that all perceptions were either about knowledge, skills, or attitudes. Therefore, three categories (knowledge-related competencies, skill-related competencies, attitude-related competencies) were determined for the whole perceptions related to competencies of qualified teachers (please see Table 17). In other words, although stakeholders listed a lot of competencies for English teachers, the competencies were noted to be related to either knowledge, or skills, or attitudes of student teachers.

Table 17.

Competencies That a Qualified EFL Teacher Should Have

Categories	Codes	f
Knowledge-related competencies	Has knowledge of subject matter	70
	Has knowledge of underlying theories of teaching (Has a repertoire of teaching techniques and approaches)	62
	Has knowledge of learners' development (cognitive, linguistic, social, emotional, and physical)	58
	Has knowledge of curriculum	33
	Has knowledge of legislation	18
Skill-related competencies	Speaks English well	78
	Transfers his knowledge into practice	71
	Conducts enjoyable lessons	70
	Uses technological tools	66
	Prepares and uses appropriate assessment and evaluation tools	65
	Understands native and/or non-native conversation	59
	Manages students and learning environment	55
	Prepares, uses, and evaluates materials	53
	Demonstrates a coherent writing	25
Attitude-related competencies	Loves his/her job	71
	Considers learners' needs	65
	Demonstrates a fair attitude towards their students	55
	Treats students as individuals	54
	Being open to professional development	52
	Being aware of the latest ELT developments (as a researcher)	44
	Has warmth and tolerance	40
	Has a good sense of humour	39
	Establishes an effective interaction, communication and cooperation with stakeholders	38
	Being democratic and humanist	33

4.1.1.1 Knowledge-related competencies

Data obtained from both focus group and semi-structured interviews showed that almost all of the participants underlined the importance of subject knowledge. That is, stakeholders stated that not only must an effective teacher, for example, analyze topics related to his/her subject but also she/he could compare and contrast different teaching techniques and methods for teaching her/his subject area. Some representative quotes are:

> A qualified teacher must possess the required knowledge of the subject s/he teaches. How can I say? A classroom environment is similar to a film stage. The actor must know the script well no matter how good s/he is…. Of course, in addition to subject knowledge, s/he should know how to teach it… (T6)

> While demonstrating micro-teaching activities, our lecturers constantly suggest us to be prepared in the subject that we would teach. We recognize the importance of these warnings and suggestions during practicum. Students get bored easily in the lesson of a teacher who does not possess the required subject knowledge or the knowledge of conveying information to learners. (ST8)

> Content knowledge and pedagogical content knowledge are a must for a qualified teacher. Teachers cannot teach their subjects if they do not have such knowledge. (PM1)

The majority of the participants stated that English language teachers should know the underlying theories of teacher education and second language acquisition. Likely, according to them, an English language teacher is expected to have a rich repertoire of techniques and methods in order to attract students' attention. They believe that if a teacher does not have such knowledge, s/he cannot teach an enjoyable lesson. As a result of this knowledge, teacher considers learners' development (cognitive, linguistic, social, emotional, and physical) while preparing, presenting, and evaluating tasks and lessons. Here are some representative quotes:

> We suggest our student teachers to spend more time on techniques, approaches, and principles in language teaching. Such knowledge helps them conduct lessons in accordance with learners' needs and levels so that their learners do like English lessons. (TT3)

> A good knowledge of second language learning theories helps teachers create appropriate tasks for students, taking into account their age, development, psychology, other important issues. They know that students are culturally and linguistically diverse in their classrooms, so they develop materials, assessments tools, and instructional teaching strategies accordingly. (EE1)

> When I case myself, considering my English teacher at the upper secondary school, I can say that if a teacher does not know the underlying principles of language teaching and learning, it is imperative for him/her to follow a one-size-fits-all lessons, ignoring his/her learners' needs. (ST1)

I graduated from a well-known university. During my first year experience at a lower secondary school, even my undergraduate knowledge about the principles and approaches of language teaching and learning helped me keep in mind my learners' needs while teaching them. (T5)

Some participants underlined the importance of curriculum and legislation knowledge. Here are some representative quotes:

An effective teacher is expected to have a good knowledge of curriculum, because their use of curriculum knowledge probably influences their teaching. (PM5)

I pay attention to curriculum. I prepare my lesson plans based on that knowledge. (T14)

4.1.1.2 Skill-related Competencies

When the data were analyzed, it was observed that all educational stakeholders of this study were not satisfied with the language proficiency of our English teachers. The most underlined component is the speaking ability. The fact that all participants mentioned speaking ability demonstrates that one of the most important competencies that an English language teacher must have is being good at speaking English. Some of the comments were:

An effective language teacher must be good at language proficiency. I mean… well… speaking, writing, reading, listening, and grammar are important components that English teachers need to have. However, especially, a good pronunciation and a fluent speaking are sine qua none of English language teachers. The teacher cannot improve his/her learners' speaking and listening skills unless s/he possesses a fluent and accurate speaking ability. (PM4)

English teachers should master the target language and have fluent and accurate speaking ability. Do you know how we benefit from a good speaking in the class? When I speak Turkish in the classroom, they do not take a positive attitude towards English language. Instead, when I speak English with my students, they start to join the lesson, and become aware of the importance of learning a new language, which contributes to conducting an efficient and fun lesson. (T7)

Without speaking, it does not make sense to use the language, so speaking skill is of importance. Then comes other skills such as listening, reading, and writing. (ST28)

Some participants criticized the teacher assessment model, since it did not gauge the productive skills. They indicated that because language proficiency of teacher is not

assessed, it is not at the desired level, and even ELTE programs do not pay attention to that issue:

> In my opinion, if you do not assess something, you cannot develop it. For example, the MoNE ask us perseveringly to enhance students' four skills. However, we are given limited course hours and students English is assessed via multiple choice questions. The only indication of the students' success is the number of correct answers they did in examination. Thus, we focus on improving their question solving strategies. (T1)

> A teacher may get high scores from a language proficiency test which was conducted in a multiple-chose format, but s/he may not be able to speak well, and have a listening and writing ability. For this reason, student teachers must be subject to four skill-based exams for language proficiency. In other words, it is not possible to assess someone by measuring his/her knowledge of grammar, vocabulary, and reading through paper-pencil exams. (TT6)

> Student teachers usually do not consider four skills during undergraduate period, because they spend all time and effort for components which are asked in high-stakes examinations. According to my experience, teachers who ignored those skills use, especially, productive skills very limited. (EE8)

Most of the participants asserted that competent English language teachers create a nice environment in which students enjoy lesson. They stated that competent English language teachers take into account students' needs and development (cognitive, linguistic, social, emotional, and physical) while preparing, using, evaluating materials. Not only are they very good at teaching language to students but also they manage students and learning environment:

> Having a good level of knowledge-related competencies does not guarantee quality teaching. Teachers need required skills for transferring what they know to their students. They carry about how to teach, so they take into account students' cognitive, linguistic, social, emotional, and physical needs while preparing materials, activities, and assessment tools. (TT4)

> For me, a good English language teacher is someone who can teach. He gives instructions, prepares activities, and chooses assessment types and tools in accordance with students' needs, uses various techniques. (ST13)

> First, a qualified English teacher should use target language and enjoy teaching it. Then, she or he can create a good learning environment in which students join the lesson and are willing to spend time. (T20)

More than half of the participants referred to the importance of technology literacy. They stated that English language cannot be taught without using technological tools, so they should be technologically competent. The term technology literacy does not mean only using and managing technological tools; instead, refers to the ability of utilizing these tools for accessing, constituting, combining, evaluating, interpreting and communicating

information. That is, they apply these tools not only for their professional development but for letting students experience a productive and enjoyable lesson:

> Now, there are smartboards in almost all classrooms. I believe that my technology literacy is very good. I am taking online courses from universities in foreign countries. I am following the forums in order to learn what's happening in ELTE programs in the world. In this case, you can conduct lessons from different perspectives thanks to these different sources. On other side of the coin, there are classes taught by teachers who use only course books. Can you imagine the difference between students of these two different classrooms? I've been using a smartboard for a year. I can see the difference between my last and this year teaching. Kids have started to like this lesson… (T2)

> With the development of technology, the new generation and its interests are changing very rapidly. We witnessed a lot of different things during practicum. For example, I heard that students were talking about a new video on YouTube. The following week, I showed them the video they had mentioned. As they were interested in that video, they all started to pay attention to the lesson. It was really an enjoyable lesson. (ST18)

4.1.1.3 Attitude-related competencies

Almost all of the participants stated that the most prominent competency that an English language teacher should have is that she or he has to love his profession. It is interesting that those who chose the teaching job due to other reasons (e.g. pragmatic) also expressed that love of profession is an indispensable inducement. That is, no matter why people choose teaching as a profession, when they start to teach English, they must have a strong loving for their job.

> We are three English language teachers in this school. Like me, Aysun Hoca (pseudonym) loves her profession. She is so happy and really conducts very enjoyable lessons; thus, students like English. However, Aykut Hoca (pseudonym) complains about being here and made a confession that he would leave the business at first opportunity and settle abroad. This unrest is reflected even in his gestures and facial expressions. Unfortunately, students get bored and are not happy in his lessons. (T9)

> …Teachers should love their profession. Otherwise, neither do students, parents nor administration want to see an ill-humoured teacher. Such a teacher himself or herself does not want to go to school anyway. S/he should love his/her job, has a smiling face, and must be open to different views. (T19)

Most of the participants asserted that a competent teacher should consider needs of the learners, demonstrate a fair attitude towards them, and respect them as individuals. Here are some extracts:

It is very important that an English language teacher be patient, has emphatic thinking, and considers learners' needs. S/he knows that they are children and so they have different views of life. That's why, s/he designs learning environment accordingly. Also, a teacher should not be impatient; instead, should care about the thoughts of others. (EE8)

An English teacher should love and value his/her students, should not be prejudiced against them, and should take into account their all needs. (ST7)

An English teacher should know that each student is unique, so the instruction process should be designed accordingly. S/he must see each of them as a jewel. (TT5)

A large number of the participants acknowledged that a competent English language teacher should be open to both personal and professional development so that s/he can aware of the latest issues regarding ELT. Thinking so, they affirmed that an English language teacher must keep himself/herself up-to-date. In other saying, an English language teacher is also a researcher who searches new ELT developments. Some examples are as followings:

A teacher should be eager to learn new things because s/he is also a researcher. She or he should not be contented with her/his previous knowledge. The more a person reads and investigates, the more his/her attitudes and perspectives towards students change. This idea lets you be aware of the learners' needs, believe that each student is unique. (EE7)

I think that the most important feature that a teacher should have is developing himself/herself. That is, after graduating from the department, s/he should not consider that s/he is competent and that there is no need for looking for what is happening in the field. Instead, the teacher should have a continuous search for learning new developments in the field so that s/he can add more things to students, because children differ from each other, they were born with very different features, and affected by its surroundings very differently, so we need to know how we should address them. (T3)

Some participants listed having warmth, tolerance, empathy, sense of humour, establishing effective interaction, communication and cooperation with stakeholders, and being democratic and humanist as competencies that an English language teacher should have. They drew attention to the personality trait so as to underline its importance for being an effective English language teacher. Some examples are below:

English language teachers are those who should respect people that have different views and prototypes. They must be democratic and humanist. It is certainly important that teachers do not let their ideologies enter the classroom. (T8)

In addition to language proficiency, the personality is of importance. The teacher should have warmth and tolerance, and demonstrate a sensitive understanding of empathy. The teacher should consider how the students think and feel about learning so that learning process becomes enjoyable and productive. (PM2)

Character, temperament, conscience, and psychology are very important in teaching. If the teacher cannot act as a child while teaching to children, s/he cannot do or continue the teaching profession. S/he should embrace a democratic classroom. This teacher establishes an effective communication with stakeholders and interacts democratically. (EE2)

Students do not want a serious and surly teacher. Instead, if they face a teacher who has sense of humour, take a positive attitude toward school, teacher, coursework and each other. I believe that since humour increases students' motivation, it enriches not only the learning environment but also students' learning opportunities. (ST25)

4.1.2 Findings Related to the Perceptions of Stakeholders of ELTAE

By asking 'What are the perceptions of educational stakeholders of English language teacher assessment and evaluation (ELTAE)?', it was aimed to examine the perceptions of educational stakeholders towards the assessment of English language teachers. Another goal of this research question was to reveal the impact of the TAE on education and stakeholders. In other words, stakeholders both informed the readers about how they perceived teacher assessment model in Turkey, and they mentioned how this model affected student teachers, in-service teachers, and education. The data displayed that while student teachers postpone some skills until they PASS the examination, in-service teachers think that it is too late (PAST) to compensate these skills appropriately. According to other stakeholders (teacher trainers, education experts, policymakers), the current TAE prevents teachers to be CEO of their profession. The acronyms (PASS, PAST, CEO) are presented in the following sub-questions. In order to visualize the whole picture and in order to make it possible for the readers to understand how these categories affect each other, the researcher presented perceptions of all participant in a single figure in the part 5.1 The Need for IDEA (please see Figure 5).

4.1.2.1 Findings Related to Student Teachers' Perceptions of ELTAE

The first sub-question of this research question was 'What are the perceptions of student teachers of ELTAE?'. Student teachers voiced the challenges they faced due to the TAE model in Turkey. After reading the qualitative data carefully, four categories were found, namely "priority of system requirements", "anxiety", "skill delay", and "shortcomings". As can be seen in Table 18, the acronym PASS emerged from the initial letters of those four categories.

Table 18.

The Impact of the TAE on Student Teachers

Categories	Codes	f
Priority of system requirements	ÖABT	31
	Literacy & numeracy	25
	Educational sciences	8
Anxiety	Stress	31
	Fear of not finishing all subjects	28
	Social pressure	15
	Family pressure	9
Skill delay	Speaking	31
	Technology use	30
	Productive skills	25
	Listening & writing	15
Shortcomings	Good pronunciation	31
	Technology literacy	30
	Teaching practice	25
	Assessment literacy	20
	Material development skill	19

4.1.2.1.1 Priority of System Requirements

In general, student teachers were observed not being happy with Turkey's TAE model. They complained about the inadequacy of the model as it did not include and gauge all required competencies. According to the student teachers, they start to pay attention to components that are assessed, because thousands of student teachers apply KPSS every year and the government employs teachers among those who have the highest scores. To

be successful in the assessment and evaluation process, student teachers both join private courses and spend their time for studying only subjects that are assessed in the KPSS, such as ÖABT, general culture, general ability, and educational sciences. Some representative examples were as followings:

> In fact, Student teachers' literacy and numeracy knowledge, their speaking ability… I mean all competencies should be assessed. And it should not be only knowledge-based, but whether they can use this information in the real classroom environment must also be taken into account. The TAE model, unconsciously, leaves us no choice but the memorization of information. (ST3)

> As you know, there are many subjects in KPSS such as Turkish, math, history, educational sciences, content knowledge, subject knowledge… and many other issues. They are mostly knowledge-based and require good memorization ability. Therefore, we need to spend lots of time to memorize these subjects. (ST15)

> We have many friends who cannot really teach, because the current TAE model does not distinguish those who can act teaching profession very well from those who cannot. Very recently, interview has been introduced, but it does not serve the intended purposes, because no rubrics or criteria are used for evaluation. So, student teachers attach importance to their appearance. For example, the interviewee wears a beautiful dress, gussies up, or wears his tie during the interviewee time so as to make a good impression. (ST10)

> I do not believe in the fairness of KPSS as it does not gauge all competencies. As a result, to be honest, I do not allocate time to topics that are not on the exam. I can handle them later; I mean when I become an English teacher. At the moment, it is time to study topics of ÖABT, general culture, general ability, and educational sciences. (ST30)

4.1.2.1.2 Anxiety

Student teachers complained about the stress and fear that aroused due to the nature of the current TAE model. They accused KPSS of having an abundance of topics that needed to be memorized and ignoring skill and attitude-related competencies. Naturally, lots of topics cause stress due to fear of not finishing all topics. Here are some representative examples:

> Each examination may cause stress. However, the case of KPSS is totally different. It only takes into account student teachers' memorization, which means that we need to memorize hundreds of techniques, methods, approaches, theories, and their underlying principles. Comprehending them requires a great deal of effort and time. I do not know whether we will cover all of them or not. This uncertainty really frightens me. (ST2)

> I am a good student at the department. Yet, I have enrolled on a course for KPSS, because there are a lot of issues to be studied. I have started to see horrible dreams. I am not sure about finishing the whole required documents. Thus, I am really afraid of KPSS and this fear brings stress. (ST9)

Some of the student teachers also stated that, in addition to stress, both family and social pressure contributed to the occurrence of anxiety. According to them, their families and surroundings expect to pass examinations as soon as they graduate from the university. Here are some comments of the student teachers:

> Once you graduated from university, your family consider as if you were an English language teacher. This expectation is causing a serious feeling of anxiety. (ST4)

> If you are not recruited to a public school even in the first year of graduation, your friends, relatives, and neighbors start to see you as an incompetent person. For this reason, even the thought of not being assigned scares me. (ST11)

> KPSS equals anxiety. Even if my parents do not tell me directly that they anticipate me to pass KPSS so as to be recruited to the public schools, I can see these expectations when I look into their eyes. (ST21)

According to some student teachers, TAE model ignores many features that need to be considered while assessing and evaluating student teachers, such as their university entrance score, their success at the university (e.g. grade-point average, practicum), and so many other important features. To clarify the issue, here is a representative example suggested by a student teacher:

> Examination conditions are not fair and equal. For example, for studying in this university, I got 434 points (T8) while my got into a university with 370 points... or for instance, my grade average is 3.85 out of 4 whereas there also students whose grade average is 2.00 out of 4. However, TAE model does not take into account these scores while evaluating student teachers. Instead, the only determinant criterion is the score they receive in KPSS. This unfairness and inequality is the source of the anxiety. (ST5)

4.1.2.1.3 Skill Delay

As mentioned before, student teachers stated that since they were not asked to demonstrate their teaching ability and language proficiency, they allocated their time to topics that were on the exam so as to response to the system requirements, which caused student teachers to feel under pressure. Hence, to overcome anxiety, they started to ignore skills that were not on the KPSS. The most ignored one is the speaking skill which was emphasized by all participants. In this respect, a few examples were:

> Although we know that speaking is a must for an English language teacher and pronunciation is an important component of speaking skill, we delay speaking skill due to the urgency of ÖABT, general culture, general ability, and educational sciences. (ST11)

92

> I really care about pronunciation and speaking but I have to admit that I do not pay enough attention to them. Nevertheless, I spend most time to other issues that we are asked to do in KPSS (e.g. teaching knowledge, content knowledge, general culture, general ability, and educational sciences), because I need to comprehend these components in order to be successful in the examination. (ST6)

All of the student teachers (except one of them) underlined the importance of technology. According to them, an English language teacher must be capable of technology literacy. However, they warned that as TAE model forced them mingle with the components of KPSS, they decided not to spend time with technology related things:

> I do believe that an English language classroom should be equipped with technological tools. Likely, an English language teacher must be technologically literate. Technology literacy includes not only the knowledge of turning on/off the computer, or using web tools, but also it refers to the ability of utilizing these tools for accessing, constituting, combining, evaluating, interpreting and communicating information. Thus, technology skill requires more interests and time. However, there is lack of time, so we feel obliged to focus on the topics we are asked to do. In this case, like many other skills, issues related to technology are postponed to another time. (ST12)

> Technology is constantly evolving and changing. Something that is very popular today may be outdated tomorrow. Thereof, the teachers should be interested in the new developments in order to be aware of them. And this interest requires more time and effort, but we do not have so much time and our technology literacy is not assessed. So, we give priority to the topics on the exam. (ST14)

Some student teachers asserted that they had to put the writing and listening skills aside due to TAE model:

> All four skills of an English language teacher should be assessed before they are employed. For example, he or she can be asked to write a coherent letter or something else. If our TAE model asked teachers to demonstrate a writing performance, all English language teachers would write coherent and cohesive essays. (ST24)

> When I did listening activities during practicum, I realized that I was having trouble with some listening parts. We do not have an intensive English speaking and listening program, unfortunately. In addition, these skills are not assessed. So, we delay these skills till we pass the KPSS. (ST16)

4.1.2.1.4 Shortcomings

As mentioned before, student teachers stated that since they were not asked to demonstrate their teaching ability and language proficiency, they allocated their time to topics that were on the exam, which resulted in delaying some skills. Consequently, they graduated from

university with lack of some skills. One of these skills is the lack of a good pronunciation that causes non-fluent and inaccurate speaking ability. All student teachers blamed the TAE model for the occurrence of the bad pronunciation and inaccurate speaking ability. A few of their statements are below:

> The biggest disadvantage of this university is the fact that it does not have an English-medium education system and Turkish is mainly used in many lessons. Actually, only a few lecturers use English intensively, so we are not very fluent in speaking English. But as we do not have enough free time due to topics of KPSS, we do not pay the required importance to the speaking skill. We find postponing speaking skill more reasonable, thinking that we can deal with it after being recruited a public school. (ST17)

> Now I am busy with some other issues… if our speaking, writing, or listening were assessed, we would take into account those skills… but after I become an English language teacher, I will spend a lot of time to enhance my, especially, productive skills. (ST27)

Almost all of them agreed that, in addition to speaking skill, technology literacy skills were also ignored:

> Technology and technologically literate teachers are sine qua none of English language classrooms. However, the lack of time for preparing KPSS related topics make us disregard technological developments. (ST19)

> I am not technologically competent, and perhaps I will graduate from university with that shortcoming, because I feel obliged to focus only on the topics that are on exam. (ST29)

Some student teachers stated that they lack teaching practice, assessment literacy, and knowledge of material development skill because of the TAE model. Some representative answers were:

> We are not asked to demonstrate our teaching ability, so we allocate our time to the requirements of the examination (KPSS). We are observed once or twice during the practicum. I have heard that some universities did not supervise their senior students. However, in this process, senior students should be observed many times and assessed regarding their teaching ability, classroom management, material preparation and development, interaction with pupils, preparing lesson plan… etc. Since those competencies are taken into consideration, we finish the department with shortcomings. (ST22)

> I wish the TAE assessed and evaluated our assessment literacy. Because if it did, both we and the lecturers would spare time to develop our assessment knowledge. Like me, a lot of friends do not know how to assess young learners. Likely, we are not competent enough to assess and evaluate the speaking skill analytically. (ST20)

4.1.2.2 Findings Related to In-Service Teachers' Perceptions of ELTAE

The second sub-question of this research question was 'What are the perceptions of in-service teachers of ELTAE?'. When analyzing the data gathered from the in-service English language teachers who were requested to describe the impact of the TAE model, it was observed that their explanations were in parallel with the shortcomings that student teachers mentioned due to TAE model in Turkey. Four categories emerged from codes, namely "practice", "assessment", "speaking", and "technology literacy". As can be seen in Table 19, the acronym PAST emerged from the initial letters of those four categories.

Table 19.

The Impact of the TAE on In-Service Teachers

Categories	Codes	f
Practice	Learners' needs	25
	Learners' grade	25
	Classroom management	19
	Effective use of time	15
Assessment	Assessing learners	25
	Assessing skills	25
Speaking	Intelligibility	25
	Fluent speaking	23
	Accurate pronunciation	23
Technology literacy	Using technological tools	23
	Using technological programs	22
	Combining technological, pedagogical and content knowledge	18
	Using web tools	17

4.1.2.2.1 Practice

In-service teachers were invited to recount the impact of the TAE model in Turkey. Like student teachers, they found the TAE model inadequate as it did not include and gauge all required competencies. They were found to dwell on shortcomings underlined by student teachers. The first one was about the teaching problems they encountered in the real classrooms. Some representative examples were as followings:

> I was recruited to a primary school as soon as I graduated from the university. The first year was very difficult, and I told myself: "Yes Aylin (pseudonym), welcome to the real world!". In fact, I had a successful practicum experience, but I did my practicum at a high school. I had never taught to young learners, who are totally different from adults. Young learners and adults differ in many respects (for example, their cognitive, social, emotional, and physical needs and development. Even though we had Teaching English to Young Learners lesson, we had difficulty in teaching to young learners in the real classes. I had to spend a few months to read articles and books in order to master their needs and development (cognitive, social, emotional, and physical). (T4)

> During the practicum, you see how the real classroom environments differ from the "virtual world" which was created in micro teaching activities. You witness the troubles that are faced in real life. Our mentor did not give us any feedbacks. To be honest, she did not even join lessons for two months. That's why, we were not informed about how to overcome problems when facing troubles. These are due to the limitations of the TAE model, because if the practicum had been considered when recruiting English language teachers, all institutions and stakeholders would have paid attention to our teaching abilities. (ST10)

> While performing micro teaching tasks, our lecturers were instructing us carefully. In addition to feedbacks, they also guided us about how to behave in order to create a real-life-like classroom. In my turn, for example, the lecture said: "One of you will cry… You two, do not join the lesson… etc.". Because of the fact that we were trying to help our friends during micro teaching activities in order to prevent them getting bad scores, every stage seems to work well. However, in reality, we had both time related problems and problems regarding classroom management. (T25)

4.1.2.2.2 Assessment

In-service teachers declared how they were faced with assessment problems due to the nature of the TAE model. According to them, the assessment related questions in KPSS gauge candidates' knowledge of certain terms (e.g. validity, reliability, mod, median, and formulized questions). Some of them stated that the scope of the testing, assessment and evaluation related lessons should be extended so that teachers could be able to assess not only young learners and adults, but also they could gauge learners' skills such as speaking, writing, listening, reading, grammar… etc. In-service teachers stated that the reason behind

96

the in-service teachers' being lack of assessment literacy was due to the TAE model since it does not consider such issues. Here are some representative examples:

> Theoretically, we learned lots of things about how to assess both young learners and adults. But that information does not work well unless you use it for real interlocutors. In my opinion, it resembles a driving test. You may get the highest score, but it does not guarantee that you are a good driver. (T21)

> I was recruited to a pilot school in which there was an intensive English program. Fifth grade students were taught English for 15 hours a week, and teachers focused on each skill (for example, speaking, writing, listening, reading, grammar) separately. During the first year of my teaching experience, I realized that I was assessing both students and their skills holistically. Last year, I join an ELT conference where I noticed my assessment shortcoming. (T11)

4.1.2.2.3 Speaking/Pronunciation

As mentioned before, in-service teachers were not happy with the TAE model since they were not asked to demonstrate their speaking ability and language proficiency. According to them, speaking is a topic that needs to be focused on, especially during undergraduate period. As it was discussed in 4.1.2.2 Findings Related to Student Teachers' Perceptions of ELTAE, speaking ability was delayed by student teachers due to the ignorance of the TAE model, and speaking skill was believed to be developed later. However, in-service teachers indicated that speaking skill must be dealt with during undergraduate education. Otherwise, it becomes too late to overcome that trouble. In this respect, a few examples were:

> I strongly believe that our language proficiency should be assessed via examinations such as TOEFL and IELTs. I do believe that there is a positive correlation between the level of our language proficiency and the TAE model. I have problems with my proficiency, especially with pronunciation. And, to me, after graduating from the department, it is not easy to overcome it. We should have managed it during undergraduate period. (T15)

> I have friends who had terrible pronunciation. These friends are English language teachers. I really feel sorry for the students in their classes. This situation is a proof of the insufficiency of the TAE model in selecting and recruiting teachers. (ST16)

> I have read somewhere that if you do not test a skill, the related skill will be ignored by teachers and lecturers. Consequently, expecting that skill be developed is nor reasonable. The TAE model and the situation we are in are a good example of that saying. Our skills are not assessed when being recruited to public schools. That's why, neither did us nor lectures pay attention to those skills, specifically speaking ability. The result, as you know, is not pleasant. (T12)

4.1.2.2.4 Technology literacy

In-service teachers stated that they graduated from university with limited knowledge of technology. In-service teachers blamed the TAE model for this limitation. Most of them agreed that technology literacy skills were ignored, so they are not technologically competent. The perceptions of some in-service teachers support the idea that if the examination board starts to give more place and importance to the candidates' ability of technology use, English language teaching programs will revise their curriculum by adding new technological courses so that not only will teachers be aware of the new technological development (e.g. hardware, software, web tools etc.) but also they will be able to use technology for teaching English. A few of their statements are below:

> We were not exposed to the technology when we were students at the university. The only technology related lessons were Computer I and Computer II in which we learned how to use Microsoft word program. I think that there should be lessons that would enhance student teachers' technology literacy. I graduated from the university with technological shortcomings, because of focusing only on the topics that were on examinations. (T17)

> The undergraduate lesson might contribute to our knowledge about the technology but we are not able to use it effectively. We need to be supported to gain experience using the technology so as to be able to accomplish technological tasks. (T13)

> Basically we might learn what this technology could do, (which is about the declarative knowledge). However, the TAE model should ask us to answer how we use this technology, (which is a form of procedural knowledge). If we become able to answer the second question, our technology literacy will improve. (T23)

> In this era, each student deserves a technologically literate English language teacher. There should be a continuing support via both undergraduate lessons and INSET programs. In addition, the technology literacy must be measured to reinforce its importance. (T18)

4.1.2.3 Findings Related to Teacher Trainers' Perceptions of ELTAE

The third sub-question of this research question was 'What are the perceptions of teacher trainers of ELTAE?'. The data gathered from the teacher trainers were transcribed carefully in order to display their perceptions towards the TAE model in Turkey. The perceptions of teacher trainers play an important role, because the future teachers are educated, prepared and graduated by teacher trainers. According to them, the current TAE model does not gauge a student teacher's strengths and weaknesses through multiple choice questions. Instead, it chooses the student teachers who are good at memorizing knowledge. The biggest problem, as they stated, is that it does not provide any information

about whether the candidate is able to teach or not. That is, it is inadequate and fails to differentiate among student teachers. Teacher trainers had the following opinions:

> A teacher's literacy and numeracy are important, and they are assessed in KPSS. There seems to be a connection between ÖABT and the courses that they study at the undergraduate level, so ÖABT can be welcomed. Nevertheless, these components are not informative in revealing what a teacher can do in classroom. Do you know what it means? It means that the current TAE model in Turkey does not distinguish those who can teach well regarding learners' needs from those who are far away from the teaching profession. If it is a teacher evaluation exam, it should take into account the teacher's teaching performance, language proficiency based on four skills, classroom management… etc. I mean a student teacher's performance must be observed in order to see whether s/he can interact with students, whether s/he can teach, whether s/he can speak target language fluently, whether s/he is competent in required subjects or not… etc. (TT1)

> Turkey is experiencing an oversupply of English language teachers, so there should be some kind of examinations. However, the problem is related to the structure of the TAE model. It gauges knowledge of student teachers through paper-pencil examinations. Any graduates of ELT programs (or English language and literature, American Culture and Literature, Department of Translation and Interpreting, Department of English Linguistics as long as they have pedagogical formation) can pass the examination, if they are good at memorizing information. We need a system which will eliminate unqualified student teachers whereas choose the right person for the teaching profession. (TT9)

Some teacher trainers underlined the importance of teachers' personal and professional development. According to them, the TAE model of a country must evoke the necessity of looking for new ELT developments to the student teachers. They stated that if student teachers gain positive attitude towards openness to personal and professional development, they will always follow new developments even after being an English language teacher. And they warned that if not, they may act teaching profession based on their prior knowledge:

> The current system dictates unconsciously that a student teacher must spare his/her time for studying knowledge-based information due to the format of the TAE model. That's why, student teachers do not gain the habit of developing themselves. As a result, when they become teachers, they try to conduct lessons based on their undergraduate knowledge. But, we know that teachers, especially English language teachers, must be technologically literate to make lessons both productive and enjoyable. For having required technological knowledge and for being aware of the new developments in language teaching, an English language teacher needs to be open to personal and professional development. (TT2)

As stated by some teacher trainers, although competencies that an effective English language teacher should have may differ from country to country, the underlying principles are common in all contexts, such as having good language proficiency, being

technologically literate, having required knowledge and skills of effective teaching…etc. They suggested that the TAE model should be designed or revised accordingly. They added that assessing and evaluating student teachers in our country is used for the purpose of eliminating candidates via multiple choice questions. So, those who do not deserve teaching job may be recruited to public schools because of indistinguishability of the TAE model:

> On the one hand, we suggest our student teachers to enhance their language proficient, follow recent ELT developments, improve their assessment and technology literacy, pay attention to teaching practices, know how to prepare and use materials appropriately, and be informed about curriculum and legislation, but on the other hand, nearly none of these competencies are kept in mind while recruiting teachers to public schools. We urgently need a teacher assessment system that can differentiate among student teachers. In order to do so, multiple sources should be applied:
>
> 1- student teachers' scores of micro teaching activities,
> 2- their scores of material evaluation course,
> 3- their practicum scores,
> 4- their language proficiency scores based on four skills. (TT8)

4.1.2.4 Findings Related to Education Experts' Perceptions of ELTAE

The fourth sub-question of this research question was 'What are the perceptions of education experts of ELTAE?'. The fourth sub-question aimed to unveil the education experts' perceptions of ELTAE as well as benefit from their suggestions. These education experts were chosen purposefully for three reasons: (a) they contributed to the last national education councils, (b) they were those who played active role in determining the content, duration, and format of ÖABT, (c) they have been invited to almost all teacher and teaching related workshops organized by the MoNE, ÖSYM, and YÖK since the last decade. So, their perceptions of Turkey's current TAE model are of importance. According to them, the rationale behind the current TAE model is due to the practicability (or manageability) issues. They declared that, notwithstanding the issue of practicability, the current evaluation model should be revised immediately or new components should be added with the aim of developing a new TAE model that can determine whether the student teachers are able to teach and whether they are the right persons for the teaching profession or not. Education experts had the following opinions:

> Both in the national education councils and in development plans, all stakeholders emphasize the importance of teacher effectiveness since an effective teacher motivates and engages his or her students in learning, and uses techniques which best fit his/her students' learning

needs. From this point of view, having effective teachers in all classrooms becomes a hot issue in these councils and development plans. We know that without a good teacher assessment and evaluation model, it is not possible to bring students and effective English language teachers together in a classroom. So, a great attention should be paid to the way how a language teacher is selected and recruited to the public schools. The current TAE model in Turkey does not distinguish effective student teachers from ineffective ones, because it does not consider the student teachers' teaching performance and especially their language proficiency based on four skills. There are English language teachers who cannot speak target language fluently. Likely, there are English language teachers who just do the job for economic concerns. The TAE model should ensure that the right person is selected. It may not be as practical as the current one, and it may also require more time, effort and money; however, it worth developing and conducting such assessment and evaluation model for a quality education. (EE3)

In all educational conferences, the importance of 21st century skills is emphasized. As it is known, we are in a digital society, and in order to contribute to students' success in this rapidly changing environment, we should provide them with 21st century skills, such as technology literacy, leadership, productivity, communication. However, the current TAE model does not allow us to assess these skills. This model should be revised as soon as possible so that it could take into consideration the student teachers' communication ability, classroom management, leadership, technology literacy, curriculum literacy…(EE8)

In our context, the TAE model may require some kinds of examinations for the purpose of evaluating student teachers, but it should not only consist of paper-and-pencil tests. Instead, the content of the exam should be more practice oriented in terms of quality and scope. Because it is not possible to assess all competencies that an English language teacher should have via paper-and-pencil tests. From this aspect, the current TAE model has deficiencies. (EE5)

Like teacher trainers, education experts stated that the underlying principles of competencies that an effective English language teacher should have are common and they can be listed as having required assessment knowledge, having a good language proficiency, being technologically literate, having required knowledge and skills of effective teaching. They added that although the TAE model is being implemented due to the conditions of the country, it needs to be revised as soon as possible:

The aim of the current TAE model is to find some solutions for the existing complaints and reactions. There are a lot of candidates, and you just allege the results for the rationale behind the elimination. That is, this model does not help you choose student teachers who have required assessment and technology literacy; who know how to prepare and use materials appropriately; who always look for new ELT developments; who have the required information, ability and attitude regarding teaching English; who have the required information of curriculum and legislation. From this point of view, the current TAE model should be revised as soon as possible so that it could take into account the aforementioned competencies. (EE8)

4.1.2.5 Findings Related to Policymakers' Perceptions of ELTAE

The last sub-question of this research question was 'What are the perceptions of policymakers about ELTAE?' In order to display the perceptions of policymakers towards the TAE model in Turkey, the data were transcribed carefully. Their perceptions are of importance, because the final decisions are shaped by them. Like other stakeholders, they believe that it is not possible to assess and evaluate student teachers via current TAE model as it does not distinguish effective teachers from ineffective English language teachers through multiple choice questions. Instead, it is applied, as they stated, to eliminate the student teachers without providing any information about whether the candidate is good at teaching or not. Policymakers voiced the following opinions:

> Such examinations have two drawbacks. Firstly, there is no threshold level for passing the examination. For example, ÖABT has fifty questions. If there is teacher shortage, even those who have only five correct answers may be considered to be successful. However, in the case of oversupply, even those who have more than forty-five correct answers may not pass the exam. This shows that the current TAE model is not utilized to choose effective teachers. Ten thousand student teachers want to be assigned to public schools while you need one hundred teachers. In this sense, from the perspective of the surveyor, the difficulty level of the questions becomes important, because as you know, it's compulsory to publish the asked questions as a result of court orders. So, what happens then? Once you ask a question in KPSS or ÖABT, you cannot ask it again. The struggle to find new questions may cause irrelevant and difficult questions. As ministry, we expect teachers knowing subject knowledge and some basic issues, but irrelevant information and questions with poor quality are asked instead of basic and qualified ones due to aforementioned reasons. Second drawback is the indiscriminate side of these exams. For example, let's assume that a student teacher has answered the whole question correctly; does it mean that s/he is a good teacher? Absolutely not, because knowing all correct answers shows the student teacher's analytic intelligence, not teaching ability. And it does not always mean that s/he will be a very good teacher. In order to display the teaching ability, we need other sources, such as interview, a sample lesson, and so on. (PM5)

Some policymakers underlined the importance of speaking, writing and speaking skills and stated that these skills were sine qua none of an effective English language teacher. They warned that some English language teachers had difficulty with basic skills, especially speaking, and that the TAE model should be revised as soon as possible so that it could display the English language proficiency of student teachers:

> As you know, some foreign institutions (e.g. British Council) do research on the state of English in schools in Turkey, specifically in public schools. According to some of these reports, some English language teachers who works in public schools do not feel themselves adequate in English. That is, the TAE model does not select student teachers who have required language proficiency. Can you, please, convince me that those teachers who lack of

speaking skill can make our children speak English well? So, the current TAE model should be revised immediately in order to ensure that English language student teachers are good at speaking and writing English, and they can understand native and nonnative English conversation. (PM3)

Policymakers indicated that the current TAE model had some limitations and they added that it was aimed to remedy this deficiency with the introduction of ÖABT and interview, but there is still a need for adding new components to the TAE model so that it can distinguish those who are good at teaching English, interacting with stakeholders from those who are not. Here are some representative opinions:

> As you know, in order to be a teacher, a student teacher should have required field knowledge in addition to basic knowledge of education. Motivating from that fact, ÖABT has been introduced. However, it also has limitations because it is prepared in multiple choice questions, which is considered to be more economical and practical. But ÖABT is insufficient to decide whether a person can teach English well. Let me share an experience with you. This petition belongs to a person who studied at Turcology department in Macedonia. He asks for accreditation so as to work in his country (Turkey). Please look at that petition! He can be gauged via that petition. In my opinion, assessing him in such way is more valid than using tests. (PM2)

> Those who graduated from a teaching department consider themselves as a teacher. But this is not the case. Graduating from those programs displays that you acquire objectives in that field. It does not warrant that you can perform that job. In many countries, particularly in Europe, a number of processes were run during teacher selection. We have a misconception: Anyone finished a teaching program should be recruited as a teacher. Everybody receives a pedagogic formation certificate and asks government to assign them to the public schools. We should change such kind of opinions. We should distinguish between one's academic knowledge and his/her ability. Therefore, there should be some different selection mechanisms, such as knowledge-based examinations (as in current model), interviews, and performance-based examinations. (PM1)

4.1.3 Findings Related to the Appropriate ELTAE

'What might an appropriate EFL teacher assessment and evaluation be?' is the fourth and last research question. The aim of this question was to examine the perceptions of educational stakeholders about how student teachers should be assessed and evaluated in order to be English language teachers in public schools. Education stakeholders' perceptions of education policy and their opinions on the appropriate English language teacher assessment and evaluation model may contribute to the test validity. That is, education stakeholders' perceptions and opinions of ELTAE were used to develop a valid and fair TAE model. Since stakeholders seemed to reach a consensus regarding the

appropriate teacher recruitment model, we decided not to present the opinions of all stakeholders according to their status (e.g. student teacher, in-service English teacher, teacher trainer, education expert, and policymaker). Instead, we categorized the data according to the features that an appropriate TAE model should have. The data suggested an IDEA to the authority about how to develop a fair, valid and reliable evaluation process for recruiting English language teachers to the public schools. The acronym IDEA emerged from the initial letters of four features that a good TAE model should have; namely, *inclusive, distinguishable, equitable/fair, accurate*.

4.1.3.1 Inclusive

The term inclusive refers to the features that should be involved in an appropriate TAE model. It informs the authority about the competencies that need to be assessed for a good evaluation. According to education stakeholders, not only must an appropriate TAE model include issues related to student teachers' knowledge and skills of teaching profession, but also their attitudes towards it. They provided good example of using different sources for assessing and evaluating student teachers so as to select the right person for the teaching profession. Here are some representative statements:

> An appropriate TAE model must assess all teaching related issues. It should gauge not only student teachers' knowledge of general culture and general ability, but also their speaking skill. The current components are only knowledge-oriented. The interview is conducted for the sake of formality. Likely, those who do interviews are not experts. The practicum is not an informative process, and it does not give any idea about whether I will be a good teacher or not. Therefore, student teachers should be asked to demonstrate a sample lesson before being recruited to public schools. (ST5)

> The current system pays attention only to student teachers' memorizing ability. Thus, we face teachers who cannot speak English or who have terrible pronunciation. In my opinion, an appropriate TAE model must ask student teachers to prove their language proficiency via TOEFL or IELTs equivalent exams. Because since the current TAE model does not take into account the student teachers' language proficiency level, neither English language teaching programs nor student teachers do not pay attention to it. (T8)

Some of the participants criticized the current model for being lack of providing information about practice. They claimed that since the current model does not include the required components for evaluating student teachers, its decisions of 'pass or failure' does not guarantee an individual's good or bad teaching.

The current model consists of questions about student teachers' literacy, numeracy, field knowledge, language proficiency, and many other topics. All questions are in multiple choice formats. Imagine that a student teacher has answered the whole questions correctly; does it mean that he will be a good teacher? Absolutely not, because it presents nothing about whether the student teacher can transfer his/her knowledge into practice. I mean we do not know whether he will be able to teach English to students appropriately. That's why, we need to benefit from other resources such as observing the student teacher while teaching. (PM2)

You know, a teacher is also considered to be a researcher. So, a researcher teacher should have good technology literacy so that s/he could search new developments in the field. In addition, we also value 21st century skills as English language teachers need them in order to respond to the needs of their students. Among these skills are technology literacy, leadership, productivity, collaboration and communication, and creativity. In order to assess all of these characteristics, we should redesign the assessment and evaluation system. For example, student teachers' undergraduate experience and practicum may be kept in mind while evaluating them. (EE8)

Some of them suggested that the use of practicum and micro-teaching process could be used as informative data for student teachers' teaching abilities and their personality traits:

The evaluation model must be revised to give information about the student teachers' skills of and attitudes towards teaching. In this sense, practicum process can be very informative. Likely, universities can collaborate with the ministry by using the information observed during that process. They may also use student teachers' marks of micro-teaching lessons. As ministry, we have revised the practicum system by adding a form, which will be filled out by both mentor and supervisor. That form will be saved electronically and it will be utilized during interview phase. (PM1)

4.1.3.2 Distinguishable

According to participants, an appropriate TAE model must distinguish effective teachers from ineffective teachers. They emphasized that the current model is inadequate in displaying the skills and attitudes that an effective English language teacher must have, because it consists of multiple choice questions which gauge only student teachers' knowledge while ignoring their teaching skills and their attitudes. They insisted that an appropriate TAE model must choose only those who were born for teaching profession. In order to achieve this goal, not only knowledge but also skills and attitudes of student teachers must be considered while assessing and evaluating them.

I have a lot of friends who will not be able to teach English. However, during interview process, they pay attention to their wearing in order to give good impression and unfortunately such kinds of behaviors sometimes work. As a result, they may be recruited. Yet, normally, I do not want them to be given any chance of teaching English to children, but

they may be chosen for their nice appearance… I mean, as you know the interview is determinant in deciding whether the student teacher fails or passes, so the appearance may cause halo effect. During my practicum experience, I faced some English language teachers who could not speak English. They were recruited because of their good scores of the KPSS. (ST3)

The examinations should be revised so that they could distinguish those who are good at writing, listening, and speaking from those who can do teaching profession well. If a person is good at English language and culture, I think that s/he can teach it easily. Personality, conscience, and psychology are important terms for teaching. A teacher should know how to behave according to students' age and level, if not, then, s/he cannot teach English appropriately. If the practicum process is well evaluated, the whole aforementioned features can be taken into consideration. (T11)

Some voiced that the current TAE model should be revised as soon as possible so that it can measure student teachers' skills and attitudes. That is, it is not possible to distinguish effective student teachers from others unless their skills and attitudes are taken into account:

Knowing what to teach and how to teach it is sine qua non of an effective English language teacher. In addition to them, personality, psychology, loving teaching job, conscience, assessment literacy, communication and collaboration, and technology literacy are a must. If we want our children to be taught by language teachers who own these features, we must assess those characteristics. Because once we ask student teachers to demonstrate the aforementioned aspects, they will attach importance to the asked aspects. Let me give you a specific example, after the introduction of ÖABT, the final year students rejoin the linguistics courses with freshmen. They gain a chance to reconsider the first year subjects from the last year viewpoint. In the case they cannot study, they may prefer to come and listen to lessons if the weekly course schedules are also available. This situation constitutes a good model for freshmen. They realize that, even after graduation, they will need the information they learned in the courses they are taking. (TT3)

4.1.3.3 Equitable/Fair

As stated by education stakeholders, since teacher assessment and evaluation plays a crucial role in education, whether the assessment and evaluation practices reflect and promote fairness should be remained on the agenda of educational policy. The term equitable/fair refers to a test's unbiased decisions. Considering test fairness, bias may refer to two vital threats to validity, namely "construct under-representation or construct-irrelevant components of test scores that differentially affect the performance of different groups of test takers" (American Educational Research Association et al. 1999 as cited in Weir, 2005, p. 210). In other saying, in order to evaluate student teachers fairly, firstly, the test should represent the construct; then, precautions must be taken to prevent certain

groups of test takers (e.g., English language and literature student teachers) from being more affected by a construct-irrelevant source of variance than other groups and to avoid unintended constructs (construct-irrelevant sources), because they may cause incorrect inferences about those who take tests and, therefore, decrease validity. Here are some representative statements:

> I sometimes have difficulty in understanding the present assessment and evaluation system. I was so glad when ÖABT was introduced, because for the first time there was an attempt to choose student teachers who are good in the field. However, it does not serve that purpose in reality. As you know, literature related questions require more specific knowledge, so answering them correctly plays a huge role in getting high scores. Thus, unfortunatelly, graduates of English language and literature have the edge on us. They spend four or five years particularly on literature and, therefore, know almost all literature related specific details. (ST11)

When the data were analyzed, it was noted that student teachers or in-service teachers of English language and literature programs feel themselves more comfortable and relaxed in terms of ÖABT. According to their statements, they are aware of the fact that the most specific and memory-based questions are those related to literature, and those who are not graduates of literature departments have difficulty in answering them correctly. This provides advantages to student teachers or graduates of English language and literature departments:

> To be honest, thanks to ÖABT, I really feel secure for some reasons. Firstly, as you know, ÖABT consists of four components: language proficiency, linguistics, methodology, and literature. During undergraduate education, we become familiar with all these components except methodology, but we studied subjects related to methodology via pedagogical formation training. We are studying English language and literature, so we are successful in doing questions both related to literature and those related to language proficiency since we are always asked to read and write essays during the four-year education. Last but not least, we have the whip hand owing to literature questions. (ST19)

One teacher trainer accused some ELT programs of ignoring productive skills and producing graduates who cannot speak English due to the nature of the current recruitment model:

> One of my MA student studied self-efficacy and beliefs of the senior students of departments of ELT and English Language and Literature departments and examined whether they feel themselves competent. The results were surprising. In language proficiency, the senior students of English Language and Literature department were found to feel more confident. Considering the senior students of ELT department, there was a lack of confidence in the language proficiency, but they felt competent in method and field knowledge. By the same token, the senior students of English Language and Literature department were noted to be

more successful than those of ELT departments. This calls for an urgent revision for the current model, because the ELT departments ignore four skills due to the nature of the current TAE. From that point of view, the current TAE model cannot be considered to be fair. (TT2)

4.1.3.4 Accurate

When the data were analyzed, it was observed that education experts complained about the scoring method of the current TAE model. Accurate has two sides. On the one hand, they suggested an accurate scoring, underlining the importance of using the right assessment tool for the related competency. For example, they warned that paper-pencil examinations in multiple-choice format do not guarantee that the right person will be selected for teaching profession. That is, they complained about the structure and content of the current evaluation model since it did not include skill and attitude based scores. Here is a representative response:

> I have some classmates who cannot interact with others. Few of them have psychological problems. They have achieved to get accepted into Gazi ELT department, and perhaps they will pass KPSS and become teachers of our children due to the format of current evaluation system. No attention is paid to a student teacher's personality, attitude, or whether s/he is good at teaching English. Those who spend time memorizing information that is asked in KPSS become successful in this evaluation system. However, teaching requires, at least, two or more parties. In the simplest term, a person needs students in order to teach something. So is he or she really good at teaching, collaboration and cooperation, classroom management, planning a lesson, and using the right assessment and technological tools? In my opinion, the current TAE model resembles the lottery. We do not know who will hit the bullseye! But it should not be like that. New components should be added and their sores should be used when recruiting someone so that those who like teaching be selected. (ST8)

On the other hand, they complained about the lack of using multiple assessment and evaluation sources. In other words, various methods of assessment should be used to gauge student teachers' knowledge, skills and attitudes. As they stated, relying on only multiple choice paper-pencil examinations to provide information about student teachers' teaching ability will not reflect the reality, because the current TAE model only measure 'what' student teachers know about certain topics while ignoring the 'how' and 'why' questions. Student teachers were also observed to be pessimist about interview. They criticized it for being lack of field experts and scoring rubric and, therefore, it diminishes the reliability and validity of the TAE model. A representative statement is as following:

108

You cannot create conditions in which each student is taught by a qualified language teacher unless you take into account the score of student teachers' skills and attitudes when evaluating them. The condition of our education represents a good example for that situation. The current system does not regard whether a person is motivated to teach English to students and whether that person is equipped with the required skills. Therefore, it is common to encounter English language teachers who cannot speak and/or teach English. I mean, since the scores are not accurate and do not represent the student teacher's skills and attitudes, it does not mean that those who are recruited love teaching profession, and vice versa. (EE5)

4.2 Quantitative Findings

The third research question was 'What are the attitudes of in-service EFL teachers towards ELTAE?' about the quantitative part of the study and it was aimed to investigate the attitude of in-service teachers towards the current TAE model.

4.2.1 Findings Related to In-Service EFL Teachers' Attitudes towards ELTAE

This part examined the findings related to relationships between demographic variables of quantitative phase and English Teachers' Attitudes towards Recruitment System (ETARS) together with its dimensions. In addition, in this quantitavie part, the significant levels of the scale and its dimensions were examined in terms of some demographic variables such as gender, graduated department, grade taught, reason for choosing department, and desired profession. ETARS has three dimensions/factors, namely, *positive attitude towards TAE* (D1), *being interested in TAE* (D2), and *negative attitude towards TAE* (D3). The statistical technique used in the data analysis can be in the form of either 'parametric' or 'non-parametric' tests according to the distribution of the data. The normality distribution of the relationships between demographic variables and dimensions D1, D2, D3 together with the item total score of ETARS was analyzed. The 'Z test', in which the statistical significance of the skewness index is tested, is one of the tests that can be used for this purpose. The fact that the p value calculated in the test result is less than $\alpha = .05$ proves that the scores come from the normal distribution (Field, 2009; Howitt & Cramer, 2011).

Table 20.

Test of Normality

	Kolmogorov-Smirnov		
	Statistic	df	Sig.
Total	.062	257	.018
D1	.105	257	.000
D2	.069	257	.005
D3	.140	257	.000

As can be seen in the Table 20, parametric tests were used to determine the relationship between demographic variables and dimensions D1, D2, D3 together with item total scores on the basis of the normal distribution of the sample. For this reason, t-test (Independent Samples Test) was used for the demographic variables consisting of two categories, while one-way ANOVA tests were applied for the demographic variables consisting of 3 and more categories. The sample of this study shows a normal distribution ($p>.05$).

Table 21.

Scores Related to the Effect of Gender Variable on ETARS

	Gender	N	Mean	Std. Deviation	t	df	p.
Total	male	82	68.9634	10.47094	-.182	268	.856
	female	188	69.2128	10.28458			
D1	male	79	21.1392	8.61980	.981	260	.327
	female	183	20.0492	8.08898			
D2	male	82	27.8780	7.13605	-.915	268	.361
	female	188	28.6755	6.33267			
D3	male	79	20.1013	4.20194	-.550	262	.583
	female	185	20.4054	4.07977			

Regarding Table 21, t-Test was administered to find whether there was a significant difference between gender variable and dimensions D1, D2, D3 together with item total score of ETARS. According to results of the Table 21, no significant difference was found to be between gender variable and the item total score of ETARS p= .856 (p>.05), D1 sub-dimension p= .327 (p>.05), D2 sub-dimension p=.361 (p>.05), D3 sub-dimension p= .583 (p>.05). This situation shows that, in Turkey, EFL teachers' attitudes towards the current recruitment system did not differ according to gender variable.

Table 22.

Scores Related to the Effect of Graduated Department Variable on ETARS

Dimensions	Group	N	Mean	Std. Deviation	t	Sd	Sig.
Total	ELT	218	68.9174	10.13333	-.661	269	.509
	others	53	69.9623	11.05226			
D1	ELT	214	20.3879	8.04821	.063	261	.950
	others	49	20.3061	9.09946			
D2	ELT	218	28.3670	6.58466	-.310	269	.757
	others	53	28.6792	6.57727			
D3	ELT	212	20.2547	4.26172	-.433	263	.665
	others	53	20.5283	3.42841			

As presented in Table 22, t-Test was utilized to investigate whether there was significant difference between graduated department variable and dimensions D1, D2, D3 together with item total score of ETARS. These results show that there was no significant difference between graduated department variable and the item total score of ETARS p=.509 (p>.05), D1 sub-dimension p=.950 (p>.05), D2 sub-dimension p=.757 (>.05), D3 sub-dimension p=.665 (p>.05). According to these findings, EFL teachers' attitudes towards the way they are recruited did not differ according to graduated department variable in Turkey.

Table 23.

Scores related to the Effect of Grade Taught Variable on ETARS

Dimensions	Group	N	Mean	Std. Deviation		Sum of Squares	df	Mean Square	F	Sig.
Total	primary	53	70.377	9.66601	Between Groups	162.002	2	81.001		
	lower secondary	124	69.266	11.31594	Within Groups	28518.979	268	106.414	.761	.468
	upper secondary	94	68.223	9.22264	Total	28680.982	270			
D1	primary	50	20.020	7.66516	Between Groups	110.370	2	55.185		
	lower secondary	120	21.066	8.51764	Within Groups	17661.113	260	67.927	.812	.445
	upper secondary	93	19.6667	8.17517	Total	17771.483	262			
D2	primary	53	29.377	5.82510	Between Groups	65.748	2	32.874		
	lower secondary	124	28.0484	6.84421	Within Groups	11596.599	268	43.271	.760	.469
	upper secondary	94	28.3936	6.61575	Total	11662.347	270			
D3	primary	52	21.1154	3.07841	Between Groups	43.429	2	21.714		
	lower secondary	119	20.1849	4.36272	Within Groups	4403.198	262	16.806	1.292	.276
	upper secondary	94	20.0213	4.24259	Total	4446.626	264			

Considering Table 23, one-way analysis of variance (ANOVA) was applied for the relationship between grade taught variable and dimensions D1, D2, D3 together with item total score of ETARS. According to the results of one-way analysis of variance (ANOVA), no significant difference was noted to be between grade taught variable and the item total score of ETARS p=.468 (F= .761 p>.05), D1 sub-dimension p=.445 (F= .812; p>.05), D2 sub-dimension p=.469 (F=.760 p>.05), D3 sub-dimension p=.276 (F=1.292; p>.05). When Table 23 was investigated, Turkish EFL teachers' attitudes towards the current TAE did not differ according to grade taught variable

Table 24.

Scores related to the Effect of Reason for Choosing Department Variable on ETARS

Dimension	Group	N	Mean	Std. Deviation		Sum of Squares	df	Mean Square	F	Sig.
Total	love	172	69.9651	10.27169	Between Groups	374.566	3	124.855	1.178	.319
	guided	43	67.5349	9.09303	Within Groups	28306.415	267	106.017		
	English teacher	32	68.4688	12.49254	Total	28680.982	270			
	other	24	66.7917	9.21709						
D1	love	164	20.3963	9.02391	Between Groups	3.708	3	1.236	.018	.997
	guided	43	20.1628	6.91420	Within Groups	17767.775	259	68.601		
	English teacher	32	20.5938	6.74350	Total	17771.483	262			
	other	24	20.2917	6.84283						
D2	love	172	29.0814	6.57697	Between Groups	256.793	3	85.598	2.004	.114
	guided	43	26.8372	5.28218	Within Groups	11405.554	267	42.717		
	English teacher	32	28.3750	8.01511	Total	11662.347	270			
	other	24	26.6667	6.06248						
D3	love	169	20.3787	4.24948	Between Groups	27.244	3	9.081	.536	.658
	guided	42	20.7619	3.63459	Within Groups	4419.382	261	16.932		
	English teacher	30	19.6667	4.60385	Total	4446.626	264			
	other	24	19.8333	3.15769						

Table 24 shows that one-way analysis of variance (ANOVA) was applied for the relationship between reason for choosing department variable and dimensions D1, D2, D3 together with item total score of ETARS. According to the results of one-way analysis of variance (ANOVA), no significant difference was noted to be between grade taught variable and the item total score of ETARS p=.319 (F= 1.178 p>.05), D1 sub-dimension p=.997 (F= .018; p>.05), D2 sub-dimension p=.114 (F=2.004 p>.05), D3 sub-dimension p=.658 (F=.536; p>.05). This situation shows that, in Turkey, EFL teachers' attitudes towards the way they are recruited do not differ according to reason for choosing department variable

Table 25.

Scores related to the Effect of Desired Profession Variable on ETARS

Dimension	Group	N	Mean	Std. Deviation		Sum of Squares	df	Mean Square	F	Sig.
Total	teacher	146	69.8356	10.38037	Between Groups	356.411	2	178.206		
	academician	73	69.3425	10.87712	Within Groups	28324.570	268	105.689	1.686	.187
	other	52	66.8077	9.05547	Total	28680.982	270			
D1	teacher	144	20.7431	8,67014	Between Groups	45.567	2	22.784		
	academician	71	20.0282	8.41083	Within Groups	17725.916	260	68.177	.334	.716
	other	48	19.7708	6.56308	Total	17771.483	262			
D2	teacher	146	29.0411	6.55783	Between Groups	208.085	2	104.042		
	academician	73	28.4247	6.55091	Within Groups	11454.262	268	42.740	2.434	.090
	other	52	26.7115	6.46055	Total	11662.347	270			
D3	teacher	144	20.0417	4.51086	Between Groups	35.648	2	17.824		
	academician	69	20.9130	3.70905	Within Groups	4410.978	262	16.836	1.059	.348
	other	52	20.2500	3.33064	Total	4446.626	264			

According to Table 25, one-way analysis of variance (ANOVA) was applied for the relationship between desired profession variable and dimensions D1, D2, D3 together with item total score of ETARS. According to the results of one-way analysis of variance (ANOVA), no significant difference was noted to be between grade taught variable and the item total score of ETARS p=.187 (F= 1.686 p>.05), D1 sub-dimension p=.716 (F= .334; p>.05), D2 sub-dimension p=.090 (F=2.434 p>.05), D3 sub-dimension p=.348 (F=1.059; p>.05). According to these findings, whether or not in-service English language teachers perform the job they desired does not affect their attitudes towards the current TAE.

4.3 Discussion of the Findings

In this section, the discussion of the findings is provided with respect to research questions. Unlike the the finding section, the research question will be discussed according to the order they were asked to the participants.

4.3.1 Discussion on Competencies That a Qualified EFL Teacher Should Have

After examining the qualitative data gathered from education stakeholders, three categories occurred as a result of the perceptions of participants about the competencies that qualified teacher must have, namely *knowledge-based competencies, skill-based competencies,* and *attitude-based competencies*. Thus, it can be concluded that education stakeholders think that a language teacher should have the required competencies related to knowledge, skills, and attitudes. These findings are consistent with "General Competencies for Teaching Profession" that was published by the Ministry in 2017. In the previous years, in addition to the generic teacher competencies, the Ministry collaborated with various stakeholders in order to determine field/subject specific competencies for teaching Profession. However, the Ministry has terminated that tradition and created a single and comprehensive text, by adding subject matter knowledge and pedagogical content knowledge to the general competencies. Therefore, competencies have been updated and categorize under three domains; namely "professional knowledge", "professional skills", and "attitudes and values", including 11 competencies and 65 indicators (for more details, please see MEB, 2017b). Likely, Mohamed, Valcke and De Wever (2017) analyzed existing frameworks of eleven countries and stated that competencies were generally clustered across three main domains; "(1) knowledge: what teachers need to know, (2) skills: what teachers are able to do and (3) attitudes or behaviours: how teachers would put their knowledge and skills in practice." (p. 155).

4.3.1.1 Knowledge-related Competencies

Regarding knowledge-related competencies, according to education stakeholders, an effective language teacher should have knowledge of subject matter, pedagogical content knowledge, a repertoire of teaching techniques and approaches, knowledge of underlying theories of teaching, knowledge of learners' development (cognitive, linguistic, social,

emotional, and physical), knowledge of curriculum, and knowledge of legislation. From this point of view, considering teaching foreign language, it can be said that having the required knowledge of some issues is a must and is prerequisite for teaching. This confirms the study of Richards, Conway, Roskvist and Harvey (2013) who examined "the link between foreign language teachers' subject knowledge and their language teaching classroom practice" (p. 235). According to them, teachers' subject knowledge is sine qua non of an effective teacher, because what s/he knows has a direct impact on the classroom environment (Richards et al., 2013, p. 233). That is, a teacher with the required subject matter knowledge can present that knowledge so eloquently that it becomes easy for students grasp the subject (McNamara, 1991).

The participants of the current study considered pedagogical content knowledge as one of the competencies that an effective teacher should have. Considering language teacher competencies, Richards (2011) mentioned pedagogical content knowledge which presents a basis for language teaching. According to him, pedagogical content knowledge "could include course work in areas such as curriculum planning, assessment, reflective teaching, classroom management, teaching children, teaching the four skills, and so on." and those who have that knowledge "should consequentially be able to make better and more appropriate decisions about teaching and learning and to arrive at more appropriate solutions to problems than a teacher without such knowledge." (Richards, 2011 pp. 3-4).

According to education stakeholders who participated in this dissertation, English language teachers must have a good knowledge of underlying theories of teaching and a repertoire of teaching techniques and approaches in order to perform effective lessons. Larsen-Freeman and Anderson (2013) pay attention to the knowledge of method as it contributes to teachers' repertoire of teaching techniques and, in order to underline the importance of that repertoire, they cited Arends (2004) as "effective teachers who are more experienced and expert have a large, diverse repertoire of best practices…, which presumably helps them deal more effectively with the unique qualities and idiosyncrasies of their students." (Larsen-Freeman & Anderson, 2013 p. xii).

Finding of this dissertation showed that knowledge of legislation and curriculum is of paramount importance in order to be an effective teacher. Participants stated that teachers should be aware of the education law which gives information about teachers' professional role, and they explained that curriculum knowledge helps teachers how to achieve

educational objectives. According to them, there is a correlation between teachers' practice and their knowledge of legislation and curriculum. In addition, Pillay (2014) examined this link in order to reveal whether teachers' knowledge of legislation and education law has a significant influence on teachers' practice. The findings of the dissertation displayed that teachers' knowledge of legislation and education law constitutes confidence and self-assurance and power (p. 145). According to Shulman curriculum knowledge can be described "as the 'tools of the trade' of teachers." (as cited in Brant, 2006). So, it can be concluded that curriculum knowledge refers to teachers' knowledge of what and how to teach and, therefore, it is needed to achieve educational standards.

4.3.1.2 Skill-related Competencies

Participants of this study listed the skill-related competencies as followings: Speaking English well; being good at teaching; conducting enjoyable lessons; using technological tools; preparing and using appropriate assessment and evaluation tools; understanding native and/or non-native conversation; managing students and learning environment; preparing, using, and evaluating materials; and demonstrating a coherent writing. Although language proficiency involves a lot of components (such as listening, grammar etc.), the most underlined one was speaking. Thus, it can be concluded that the speaking skill is one of the most important features that an effective English language teacher should have. The importance of oral proficiency was examined by some researchers (for example, Akcan, Aydın, Karaman, Seferoğlu, Korkmazgil, Özbilgin & Selvi, 2017; Chambless, 2012; Faez & Karas, 2017; Freeman, 2017; Richards, 2017a, 2017b; Richards, et al., 2013; Tsang, 2017; Van Canh & Renandya, 2017).

In their study, Richards et al. (2013) noted that teachers who do not have a desired TL proficiency level "will only be able to provide beginner-level students with a limited introduction to the foreign language" (p. 244). Regarding the viewpoint of Richards et al. (2013), a student can be proficient in speaking unless s/he is instructed by a teacher who is proficient speaker. The correlation between having a good speaking level and using it was also mentioned by some participants of this study (for example, ST8: …*if you are not good at speaking TL, you cannot use it in your classes*., and PM4: … *especially, a good pronunciation and a fluent speaking are sine qua none of English language teachers. The teacher cannot improve his/her learners' speaking and listening skills unless s/he*

117

possesses a fluent and accurate speaking ability.). According to them, since the speaking ability of student teachers is not assessed, they do not pay attention to it. As a result, they become language teachers who lack of speaking and, therefore, their students do not receive enough input. The aforementioned excerpts remind the Sullivan's (2011) assertion that "all can agree that the French teacher who cannot speak French will not be a successful teacher of French" (p. 241 as cited in Taner, 2017, p. 168). Likely, Chambless (2012) states that language proficiency is sine qua non of English language teachers and warns that English teachers should speak TL well if they want to teach it. She further points out that those with low level of language proficiency would fail to create an input-reach environment for their students. According to Akcan et al. (2017), teachers' language proficiency is among the fundamental qualifications that must be considered when recruiting teachers in higher education settings.

As reported by the education stakeholders, the second most emphasized skill-related competency is being good at teaching. In other saying, a teacher's having a good knowledge of underlying theories of teaching and learning does not guarantee that she or he will perform well in classroom or demonstrate enjoyable lessons. Especially, for novice teachers, the initial challenge is to acquire the required basic classroom skills for presenting and navigating their lessons. From this point of view, Richards (2011) voiced that teaching is an act of performance, and teachers require "techniques and routines, including routines and procedures for such things as opening the lesson, introducing and explaining tasks, setting up learning arrangements, checking students' understanding, guiding student practice, making transitions from one task to another, and ending the lesson." (p. 4). Seferoğlu (2006) conducted a study with 176 senior year students at the Middle East Technical University in Ankara, Turkey in order to investigate teacher students' reflections on the methodology and practice components of a pre-service English teacher training program in Turkey. Her study showed that the student teachers complained about the lack of close relationship between the course materials and practical application in real classrooms, and that they should be provided more opportunities for micro-teaching and practice teaching. It seems that knowledge itself is not enough until it turns into practice.

It was noted that the participants focused on the technology literacy. According to them, it contributes our instruction, so makes a better teaching and learning environment. As

Davies (2011) stated "[t]echnology literate people know what the technology is capable of, they are able to use the technology proficiently, and they make intelligent decisions about which technology to use and when to use it" (p. 47). According to U.S. Department of Education (2001), there are some benefits of integrating instructional technology into the teaching and learning environments, such as "(1) that learning is enhanced through the use of technology and (2) that students need to develop technology skills in order to be productive members of society" (as cited in Davies, 2011 p. 46). But not all education practitioners and school administrators value technology. For example, Bauer and Kenton (2005) do not care about integrating technology into classroom practices as much as theory suggests it should. They conducted a qualitative study with 30 "tech-savvy" teachers who were proficient with technology and who used computer technology in their instruction. However, those teachers were observed not to integrate technology on a consistent basis as both a teaching and learning tool due to two reasons; (1) their students lacked enough time at computers, (2) teachers required extra planning time for technology lessons. They also listed some other problems such as outdated hardware, lack of appropriate software, technical difficulties, and student skill levels. Their results displayed that schools fail to true technology integration.

4.3.1.3 Attitude-related Competencies

The last category consisted of competencies related to student teachers' attitudes. Regarding attitude-related competencies, the education stakeholders claimed that an effective English language teacher is a person who

- loves his/her job,
- considers learners' needs,
- demonstrates a fair attitude towards their students,
- treats students as individuals,
- is open to professional development,
- is aware of the latest ELT developments (as a researcher),
- has warmth and tolerance,
- has a good sense of humour,

- establishes an effective interaction, communication and cooperation with stakeholders,
- is democratic and humanist.

Thus, it can be concluded that a person needs to know the answer of 'why' in order to become an effective teacher in addition to being aware of 'what and how to teach'. That is, knowing something or being able to perform it is not enough for a person to be an effective English language teacher; instead, it requires positive attitudes towards teaching profession. These views are parallel with the findings of Bağçeci, Yıldırım, Kara and Keskinpalta (2015) who introduced the existence of a positive correlation between positive attitudes and an effective teacher. Like findings of Sakurai (2012) and Taner (2017), the results of this study indicated that among the qualities of teachers with positive attitudes are having required subject and field knowledge, being able to successfully transfer their knowledge into practice, and being open to professional development. Taner (2017) stated that 10% of her participant "talked about attitudes, and especially towards profession with the code *loves teaching*" (p. 172, emphasize in original). However, when the frequency of similar code (labelled as *Loves his/her job* in our study) was examined, the ratio was above 94 %. That may be because of the fact that they value only those who love teaching profession.

When the data obtained from education stakeholders who participated the present study examined, it was noted that they considered some codes such as *considers learners' needs, demonstrates a fair attitude towards their students*, and *treats students as individuals* as competency areas. These results draw parallel lines with Yalçın İncik and Akay (2015) who conducted a study to examine the opinion of student teachers (enrolled at the faculty of education and pedagogical formation program) on teaching profession competence. In their study, student teachers found teaching competencies inadequate and suggested some new competencies such as *tolerance, developing empathy, love of students, establishing a good communication with students*. Establishing positive human relations with stakeholders and having reasonable attitudes for students remain on the agenda of international competence framework proposals (e.g., Pantić & Wubbels, 2010) which also shared some other features frequently emphasized by the education stakeholders of the current study.

Another important code determined by participants is *being aware of the latest ELT developments (as a researcher)*. Thus, it can be reported that stakeholders pay attention to teacher research and consider an effective language teacher as a researcher, and they also recommend teachers to look for new developments in the field of English language teaching. This ratifies the study conducted by Yeşilçınar and Çakır (2018) who examine the views and practices of English language teachers on teacher research. Respondents considered English language teachers as researchers and they believed that teacher research was sine qua non of their professional development. Besides, this view can be discussed from the perspective of *Openness to Experience* one of the "Big Five" personality dimensions, which is also a category (labelled as *Openness to personal and professional development*) of the current study. The Big Five personality traits examines personality in five dimensions (Extraversion, Emotional Stability, Agreeableness, Conscientiousness, and Openness to Experience). 'Openness to Experience' is described as the extent to which a person's actively looks for and welcomes different experiences (Pervin, 2002) and may promote the use of effective learning strategies which, in turn, influences academic success (Educational Testing Service, 2012). People who are open to experiences define themselves as independent and original, and they are generally creative, liberal, have a wide area of interest, and like change (McCrae & Costa, 1987).

In terms of personal characteristics, participants listed many features, such as being patient, tolerant, honest, leader, democratic and humanist, and having warmth and a good sense of humour. The link between humour and teaching, learning, classroom environment, or students' subject matter was examined by some researchers (Beard & Wilson, 2013 p. 216; Berk, 2002 p. 10; Gurtler, 2002; Henry, 2000; Levine, 2006; Lovorn, 2008, 2009; Paterson, 2006; Posnick-Goodwin, 2009; Shiyab, 2008; Walker, 2008). The findings of these studies can be summarized as followings; using humour as a teaching tool may

- corroborate students' attention,
- improve their motivation,
- lead enhancements in their learning,
- create a more relaxed and comfortable classroom environment,
- help students bear positive attitudes toward learning process,
- encourage both academic and behavioral success,
- contribute to higher order thinking,

- contribute to subject knowledge, and
- increase creativity.

4.3.2 Discussion on the Perceptions of Stakeholders of ELTAE

The content analysis of the qualitative data of this research question, which examined how education stakeholders perceived the way student teachers are assessed and evaluated in order to be recruited a public school, revealed that none of the participants were happy with the current TAE model in Turkey. By asking this research question, it is also aimed at displaying the impact of the current TAE model through the perceptions of education stakeholders. Thus, it can be concluded that the fairness, selectiveness, inclusiveness and validity of the current TAE model are investigated by the help of stakeholders' perceptions. As the qualitative data show that the current TAE model is not appropriate for selecting and recruiting student teachers. Although all of the stakeholders arrived at a consensus on inappropriateness of the current model, the reason behind its impropriety differs from person to person. In this sense, how they perceive TAE and how TAE model influences the stakeholders will be discussed separately in parallel with sub-questions of the related research question (for example, perceptions of student teachers, perceptions of policymakers etc.).

4.3.2.1 Perceptions of Student Teachers

The content analysis of student teachers' perceptions of ELTAE revealed dissatisfaction of student teachers with the current TAE model in Turkey. They found the current recruitment model inadequate and criticized it for not including and gauging all required competencies. Consequently, this situation does not have a good impact on their life due to the nature of examinations or components used in KPSS as assessment and evaluation tools. Student teachers stated that they feel obliged to pay attention to only competencies that are assessed in KPSS in order to get higher scores since it is a high-stakes test. This means that student teachers join private courses and spend their free or afterschool time for studying only subjects that are assessed in the KPSS, such as ÖABT, general culture, general ability, and educational sciences and, therefore, ignore vital competencies and

delay important skills such as speaking, technology use and assessment literacy. We labelled the codes related to this situation as *priority of system requirements*, because although curriculum asks student teachers master many characteristics, they start to take into account only the features that KPSS includes due to its high-stakes nature. That is, they feel obliged to restrain the curriculum so as to have enough time for studying issues that the current TAE asks. This confirms findings of studies in other countries where high-stakes regimes were found to have influence on constraining the curriculum and narrowing students' educational experiences. For example, Polesel, Rice and Dulfer (2014) conducted a study to investigate the impact of NAPLAN (National Assessment Program – Literacy and Numeracy) on Australian schools and students. The NAPLAN is a national testing scheme, in which all Australian students in Years 3, 5, 7 and 9 should participate. The findings showed that high-stakes examinations can deform instructional practices, narrow the curriculum and students' educational experiences. In Greece, Tsagari (2011) carried out a study to explain the link between the intended impact of the FCE (First Certificate in English) exam administered by Cambridge ESOL, teachers' perceptions of the exam and their classroom practices. The Greek State officially recognizes the FCE certificate as a language qualification for the recruitment and promotion of employees who work not only in the public and private sectors but also in various educational fields (Tsagari, 2011, p. 433). The results of the study showed that although high-stakes tests are powerful in changing educational context, they might not be the sole agents for that change. In addition, she presents teachers, textbooks, parents and educational institutions as factors that have equal role in the occurrence of washback.

The emerging categories regarding the ELTAE-related perceptions of student teachers (*priority of system requirements, anxiety, skill delay,* and *shortcomings*) were interrelated. The student teachers of this study stated that being forced to focus on the priority of system requirements causes the occurrence of anxiety. In order to overcome the anxiety or lessen its influence, as they stated, they tried to postpone some skills with the aim of allocating all of their time to only issues of the current evaluation model. Although delaying skills seems to provide advantage because of saving time, student teachers regret of ignoring skills, because the ignorance of skills results in graduates who have some vital shortcomings. The literature supports these findings and informs us that if something is tested, then testees start to attach importance to it; if not, they may ignore it. Regarding critical language testing (CLT), being parallel with the perceptions of student teachers of the current study,

Shohamy (2017) states that the power of high-stakes tests causes both test takers and educational systems to alter their instructional behaviors. In this dissertation, student teachers mostly underlined the importance of the speaking skill which was also a delayed competence. This was confirmed by Tsagari (2011) whose participants used traditional methods while ignoring communicative methodologies as they found them incompatible with the principles underlying the exam. In her study, students asked their teachers provide them with 'repeated items' and if their teachers do not prepare them according to subjects come in the exam, they would get upset.

4.3.2.2 Perceptions of In-Service Teachers

As mentioned above, while visualizing the impact of TAE model, student teachers expressed that TAE model might cause them to graduate with some shortcomings. Therefore, during the interview with in-service English language teachers, the main focus was to reveal whether they confirm those shortcomings, and if yes, learn what they are. The interview data showed that in-service teachers recognized remonstrance of student teachers. According to their statements, the delayed competencies are *practice, assessment, speaking*, and *technology literacy*. It may be due to the memory-based nature of TAE model. That is, all components of the current teacher recruitment or evaluation model (KPSS) are designed to gauge only student teachers' memorization ability while excluding skill or attitude-based competencies.

In addition to the findings of this dissertation, previous research shows that there is a relationship between the type of washback (e.g. positive, negative) and the nature of the exam. In other words, for example, if the impact of a paper-pencil exam on individuals' speaking skill is examined, most likely negative washback will occur. For instance, Allen (2016) examined the consequential validity of IELTS, specifically targeting on washback upon learners' test preparation strategies. According to findings, "[t]he IELTS Test created positive washback on learners' language ability and test preparation strategies, specifically regarding productive skills, which learners in the study context had neglected in their previous language study." (p.1). It may be due to the fact that IELTS Test includes also productive skills. However, when Özmen (2011) examined the washback effects of a high-stakes test (Inter-university Foreign Language Examination) on candidate academics in Turkey, he found that that exam had harmful and negative washback effects. He attributed

these results to the content and style of the high-stakes test which consists of a multiple-choice test. One of his participants complained about the nature of the exam and stated "... I used to speak English better, but now I completely forgot how to do it because ILE [Inter-university Foreign Language Examination] gets me to study grammar, grammar and grammar." (p. 221). So, it can be concluded that if the individuals are evaluated through knowledge-based paper pencil examinations, they may neglect skill and attitude-based issues. As occurred in Tsagari (2011), teachers explained that "exam influenced students' attitudes towards teaching, that they were negative towards communicative activities and materials or activities that did not relate to the exam because students believed that such activities did not relate to the requirements of the exam." (436). In addition, Mahmoudi (2015) investigate the washback effect of a vital high-stakes test in Iran, the Iranian National University Entrance Exam (INUEE). According to findings, students were interested in language areas that the INUEE tests. The INUEE includes two essentials areas, namely vocabulary and grammar, so students allocated their time to memorize lists of isolated vocabulary items and formulaic grammar rule. The aforementioned studies justify the reason behind both why the student teachers delayed some skills and why in-service teachers have some shortcomings such as practice, assessment, speaking, and technology literacy.

4.3.2.3 Perceptions of Teacher Trainers

According to the perceptions of teacher trainers, the current TAE model fails to measure student teachers' strengths and weaknesses through multiple choice questions. Instead, they believed that it usually selects the student teachers who are good at memorizing knowledge. The biggest problem of the current model, as they stated, is its inability in providing information about student teacher's teaching ability. That is, it can be claimed that teacher trainers perceive the assessment and evaluation model inadequate in terms of differentiating among student teachers. Thus, it is fair to announce that teacher trainers find the current evaluation model inadequate as it does not represent the performance and attitude-related evaluation methods.

The results of the present study seem to be parallel with that of Tovar (2011) who asked foreign-language teachers and supervisors to determine the strategies and processes that inform teachers' professional development. In her dissertation, the participants were found

to consider classroom observation and students' performance reviews as the most popular evaluation strategies. This shows that the best evaluation strategies are those which provide information about the individual's performance, as stated by a teacher trainer of this study: "…*I mean a student teacher's performance must be observed in order to see whether s/he can interact with students, whether s/he can teach, whether s/he can speak target language fluently, whether s/he is competent in required subjects or not… etc*. (TT1)". Beside, according to literature, investigating teacher performance and competency is important for changing and improving educational practice (Medley, Coker & Soar, 1984). In order to understand school leaders and teachers' attitudes and experiences, Stiell, Willis, Culliney and Coldwell (2018) conducted a study on international teacher recruitment, using two data gathering strands. The first one was telephone interviews with 44 head teachers, principals, Human Resources (HR) representatives, and other school leaders in secondary schools across England whereas the second one was a survey of international teachers, applied to 13,436 teachers who trained in another country and obtained Qualified Teacher Status (QTS). It was observed that school leaders who had recruited internationally within the last three years either took into account "teachers holding QTS" or "factors such as teachers having good subject knowledge, their potential ability to facilitate student learning, being enthusiastic about their subject, and enjoying working with children." (Stiell et al., 2018, p. 10). Motivated from the aforementioned reasons, some teacher trainers suggested that the TAE model should be designed or revised accordingly, because it cannot eliminate student teachers who are unworthy of teaching job.

As mentioned in previous parts, a competent English language teacher should be open to both personal and professional development for being aware of the latest ELT developments. This means that keeping himself/herself up-to-date is a must for an effective English teacher, and that a successful TAE model of a country must contribute to the student teachers' research engagement. As stated by TT2, the current TAE model affects student teachers in a negative way because of its memory-based nature. It causes student teachers to ignore the habit of developing themselves. This gives rise to a tendency of using traditional techniques and method even when they become teachers. TT2 added that "… *we know that teachers, especially English language teachers, must be technologically literate to make lessons both productive and enjoyable. For having required technological knowledge and for being aware of the new developments in language teaching, an English language teacher needs to be open to personal and professional development*". Thus, it can

be concluded that there is a need for an effective TAE model since the current TAE model does not contribute to student teacher' personal and professional development. This finding supports some studies which did not find the current practices as the indicator of an effective teacher selection (Ballou, 1996; DeArmond, Gross & Goldhaber, 2008; Engel & Finch, 2015; Kersten, 2008; Papa & Baxter, 2008; Rockoff, 2004; Trimble, 2001).

4.3.2.4 Perception of Education Experts

Education experts complained about the lack of language proficiency of student teachers, specifically in speaking. They stated that being proficient in speaking English is sine qua non of an effective English language teacher. They added that although the interview has been introduced to assess student teachers' communication skill, it does not reach its goal because the interview process takes place in Turkish. Besides, they indicated that the lack of good rubrics and field experts in interview would thwart the success rate of selecting student teachers with required communication skills. Thus, based on the study findings, it can be said that integrating an oral examination is a must in order to train student teachers with fluent speaking. In this context, this study shares similar finding with Garcia Laborda, Litzler, Gonzalez Such, Bakieva and Otero de Juan (2012) who found Spanish PAU (University Entrance Examination) insufficient in describing their language use as it is based on written aspects of the language. They studied on the OPENPAU Project with the aim of enhancing student teachers' speaking skill. They believed that the inclusion of oral tasks could have a positive influence on Spanish student teachers' overall educational and professional expectations. According to them, the OPENPAU project would help student teachers improve their communicative skills successfully. They concluded that the new PAU would promote achievement of a higher grade of language knowledge both in speaking and written tasks. One criticism may be about the cost of an oral examination. However, recruiting an effective English language teacher is clearly worth the investment (Pillsbury, 2005). This explanation substantiates the statement of one participant (EE3) of the current study who uttered that "*an oral examination may require more time, effort and money; however, it worth developing and conducting such assessment and evaluation model for a quality education.*"

Some education experts drew attention to the literature-related questions in ÖABT and voiced how those questions militated in favor of graduates of English language and literature. They warned that those questions require very specific knowledge of literature and that they could be easily answered by graduates of English language and literature whereas ELT graduates found it too specific to answer. Thus, it can be concluded that the current TAE model displays unintended consequences, which diminishes fairness and validity of the current evaluation model. Henning (1987) argued that the misused of tests or misinterpretation of score meaning are the most apparent threats to validity, which is called as construct-irrelevant variance (Messick, 1989). He expressed the construct-irrelevant variances as variables unrelated to the construct being measured, which occurs when factors irrelevant to the construct influence the test scores. According to Bachman (1990), the test taker's personality, background knowledge, test-taking strategies, as well as general intellectual or cognitive ability may be construct-irrelevant and, therefore, effort should to be made to minimize their influences. Thus, it can be interpreted that the current model is not construct relevant, and shows a tendency of favoring graduates of English language and literature programs because of the background knowledge of those student teachers.

4.3.2.5 Perception of Policymakers

The data analysis on the perceptions of policymakers showed that selecting effective English language student teachers via the current TAE model is not possible. They asserted that unless the results of attitude and, particularly, skill-based assessment are considered while evaluating student teachers, the TAE model will fail to distinguish qualified English language teachers from ineffective ones. In other words, the current TAE model may cause the existence of unskilled English language teachers. Thus, it can be inferred from these explanations that authorities should immediately identify what is targeted by testing at higher stakes, because the answer of this question

> would strongly determine what gets taught and learnt at lower stakes, with degrees of quality and intensity though. As long as their results are used for important decision making, high-stakes tests such as university entrance exams are widely perceived to have immense importance for the people and institutions involved, in the sense that they induce a rather equal curriculum through a renewed focus on what is measured. In practical term, high-stakes tests, powerful as they are, have the potential to shape curricular teaching and learning. (Kiany, Shayestefar, Samar & Akbari, 2013 p. 326)

128

From this point of view, establishing a fair, valid and inclusive evaluation system may be helpful not only for the process of selecting the right person for the teaching profession but also for redesigning curricular objectives and controlling education outcomes. It may also contribute to the quality of education by selecting teachers who have required skills and language proficiency. Although some countries, such as England and the United States, use open entry policies, they exerted pressure through subject matter and performance tests during certification or tenure stage after graduation (Ingvarson & Rowley, 2017). The basic finding in the Wang, Coleman, Coley and Phelps (2003) report exhibits that high-performing countries on TIMSS utilize high-stakes filters across all stages in the pipeline.

Policymakers in the current dissertation complained about the limitations of the current TAE model and pointed out a need for adding new components for distinguishing those who are good at teaching and speaking English, interacting and communicating with stakeholders from those who are not. Thus, it can be concluded that since an effective English language teacher needs to have required teaching competencies, subject matter knowledge and language proficiency, the current TAE model should be rearranged accordingly. That is to say, using more assessment tools and techniques will most likely result in having qualified teachers. This supports the case in Chinese Taipei where individuals experience severe set of assessments before gaining a tenured teaching position (Hsieh, Ling, Chao & Wang, 2013). They must both pass the Ministry of Education's Teacher Qualification Assessment and undergo two stages: "(a) written tests designed to assess education professional knowledge or subject matter knowledge and (b) a demonstration of their teaching competency and a personal interview, both of which are evaluated by a panel of teachers and a principal." (as cited in Ingvarson & Rowley, 2017, p. 4). This indicates that the current recruitment system does not contribute to the selection and recruitment of effective teachers, and that new assessment tools should be introduced immediately so that it could take into account the representative scores of student teachers' performance, personality traits, and attitudes as well as their knowledge of specific subjects.

4.3.3 Discussion on the Attitudes of In-Service EFL Teachers towards ELTAE

This part consists of the quantitative side of the study and data collected by an attitude scale developed by the researcher since literature does not provide any attitude instruments on teacher assessment and evaluation system in EFL contexts where the recruitment of teachers is actualized through centralized system. The findings of this research question showed that in-service English language teachers do not have positive attitudes towards the way they are assessed and evaluated (regardless of demographic variables such as gender, graduated department, grade taught, reason for choosing department, and desired profession) in order to be recruited to public schools. This shows that in-service English language teachers considered the current TAE model inadequate, unfair and invalid, and that they favored some other additional tests that could determine their skills and attitudes in order to reach the most appropriate candidates for the profession.

A similar result was found in the study carried out by Karaca (2011). He examined the attitudes of 258 student teachers (studying in different departments in 2009) towards KPSS via a 75-item questionnaire developed by the researcher. The research findings displayed that student teachers generally had negative attitudes towards KPSS and they all reached consensus on an urgent need of an alternative examinations. In this sense, they believed that KPSS failed to gauge the quality of teachers. In their qualitative study which consisted of 180 student teachers of Department of Secondary Science and Mathematics Education, Uyulgan and Akkuzu (2015) investigated their views of teacher selection and recruitment. The content analysis revealed that they had negative opinions regarding ÖABT, and that the exam caused them feel anxious and inadequate. In addition, they added that it is not fair to recruit teachers based on only the scores of knowledge-based exams.

Sezgin and Duran (2011) conducted a study to determine student teachers' perceptions about KPSS, whether this exam affected their lives, and whether it influenced student teachers' perceptions about teaching profession and the content of undergraduate courses. The content analysis reported that participants found KPSS as an inappropriate examination in terms of selecting qualified teachers. In addition, they stated that KPSS negatively affected their communication with surroundings, social relations, and joining to artistic, cultural and sporting activities. Their participants offered oral and written examinations instead of KPSS. Yilmaz and Yaşar (2016) studied with 87 student teachers to determine their perceptions about KPSS. While there was no significant difference

between gender and their perceptions, a significant difference was found to be among departments where they were studying.

The aforementioned results suggest a call for performance based scores. It is clear in literature that schools which are interested in quality education and teachers generally pay attention to performance indicators. For example, Jacob, Rockoff, Taylor, Lindy and Rosen (2016) listed four stages of evaluation, as they changed somewhat from year to year, that are considered by Washington DC Public Schools while recruiting teachers: general essays, subject-specific assessments, an interview, and a teaching audition. They examined the relationship between applicants' characteristics, hiring outcomes and teacher performance. Findings showed that both background characteristics and a demonstration lesson are informative about teacher effectiveness, and that "there exists considerable scope for improving teacher quality through the selection process" (p. 37).

4.3.4 Discussion on the Appropriate ELTAE

The qualitative data related to the last research question revealed that the education experts generally valued a teacher assessment and evaluation (TAE) model which gauges not only student teachers' knowledge but also their skills and attitudes. When the emerging codes were scrutinized, four categories were observed to outline the characteristics of an appropriate TAE model, namely *inclusive, distinguishable, equitable/fair*, and *accurate.* As can be noticed from the initial letters of the categories, the data suggested an IDEA to the authorities about the characteristics that an appropriate TAE model need to have while recruiting English language teachers to the public schools. To enlighten each of these categories, they are discussed here under separate headings.

4.3.4.1 Inclusive

Education stakeholders stated that in order to recruit the right person for the teaching profession, one of the basic feature is that the evaluation process should include competencies related to student teachers' knowledge, skills, and attitudes that an effective language teacher should have. So, it is useful to identify what makes a teacher effective. It has been a subject of importance for many educational scholars. While Pettis (1997)

believes that a professionally competent teacher, at first, needs to be principled, knowledgeable, skillful and open to develop, Walsh (2002) see discipline knowledge, pedagogical knowledge, knowledge of context, and personal knowledge as a must for an effective English language teacher. Some indicated that an effective language teacher should have good language proficiency, personality and language-teaching skills (Brown, 2001). Other considered personal teacher dispositions, such as caring, as a vital factor in teacher effectiveness (e.g., Averill, 2012; Noddings, 2013). From these examples, it is clear that an appropriate TAE model should include the whole aforementioned characteristics. That is, an appropriate TAE model must be informative about all teaching related issues, so different sources need to be applied to assess and evaluate student teachers with the aim of selecting affective English language teachers while eliminating ineffective ones. As one of them (PM2) emphasized, "…*imagine that a student teacher has answered the whole questions correctly. Does it mean that he will be a good teacher? Absolutely not, because it presents nothing about whether the student teacher can transfer his/her knowledge into practice…*"

This statement explains the reason why an appropriate TAE model must be inclusive and it shares common features in line with literature. According to Harris and Sass (2014), the link between teacher skills and productivity can be the best way of teacher performance. That's why, a lot of countries usually apply interview including sample lesson. For example, in Japan, private schools decide their recruitment criteria as in Turkey. But each Prefectural Board of Education is empowered to recruit their teachers for public schools through a recruitment test which generally includes both a paper test and an interview including a teaching demonstration (Ramírez Carpeño & Mekochi, 2015). In Spain, in order to be a public school teacher, a nationwide competitive examination, which has two phases, is required. Student teachers have a practical exam in the first part, but the second part measures their pedagogical knowledge by asking them to "present a teaching program based on the first exam" and to "do an oral presentation of a teaching unit" (ibid., p. 117). In addition, in Korea (see Kang & Hong, 2008 for more details), student teachers require passing the National Teacher Employment Test (NTET) which assesses their knowledge, skills, and dispositions via three phases and the last one consists of teaching performance and interview. In USA, Education Testing Service (2005) has also considered the importance of teacher's performance while establishing the Praxis Series of certification tests for beginning teachers. It includes three series which consists of (a) literacy and

numeracy; (b) subject-content tests; (c) Classroom Performance Assessments, and the last one gauges a teacher's in-class ability as a practicing teacher (Coniam & Falvey, 2007). Thus, it can be concluded that countries having good educational outcomes pay great attention to their recruitment model since ensuring that each classroom has a quality teacher starts with the use of efficient teacher evaluation model (Rivkin et al., 2005). According to Danielson (2007), classroom practices based teacher-quality measures "are more closely linked to professional standards and therefore assess teachers from expectations that are known and shared by the teaching community" (as cited in Santelices, Valencia, Gonzalez & Taut, 2017).

In Turkey, entering an English language teaching program requires successful scores of a centralized nationwide examination based on multiple-choice questions. There are no exit exams. Anyone graduated from these programs can be in line for the teaching job in public schools, and graduates of English Language and Literature, American Culture and Literature, Department of Translation and Interpreting, and Department of English Linguistics deserve to be teacher in public school as long as they have pedagogic formation certificate. The education stakeholders criticized the way pedagogic formation certificates are provided, and suggested authorities to make regulations for taking into account the teaching experience of student teachers (performance score of micro-teaching of of school experience and practicum). The perceptions of education stakeholders of this dissertation confirm Leshem and Bar-Hama (2008) who focused on how teachers should be assessed and provided three options as an answer: (1) by observable behavior and measurable components, (2) taking into account the practicum, or (3) by their formal knowledge.

In a review of prior research on teacher personality, personality was noted to be both the main predictor of teacher success (Polk, 2006) and student achievement (Zhang, 2007). One of the in-service teachers (T11) voiced that "…*personality, conscience, and psychology are important terms for teaching...*". In parallel with T11, recent labor economics related studies have underlined the importance of personality traits in in determining worker productivity (Borghans, ter Weel, & Weinberg, 2008b; Cunha, Heckman, Lochner & Masterov, 2006; Heckman, Stixrud & Urzua, 2006 as cited in Harris & Sass, 2014). In education, personality traits that effective teachers have can be labelled as caring, avid, motivated, fair, respectful, reflective, and humorous (Black & Howard-Jones, 2000; Delaney, 1954; National Association of Secondary School Principals

[NASSP], 1997; Peart & Campbell, 1999 as cited in Stronge and Hindman, 2006). Although measuring personality traits objectively is not easy (Borghans, Duckworth, Heckman & ter Weel, 2008), these traits may be noticed via direct observation (Harris & Sass, 2014). Thus, it can be suggested that there is a need to consider the student teachers' score of practicum and some courses (e.g., micro-teaching, material development and evaluation) when evaluating them.

4.3.4.2 Distinguishable

The term distinguishable here refers to the capacity of a TAE model to distinguish effective teachers and ineffective teachers. The methods for measuring teacher quality and the way to implement teacher performance assessment differ from context to context. For this reason, it is recommended for a nation to outline the characteristics that an effective language teacher must have so that it could develop related approaches for selecting effective teachers while eliminating ineffective ones. In other words, the Ministry needs to define the ideal student teacher that is aligned to the country/education's vision, need, culture, and performance expectations, because it is important to identify and recruit student teachers who display the qualities of effective teachers in order to maximize the impact of the assessment and evaluation process (Stronge & Hindman, 2006). Although there is not a unique definition of a qualified or effective teacher (Ingersoll, 2001, p. 42), as it is mentioned somewhere in this dissertation, an effective English language teacher needs to be good at competencies related to knowledge, skills, and attitudes. So, the current TAE model should be revised accordingly.

In *Good to Great* (2001), which provides many assumptions for how to achieve greatness, Jim Collins states that "when in doubt, don't hire—keep looking" (p. 54). However, the question to be considered, as Stronge and Hindman (2006) asked, is how we distinguish effective applicants from less or ineffective ones. Moreover, how do we ensure that we select a good teacher? They indicated that what is pursued "in teacher selection is *predictive validity*—that is, the ability to use the information available about candidates to make hiring decisions that result in capable and committed teachers" (p. 19 emphasize in original). On this basis, they introduced Teacher Quality Index (TQI) for teacher selection, which consists of research-based and field-tested interview protocol. They voiced that the

TQI interview protocol concentrates on gathering information from qualified applicants on their abilities, skills, and experiences in an initial screening interview and, later, in a more in-depth building level interview (ibid., p. 5).

Research focusses on four information sources that are referenced by principles in order to distinguish candidates with desired teaching qualities from others: "(a) past experience, (b) job interview, (c) professional references, and (d) resume" (Grove, 2008; Harris, Rutledge, Ingle & Thompson, 2010; Kersten, 2008 as cited in Ellis, Skidmore & Combs, 2017, p. 453). In the OECD, Musset (2010, p. 46) identifies the requirements for getting teaching certifications and entering the professional field:

> To improve teacher quality, policy makers can also allow both the rise of teacher certification requirements and the accreditation for teacher education programmes. In half of the OECD countries, having followed a teacher education programme is not enough to obtain a teaching certification. In France, Germany, Greece, Italy, Japan, Korea and Spain, aspiring teachers have to pass a competitive exam to be certified and enter the profession. The examination can be used to obtain a teaching license but also to obtain tenure in a public school. In other countries, they also have to complete a period of probation before getting their certification; in Spain and Italy this period is of one year.

Considering how to measure teaching quality, literature lists three main lines of research:

> (a) the study of observable teacher characteristics that may impact student performance (Hanushek and Rivkin 2006), often referred to as *inputs* (Goe et al. 2008);
>
> (b) the analysis of teaching quality, defined as professional practices based on professional standards and usually measured through class observation guidelines, focusing on *process* (Danielson 2007; Darling-Hammond 2012; Papay 2012); and
>
> (c) estimates of value added based on student performance on standardized tests (Millman 1997), also known as *outputs* and more traditionally associated to the concept of teacher effectiveness. (as cited in Santelices et al., 2017, p. 114 emphasize in original)

Very recent literature has exhibited a different approach to the discussion on how to measure teaching quality, putting emphasize on the amalgamation of different sources of information on teacher performance as well as it overall application in teacher performance evaluation (Cantrell & Kane, 2013; Darling-Hammond, 2012; Goe, Bell & Little, 2008). From this perspective, the aforementioned three research strands represent subcomponents of a single measure, instead of referring to different measures of teaching quality. Thus, teacher assessment and evaluation process should involve "observable teacher characteristics, professional practices, and the impact on student performance" (Santelices et al., 2017, p. 114) in order to distinguish effective teachers from in effective ones. Also,

135

Tucker and Stronge (2005) underlined the importance of observation, indicating classroom observations as the primary instruments utilized for evaluating teachers.

4.3.4.3 Equitable/Fair

Education stakeholders lay emphasis on crucial role of teacher assessment and evaluation in education, and stated that authorities should always keep in mind whether the assessment and evaluation practices reflect and promote fairness. Kunnan (2005, p. 235) presents two main principles with a set of sub-principles for test fairness; (1) *The Principle of Justice* (treating every person with equal respect), (2) *The Principle of Beneficence* (not being harmful or detrimental to society). Fairness in testing has various definitions, so a universally adopted definition is not available in the literature (American Educational Research Association, American Psychological Association & National Council on Measurement in Education, 2014, p. 49). However, its definition in the current study is in compliance with *ETS Standards for Quality and Fairness* which defines the definition of fairness as "the extent to which the inferences made on the basis of test scores are valid for different groups of test takers." (Educational Testing Service, 2014, p. 19). According to Messick (1989), construct under-representation and construct-irrelevant variance are two major threats to validity and they need to be examined in order to improve the test quality and thus promote positive washback. The former occurs when construct has some missing points, whereas the latter happens when components that are irrelevant to the construct of focus influence test score variance. These two factors both diminish test validity and affect the fairness of the test if groups are affected differentially. Thus, to establish a fair evaluation requires preventing those two factors and, specifically, preventing certain groups of test takers being more affected by a construct-irrelevant source of variance than other groups and to avoid unintended constructs (construct-irrelevant sources), because they may cause incorrect inferences about those who take tests and, therefore, decrease validity. This information confirms ST19:

> … *We are studying English language and literature, so we are successful in doing questions both related to literature and those related to language proficiency since we are always asked to read and write essays during the four-year education. Last but not least, we have the whip hand owing to literature questions.*

It shows that the current TAE model militates in favor of student teachers of English language and literature due to their background, rendering construct-irrelevant variance. In their study, *Construct-irrelevant factors influencing young English as a foreign language (EFL) learners' perceptions of test task difficulty*, Cho and So (2014) examined the interaction of primary students learning English with test materials and their response to English language test questions. They found that the complexity of language used in test questions may influence the way they demonstrate their skills on the constructs targeted in the test questions, and that perception of test difficulty may be affected by memory load and ambiguity in task descriptions.

Another finding of the study was that the components of the current TAE model (e.g., ÖABT, Interview) do not represent the construct in terms of language proficiency. Participants stated that ÖABT tries to measure language proficiency of student teachers through twenty-five questions in multiple choice formats. However, in order to ensure language proficiency, they are required to demonstrate their speaking skills as well as other skills. They added that although interview has been introduced to measure their communication skill, it does not serve its purpose as it measures English language teachers in Turkish, causing the construct under-represented. This is also mentioned by Chen, Li and Chason (2018), who stated the possible occurrence of construct under-representation when the goals of assessment are deficiently considered. According to them, low level items of cognitive function require only rote memorization in order to recall isolated facts, but they may not reflect the knowledge to represent the performance for real-world situations.

Alavi and Masjedlou (2017) conducted a study on the validity of IELTS Academic Writing Task One (IAWTO). They found IAWTO input to be decayed and inadequate, showing the construct underrepresented, i.e., reducing and narrowing the construct. They explained the reason behind the construct under-represented as followings:

> To tease apart what is at stake here, taking account of the reality that IAWTO limits the test takers to just factual content of the input diagram and does not allow any speculated explanations- and personal beliefs of the test takers- outside the given data (IELTS Handbook, 2007) sounds to us to construct under-represent. The case in attention runs diametrically counter to Harsch and Rupp's (2011) stance, holding the view that for measuring the full writing abilities of the test takers, the writing task must be open rather than limited. (Alavi & Masjedlou, 2017, p. 1098)

From these explanations, it can be concluded that validity is sine qua non of a high-stakes test. Therefore, a well prepared test should be far away from treats to validity, namely construct-irrelevant variance and construct under-representation, in order to prevent it from having unintended, negative, or no consequence (Green, 1998).

4.3.4.4 Accurate

The category *accurate* has two sides. The first one refers to accurate scoring (using the right assessment tool for the related competency). They criticized the inappropriateness of the current TAE model as it only depends on paper-pencil examination, because answering even all questions in the current model does not guarantee the selection of effective teachers. Each scoring should represent the related construct appropriately. That is, to what extent does a student teacher's score on ÖABT represent the "real or appropriate" measure of his/her English ability since multiple-choice questions may not be give information about individual's classroom performance? The second side occurs naturally as a result of the first side, and it is about the requirement of using multiple assessment and evaluation sources so that it could provide information about the student teachers' knowledge, skills and attitudes since relying on only multiple choice paper-pencil examinations does not visualize the student teachers' teaching ability. This was also emphasized by James Popham, Berliner, Kingston, Fuhrman, Ladd, Charbonneau and Chatterji (2014) who stated that scored obtained from standardized tests are inappropriate for evaluating teacher quality. According to Parsi and Darling-Hammond (2015, p. 5), current tests are not supportive, because

> [T]hey do not reflect and embody the skills students need, they cannot inform instruction that aims to fulfill these purposes. Additionally, multiple-choice tests do not offer any information about students' thinking or reasoning, their misconceptions, or ability to express their ideas. Thus they provide little information to help teachers shape instruction. Last but not least, since these assessments are primarily summative, they cannot inform continuous instructional improvement, a goal facilitated by a rich array of assessments within a comprehensive state assessment system.

These explanations prove how some policymakers (for example, PM4) were right in warning that "*Unless the system terminates the measurement of teaching quality through questions in multiple-choice format, it becomes impossible to equip each classroom with*

an effective language teacher." This may be the reason why school districts in some countries (such as, Australia, Hong Kong, and Singapore) have started to extend the use of performance assessments (Darling-Hammond & Wentworth, 2010). Mostly, the amalgamation of standardized tests and classroom-embedded performance assessments occurs in international context; for example, in Queensland and Australia, students in grades 3, 5, 7, and 9 are measured by national tests of numeracy and literacy occur while performance tasks are tested in other grades (Parsi & Darling-Hammond, 2015).

One of the educational experts in this study (EE6) expressed his dissatisfaction of the current model by uttering as fallowing: "*Without taking into account the score of student teachers' skills and attitudes when evaluating them, it seems impossible to enhance education and encounter each student with a qualified language teacher.*" This is, actually, a call for change or at least a revision for the current TAE model. Correspondingly, Kılıçkaya and Krajka (2013) examined the selection and recruitment process in Turkey and Poland. They compared teacher recruitment process in Turkey to that in Poland. They concluded that (ibid., p. 100-1)

> [W]hen what KPSS covers and what teacher candidates are expected to perform are concerned, especially language teachers in Turkey seem to have at disadvantage in that the exam does not differentiate those who have high levels of proficiency in English from those who do not. In other words, candidates' knowledge of English is in no way taken into account. As regards the recruitment process, as indicated earlier, a language teacher candidate's overall score on KPSS is what counts in the recruitment process in Turkey....
>
> It is natural that an exam is conducted when there are more applicants than vacancies in any profession. However, these exams should test what teacher candidates are expected to perform in class. In other words, we cannot talk about the reliability or validity of these tests such as KPSS in Turkey, which do not test candidates' subject knowledge.

Very recently, interview has been introduced as a new component to the current TAE model, with the aim of measuring student teachers' communication skills. However, the qualitative data analysis displayed that interview has negative and unintended consequences due to accuracy concerns. Education stakeholders focused on two vital issues that the interview process does not have, which, in turns, decreases validity. The first one is that those who give scores to student teachers in the interview process are not experts. The second issue is the absence of a detailed, valid and fair scoring rubric. Similarly, Strong and Hindman (2006) asserted that the goal for those who involved in the recruitment process should be selecting a highly effective teacher for every student in the

school. According to them, use of a scoring rubric is of importance for ensuring that the same criteria are used to evaluate responses. In addition, as cited in Stronge and Hindman (2006, p. 25), using rubrics "is likely to reduce errors in the interview process (Pulakos, Schmitt, Whitney & Smith, 1996). Rating scales may reduce bias and enhance interviewers' consistency of judgments about candidates' responses (Campion, Palmer & Campion, 1997)." Thus, these data suggest that the authorities should either terminate interview or redesign the process in which experts and well developed rubrics are common. It seems that the second option will be useful since interviews have been found to be "vital to establishing an organizational match between the candidates' knowledge, skills, and abilities and the culture and needs of the school system (Eder & Harris, 1999)" (ibid., p. 28).

Some researchers (e.g., Sautelle, Bowles, Hattie & Arifin, 2015) presented various sources that could be used as indicators of their performance and that might be considered during the evaluation process. According to them, "if selecting a candidate for a particular job, selectors predict a candidate's job performance (criterion) based on their previous experience and university scores (cues), as such cues are probabilistically related to their performance, the desired criteria." (ibid., p. 57). In a nutshell, literature displays evidence that using multiple scores results in the recruitment of more highly qualified teachers (Balter & Duncombe, 2008).

CHAPTER V

A SUGGESTED MODEL

The IDEA is developed to serve as a guide for the recruitment of English language teachers in Turkey and in EFL contexts where there is a centralized teacher recruitment system. It aims to bring a fair, reliable and valid teacher assessment and evaluation (TAE) model for English language teachers. The IDEA intends to take into account the knowledge, skills, and attitudes of student teachers through multiple sources of information with the aim of equipping each classroom with an effective language teacher.

5.1 The Need for IDEA

As discussed in previous chapters, teacher assessment and evaluation of a nation is determinative of the quality of its teachers and education. Considering the status of English in public schools in Turkey, education stakeholders were requested to reflect on the link between the way English language teachers are evaluated and their personal and teaching quality. The perceptions of education stakeholders and the attitudes of English language teachers displayed the impact of the current TAE model on student teachers, in-service teachers and education. That is to say, the quantitative and qualitative data showed how the current TAE affects and shapes the current situation of English language teachers. First, due to structure of the current TAE, student teachers were noted to feel obliged to deal with the neglected competencies only after all examinations related to current model PASS; that is, they delay some skills until becoming teachers. Thus, according to them, they graduate with some shortcomings due to the skill-delay process as they have to study for an exam based on memorization.

Second, the in-service English language teachers were interviewed in order to learn (1) whether the shortcomings that student teachers mentioned came true, and (2) their perceptions of the current recruitment model. They stated that both time and opportunities stayed at PAST in order to compensate these shortcomings. According to their statements, all these competencies should have been covered in undergraduate period, because once student teachers let these competencies get out of control, it seems to be too difficult to manage them after graduation.

Finally, other stakeholders (teacher trainers, education experts, and policymakers) contributed to the views of student teachers and in-service teachers in claiming that the current TAE model prevents teachers from becoming the CEO of their profession. They voiced that it fails to distinguish effective teachers from in effective ones. Their dissatisfaction and distrust of the current TAE model has doubled, especially, with the introduction of interview. Unless these deficiencies with TAE model are overcome, it seems to be impossible to equip each classroom with a qualified/effective teacher, which, in turns, relieves the education of the country. From that point of view, developing a valid and fair TAE model becomes a must. PASS, PAST, and CEO acronyms emerged from the initial letters of related features (please see Figure 5).

5.2 Referenced Sources for IDEA

Various sources have been applied while developing the IDEA. Some of them are as followings:

1- perceptions of education stakeholders
2- documents
3- literature
4- teacher selection and recruitment of different countries

    a- countries that perform well in PISA in reading
    b- countries that are on average in PISA in reading
    c- countries that are below average in PISA in reading
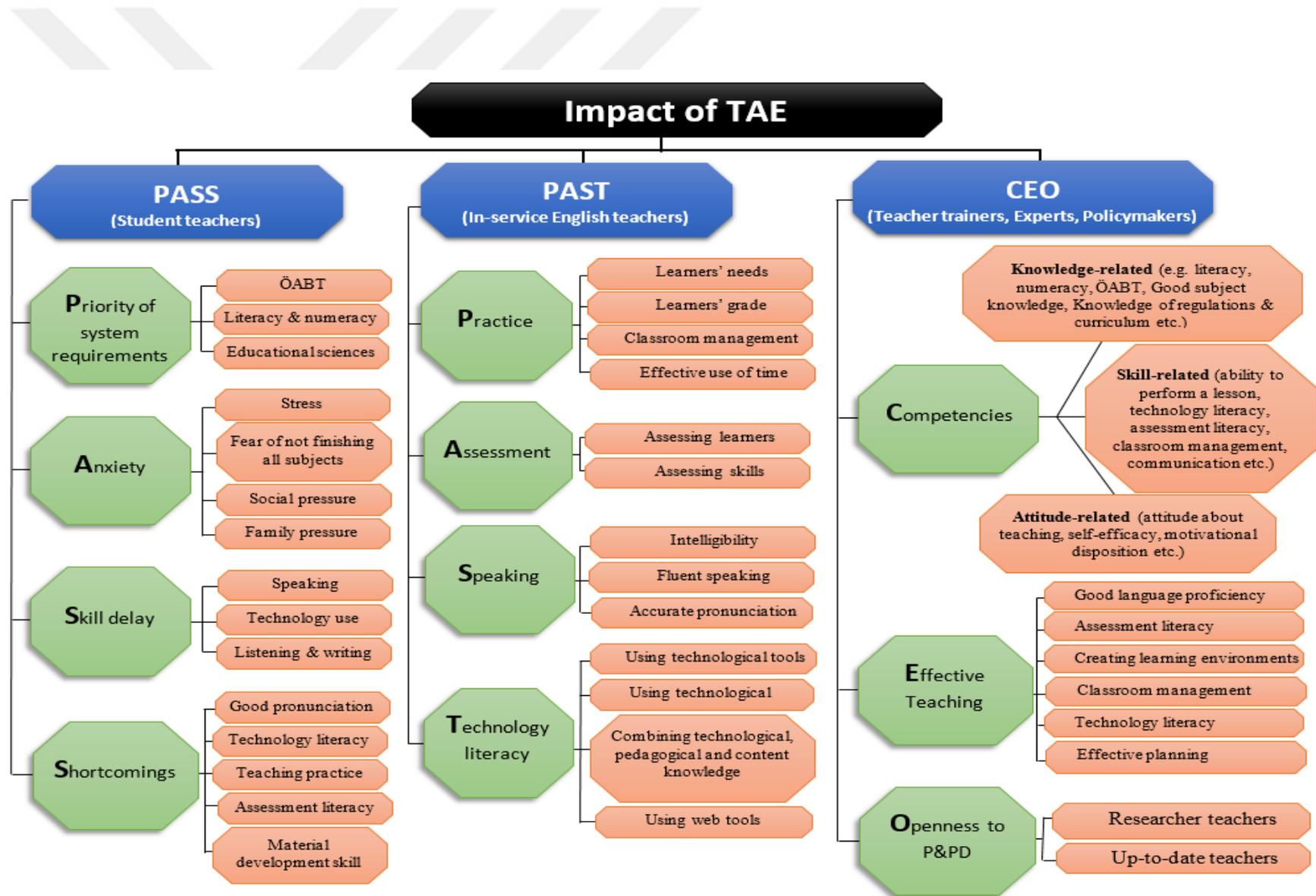
Figure 5. The impact of TAE

5.3 General Aims

The IDEA is in line with the philosophy of the consequential validity and teacher quality in that it aims at developing a construct represented and relevant TAE model for selecting and recruiting language teachers with required knowledge, skills and attitudes. It also takes into account stakeholders' perceptions, because in the case of "adopting a lens of consequential validity, it becomes important to understand public perceptions of these issues" (Hitchcock et al., p. 5). In addition, it aims to create a collaboration and cooperation between the MoNE and universities, by inviting both parties into evaluation process. In this sense, individual's teaching ability and personality traits are also assessed and these scores are considered for evaluation process. By sharing responsibility, it is targeted not only to ease the workload of the ministry but to contribute to the selection of language teachers who are good at speaking and teaching English well.

5.4 Characteristic of the IDEA

According the data, the suggested TAE model should include four features so that it can eliminate in effective student teachers. These attributes are as followings (please see Figure 6 for more details):

5.4.1 Inclusive

The IDEA includes competencies that qualified English language teachers must have. Not only does it have competencies related to knowledge, but also it involves both skills and attitude-related competencies. It aims to provide good examples of using various sources and approaches for assessing and evaluating student teachers so that it can select the right person for the teaching profession. It involves scores which represent student teacher's ability of preparing material, demonstrating teaching ability, and speaking fluently. In addition, it measures their subject knowledge as well as the knowledge of curriculum and legislation.
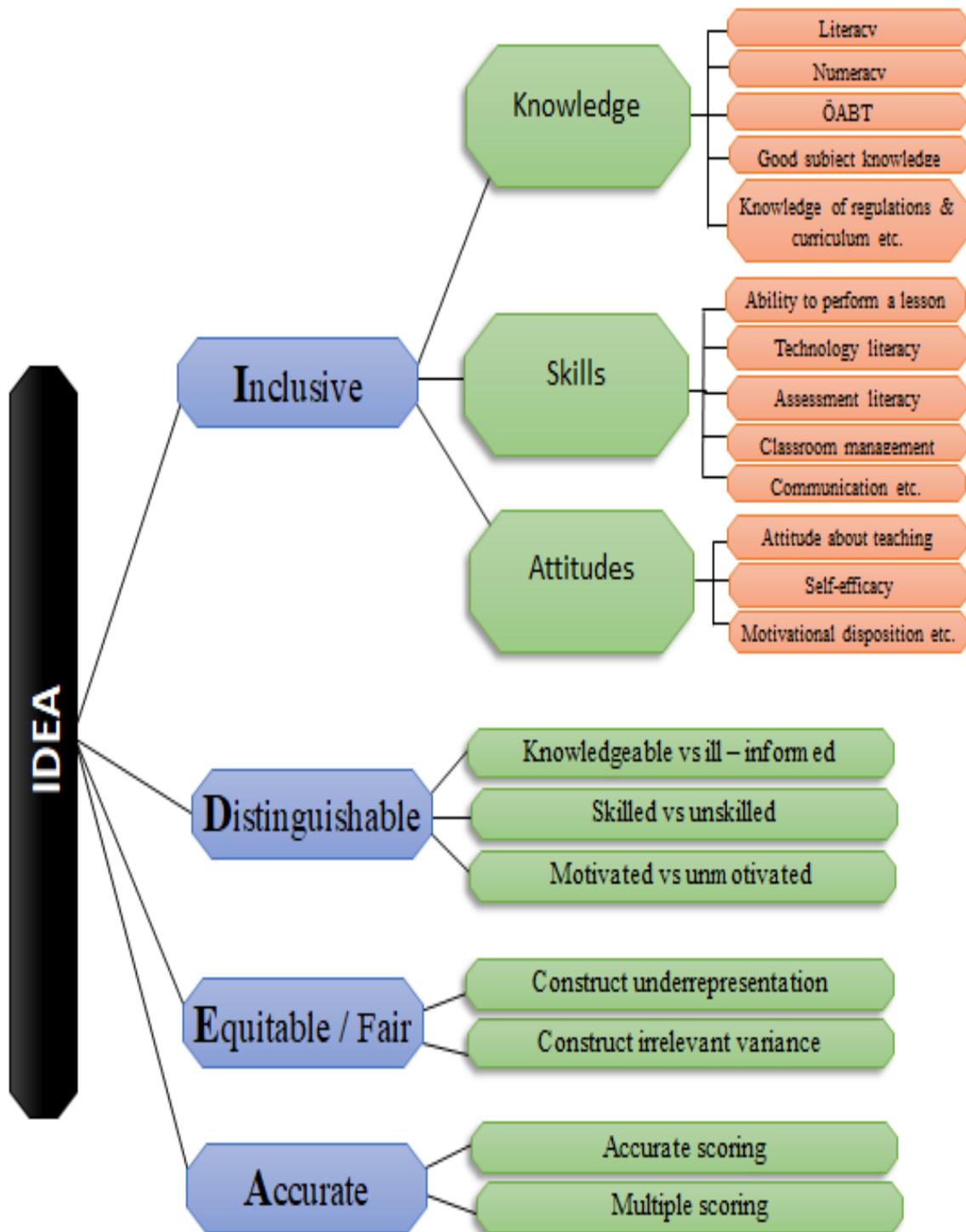
144

Figure 6. The suggested TAE model: IDEA

5.4.2 Distinguishable

The main limitation of the current TAE model is its insufficiency in selecting student teachers who are good speakers of target language, who are good practitioners, and who are highly motivated to teach foreign language to students. In this sense, the IDEA has been designed to distinguish effective English language teachers from ineffective ones. By using the aforementioned sources and approaches when assessing student teachers, it aims to distinguish knowledgeable, skilled and motivated student teachers from ill-informed, unskilled and unmotivated individuals respectively.

5.4.3 Equitable/Fair

In order to ensure fairness, the principle of being free from bias has been kept in mind while developing the suggested model. That is, the IDEA has been designed in a way that it can represent its constructs to be measured and hinder construct-irrelevant variance so that it could treat every person with equal respect in order to ensure equity and fairness in testing. In other saying, a great importance has been attached to the development process of IDEA for not being affected from the test taker's personality, background knowledge, test-taking strategies, as well as general intellectual or cognitive ability, because they may cause incorrect inferences about those who take tests and, therefore, decrease validity

5.4.4 Accurate

As discussed in previous chapters, the current assessment and evaluation model requires memorizing lots of topics and subjects and asks student teachers respond multiple choice tests. Moreover, the lack of rubrics for evaluating the scores of student teachers in interview process causes a dissatisfaction and anxiety among student teachers. Thus, the model aims to reassure that the evaluation is done appropriately, by using the right assessment tools with the related competencies. It also uses various sources of evaluation for ensuring accurate scoring.

5.5 Objectives

The IDEA has some objectives related to student teachers' knowledge, skills and attitudes. In other words, it aims to select and recruit English language teachers who have the following competencies.

5.5.1 Knowledge-Related Objectives

The English language teachers will have a required knowledge of

1- literacy
2- numeracy
3- subject/content
4- curriculum
5- legislation
6- monitoring, assessment and evaluation
7- technology
8- child development
9- underlying theories of language teaching and learning

5.5.2 Skill-Related Objectives

The English language teachers will be able to

1- perform an efficient lesson
2- integrate and use technology appropriately
3- use the right assessment tools for their students
4- speak English fluently
5- write a coherent piece of paper
6- create learning environments
7- interact with students
8- manage the teaching and learning process
9- create a good relationship with stakeholders

5.5.3 Attitude-Related Objectives

The English language teachers will be able to

1- care needs of students
2- value every student as a human being and individual
3- have positive attitudes towards students
4- open to personal and professional knowledge
5- cooperate with stakeholders
6- love their students and profession

5.6 Methodology

This part includes information about (a) what characteristics of a student teacher should be measured while recruiting him or her to the public school, and (b) what kind of ways may be followed while assessing these features. Figure 7 visualizes the knowledge-, skill-, and attitude-related competencies and provides the possible ways to for assessing and evaluating them during the assessment and evaluations process, with the aim of choosing the best student teacher.

5.6.1 General Considerations

Teacher assessment and evaluation is a multidimensional process with a variety of factors to take into consideration. The characteristics of the IDEA display that taking into account the student teachers' teaching ability, their attitudes, and their personality traits is sine qua none of recruiting or assigning an effective English language teacher. Hence, two kinds of question come to the fore; (1) *what to assess* and (2) *how to assess* (see Figure 7). The first one is about the features (related to knowledge, skills and attitudes) that an effective English language teacher should have, while the second question is about the sources or ways that may be applied for the assessment and evaluation process. To summarize it, the suggested model utilizes various sources of information for the assessment and evaluation of English language teachers to suit the knowledge, skills and personality traits of qualified teacher. The IDEA also emphasizes the productive skills of language teaching, particularly

speaking, indicating that an effective English language teacher needs to be fluent in the TL. Besides, it advocates the use of multiple scores which cohere with the related construct. Thus, it can be concluded that the IDEA aims to provide each student with a qualified language teacher.

Knowledge
(e.g. literacy, numeracy, ÖABT, curriculum etc.)

Skills
(ability to perform a lesson, technology literacy, assessment literacy, classroom management, communication etc.)

Attitude
(attitude about teaching, self-efficacy, motivational disposition etc.)

What

Assessment

How

Paper-pencil exams

Practicum
(mentor & adviser)

TOEFL or IELTs equivalent exams

Scores of undergraduate cources
(e.g. material evaluation, micro teaching

Figure 7. Competencies Need to be Assessed and Possible Ways of Assessment

5.6.2 Recommendations for What to Assess and How to Assess

Competencies that a qualified teacher should have may differ from country to country. Therefore, the characteristics mentioned here are determined by the researchers, depending on the perceptions of education stakeholders, literature and document analysis, in accordance with the need of the country.

5.6.2.1 Recommendations for Knowledge

According to the Ministry, literacy and numeracy are two important features that each teacher must own. In addition to them, the IDEA suggests that a qualified English language teacher must have a deep knowledge of subject/content, has required knowledge of curriculum, legislation, child development, testing, and underlying theories of language teaching and learning. As those issues contain a mass of knowledge, for the practicality concerns, the suggested model recommends the use of paper-pencils exams for knowledge-related issues.

5.6.2.2 Recommendations for Skills

The most problematic area of the current TAE is its being lack of performance-based scores. Although many subjects may be considered under this title regarding a student teacher's performance, basically issues seem to be of importance, namely verbal ability, teaching ability, and personality. The model proposes TOEFL or IELTS equivalent exam that will be conducted by ÖSYM for student teacher's speaking ability. To make it practical, the time for the exam should be flexible. That is, student teachers should be given opportunity to apply TOEFL or IELTS equivalent exam anytime they want (from enrolling the program to graduation).

The IDEA suggests the use of multiple sources of scores for the identification of student teacher's teaching ability due to the subjectivity concerns in performance-based evaluation. That is, it advocates the use of the scores of practicum and some lessons (e.g., micro-teaching of school experience, material development and evaluation) instead of applying a one-time observation-based exam (for example, a sample lesson). Regarding practicum, both mentors and advisors' evaluation scores should be considered during evaluation process. Three problems may arouse;

1. scores of private universities (some economical concerns may interfere the process),
2. consistency of scores (for example, does the score 80 in A university refer to the same score in B university?),
3. pedagogic formation.

For the first and second options, there may be two solutions; the Ministry should either introduce an exit exam or take into account student teachers' entry scores when evaluating them. Pedagogic formation is an enigmatic issue in Turkey. It is beyond the scope of the current study to discuss whether pedagogic formation system for outfield graduates is good or not. However, the IDEA insists that outfield graduates should spend at least one semester with senior ELT students if the pedagogic formation programs continue. Experiencing some courses (e.g., micro-teaching of of school experience, material development and evaluation) and practicum process with senior ELT students will contribute to pedagogical knowledge of outfield graduates.

5.6.2.3 Recommendations for Attitudes

The attitudes of student teachers towards teaching, their motivation, and their personality traits determine whether they will be good teachers. Thus, it can be concluded that, an effective teacher is expected to love his/her profession, care the needs of learners, respect each student as an individual, and have a sense of humour. Although it is not easy to measure personality traits objectively, a direct observation may be informative. That is, in micro-teaching lessons and practicum process, whether the individual is prone to teaching profession can be evaluated. This case is also the fifth action (Enabling students to switch to other programs from teacher training programs) which is stated in Teacher Strategy Paper (MEB, 2017a).

5.7 Conclusion

Currently, teacher assessment and evaluation depends on knowledge-based paper-pencil exams. Some vital components (teaching ability, verbal ability etc.) are ignored due to its multiple-choice nature. Very recently, interview has been introduced with the aim of selecting candidates with good communication skills. However, it does not serve its purpose because it is not possible for interviewees to encounter field experts during the interview process and it also lacks a good rubric for scoring. The individuals who want to be English language teachers mainly study lots of subjects related to numeracy, literacy, ÖABT in multiple-choice format. In addition, their communication skills are assessed in interview through few questions in Turkish.

As the findings of the previous chapter indicate, the current TAE model does not guarantee the selection and recruitment of qualified English language teachers. At the same time, the current model fails to capture the student teacher's ability of teaching and speaking English. In light of this information, a new model which considers student teachers' both skills and attitudes as well as their knowledge will contribute to the quality of teachers, and; therefore, increase the quality of education in Turkey.

CHAPTER VI

CONCLUSIONS

The present study was conducted with the purpose of (a) examining perceptions and attitudes of education stakeholders about Turkey's teacher assessment and evaluation (TAE) model in terms of consequential validity, (b) developing an attitude scale for English language teachers so as to reach a wide population with the aim of generalizing the results, and (c) suggesting a more reliable, valid and fair TAE model for EFL contexts in which the assignment process is done centrally. The participants of the study were;

- student teachers (of ELT department and other programs such as English Language and Literature, American Culture and Literature, Department of Translation and Interpreting or Department of English Linguistics graduating),
- in-service English language teacher (graduates of both ELT and other departments such as English Language and Literature, American Culture and Literature, Department of Translation and Interpreting, or Department of English Linguistics graduating),
- teacher trainers (of ELT programs),
- educational experts, and
- policymakers.

Semi-structured interviews were applied to all participants in order to gather the qualitative data whereas the quantitative data were obtained from in-service English language teachers through the 23-item attitude scale developed by the researchers. This chapter begins with a brief summary of the findings. Then, it is followed by conclusion and recommendations for further research.

6.1 Summary of the Findings

The first research question aimed to display the competencies that stakeholders (student teachers, in-service teachers, teacher trainers, education experts, and policymakers) valued. The content analysis of the data suggested three categories (knowledge-based competencies, skill-based competencies, and attitude-based competencies) that an effective English language teacher should have. Thus, it was concluded that, according education stakeholders, a qualified language teacher should have the required competencies related to not only knowledge but also skills and attitudes.

The porpuse of the second research question was to reveal stakeholders' perceptions of English language teacher assessment and evaluation (ELTAE) in Turkey. In other words, the impact of ELTAE was examined. Education stakeholders voiced some negative effects and the common agreement was on the shortcomings that occurred due to the current TAE. Participants stated that since the current model uses only knowledge-based exams, it causes unintended consequences (for example, selecting unmotivated and unskilled candidates).

The third research question was about the quantitative part of the study and its results showed that in-service English language teachers had negative attitudes towards the current TAE regardless of demographic variables such as gender, graduated department, grade taught, reason for choosing department, and desired profession. This part confirmed the qualitative data and provided an opportunity for generalizing the results.

The last research question was the sum and essence of the present study, which was used to develop a valid and fair TAE model. It was observed that stakeholders (e.g. student teacher, in-service teacher, teacher trainer, education expert, and policymaker) had reached a consensus on the introduction of a new valid and fair teacher assessment and evaluation model. The data suggested that the new TAE model should be inclusive, distinguishable, equitable/fair, and accurate. So, the acronym IDEA emerged from the initial letters of four features that a good TAE model should have. Thus, the suggested model can be regarded as a good IDEA to the authority about how to develop a fair, valid and reliable evaluation process for recruiting English language teachers to the public schools.

6.2. Conclusion

In this dissertation, the consequential validity of the current teacher assessment and evaluation (TAE) model in Turkey was investigated through the perceptions of various stakeholders (such as, student teachers, in-service English language teachers, teacher trainers in ELT departments of different universities, education experts, and policymakers) with the aim of developing a fair, reliable and valid TAE model for English language teachers who want to be recruited to public schools. The findings of this research have indicated that the current TAE model has not been appreciated and valued in education society. In addition, since there is a positive relationship between the effectiveness of teacher evaluation (or teacher selection & recruitment) and teachers' quality and job satisfaction (e.g., Johnson & Birkeland, 2003; Liu & Johnson, 2006), the current model has been accused of not contributing to the process of bringing up quality English language teachers due to its inability to distinguish effective language teachers from ineffective ones. In this sense, it can be deduced that bringing effective teachers in classrooms is only possible with a well-established TAE model.

The quantitative and qualitative data showed that the current TAE does not contribute to English language teachers' personal and professional development. Indead, it was found to damage their personal and professional development. For example, student teachers ignore some vital competencies (e.g., speaking and teaching English well etc.) due to the nature of the current TAE and they wish to deal with the neglected competencies when they become teachers. However, after becoming an English teacher, they were found to be pessimist in terms of overcoming and managing these shortcomings. In addition to them, other stakeholders also indicated that the current model damages student teachers regarding their skill-related competencies and form a basis of the existence of ineffective English teachers. This led to the development of a fair, reliable and valid TAE model, the IDEA, serving as a guide for the recruitment (evaluation) of English language teachers in EFL contexts where teachers are assessed and evaluated via a centralized system. The aim of the IDEA is to equip each classroom with an effective language teacher, so it takes into account the knowledge, skills, and attitudes of student teachers through multiple sources of information.

According to the qualitative data, the stakeholders had negative perceptions towards the current TAE. However, it was not feasible by using qualitative data to reach wider

155

populations and generalize the results. Thus, with the help of the findings of the current study as well as literature, an attitude scale was developed by the researcher of the current study in order to generalize the results, because, to the best knowledge of the researchers, no attitude scales were available in literature regarding English teachers' attitudes towards teacher evaluation or recruitment model in EFL contexts where the teacher recruitment is actualized centrally. Thus, the researchers developed ETARS (English teachers' attitudes towards recruitment system) to reveal the attitudes of in-service English teachers the way they are assessed and evaluated. The creation of ETARS is of paramount importance, because using such instruments display "whether qualitative research findings held over large groups of respondents." (Hitchcock & Nastasi, 2011; Hitchcock et al., 2005, 2006; Nastasi, Hitchcock, Burkholder, Sarkar & Varjas, 2007 as cited in Hitchcock, et al, 2015 p.5).

The findings of the current dissertation presented an urgent need for an information rich TAE model. That is, the new TAE model should provide information about not only student teachers' knowledge but also their skills and attitudes. As Danielson (2011) states, evaluation systems should be significant for teachers themselves and informative about their own practice. Stakeholders of this study believed that it is not possible to equip each classroom with an effective English language teacher unless student teachers' teaching and verbal ability is measured and unless their personality traits are taken into account. Their statements suggested a four-characteristic TAE model which could select qualified teachers while eliminating ineffective candidates. These characteristic are

- *inclusive* (refers to the features related to student teachers' knowledge, skills of teaching profession, and their attitudes towards it),
- *distinguishable* (refers to its capacity of selecting quality teachers but eliminating ineffective teachers),
- *equitable/fair* (a test's unbiased decisions. It should be construct represented and construct relevant), and
- *accurate* (refers to its capacity of scoring accurately and using the right assessment tool for the related competencies).

As the initial letters of these categories speak, the present study suggested an IDEA to the authorities about how to select qualified applicants when the number of English language teachers outweighs the number of available places.

The findings also have some implications for policymakers. It has circumstantiated the reason why the current model has been believed to be invalid and unfair, and it has also suggested possible solutions for overcoming these problems. The education stakeholders argued the unintended and negative consequences of the current TAE model as the reason behind its invalidity and unfairness. Therefore, the present study has delved into these consequences, keeping in mind that, in the case of any negative consequences, policymakers, and test organizers must investigate and correct the source of the problem (Hitchcock et al., 2015). This study might ease the job of policymakers because of its scientific and pedagogic nature. In other saying, underlying problems related to validity of the current model were investigated and identified, and accordingly a new valid TAE model was developed in order to offer possible solutions to these problems. It is valid because the use and interpretation of a test is also a kind of validity: consequential validity. "In other words, when tests are used in decision-making processes the validity of the test is tied to its use and the real consequences it has for people." (Brewer, et al., 2015).

In a nutshell, based on the underlying principles of consequential validity, the IDEA has been developed in order to make the process of teacher assessment and evaluation more valid and fair, with the aim of enhancing the quality of teaching and learning. It is assumed that if the four pillars (inclusive, distinguishable, equitable/fair, and accurate) of the IDEA are utilized appropriately by the authorities during the process of recruiting language teachers, it will not become impossible for the country to provide classrooms in which each student is taught by an effective English language teacher.

6.3. Recommendations for Further Research

In this study, the perceptions of education stakeholders (e.g. student teachers, in-service teachers, teacher trainers, education experts, and policymakers) and the in-service English language teachers' attitudes towards the current teacher recruitment system were investigated. As it has been mentioned, the suggested model offers a more reliable, fair, scientific, valid, and pedagogic recruitment system for English teachers. The model can be put into use with the decision of policymakers (such as the Minister, ministry administrators, deputies in the parliament etc.). Having considered this situation, the researcher has tried to reach as many policymakers as possible in order to discover their attitude towards such recruitment system. A further study might be based on developing

the offered model by reaching out more policymakers, especially the ones in the parliament's National Education Commission because such comprehensive model needs to be analyzed and evaluated by the related members of the parliament.

Although a great effort was spent to develop a fair, reliable and valid TAE model and even though various stakeholders were given voice for creating a reasonable methods and ways of assessing and evaluating English teachers, it is believed that a further study needs to investigate whether the IDEA works. This would lead to small-scale piloting of the model with the government support. As a result, we would have the opportunity of applying this model all around the country and this would benefit all the aforementioned stakeholders.

REFERENCES

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago Public High Schools. *Journal of Labor Economics, 25*(1), 95–135.

Adnot, M., Dee, T., Katz, V., & Wyckoff, J. (2017). Teacher turnover, teacher quality, and student achievement in DCPS. *Educational Evaluation and Policy Analysis, 39*(1), 54-76.

Akcan, S., Aydin, B., Karaman, A. C., Seferoğlu, G., Korkmazgil, S., Özbilgin, A., & Selvi, A. F. (2017). Qualities and qualifications of EFL professionals: What do Intensive English Program Administrators Think? *TESOL Journal, 8*(3), 675-699.

Akpınar, K. D., & Çakıldere, B. 2013. Washback effects of high-stakes language tests of Turkey (KPDS and ÜDS) on productive and receptive skills of academic personnel. *Journal of Language and Linguistic Studies, 9*(2), 81-94.

Akunal, Z. (1992). Immersion programmes in Turkey: An evaluation by students and teachers. *System 20(4)*, 8-12.

Akyüz, Y. (2001). *Başlangıçtan 2001'e Türk Eğitim Tarihi*. Genişletilmiş sekizinci basım. İstanbul: Alfa.

Akyüz, Y. (2009). *Türk Eğitim Tarihi*. Ankara: Pegem.

Alavi, S. M., & Masjedlou, A. P. (2017). Construct under-representation and construct irrelevant variances on IELTS academic writing task 1: is there any threat to validity? *Theory and Practice in Language Studies, 7*(11), 1097-1107.

Aldemir, A. Y. & Er, K. O. (2012). Türkiye ve Japonya'da İngilizce öğretmeni yetiştirme sistemlerinin karşılaştırılması. *Journal of Faculty of Educational Sciences 45*(1)

Alderson, J. C. 2004. Foreword in L. Cheng, Y. Watanabe, and A. Curtis (eds.). *Washback in language testing: Research contexts and methods*. London: Lawrence Erlbaum.

Al-Issa, A., Al-Bulushi, A., & Al-Zadjali, R. (2017). Student teachers' perceptions on the selection criteria of English language teachers in the Sultanate of Oman: Implications for policy reconsideration. *Journal of Language and Linguistic Studies*, *13*(2), 235-259.

Allen, D. (2016). Investigating washback to the learner from the IELTS test in the Japanese tertiary context. *Language Testing in Asia*, *6*(7), 1-20.

Allen, M. (2003). Eight Questions on Teacher Preparation: What Does the Research Say? A Summary of the Findings.

Allen, M. B. (2005). Eight questions on teacher recruitment and retention: what does the research say?. *Education Commission of the States (NJ3)*.

Allen, J. M. (2009). Valuing practice over theory: How beginning teachers re-orient their practice in the transition from the university to the workplace. *Teaching and Teacher Education, 25*(5), 647-654.

Alptekin, C., & Tatar, S. (2011). A country in focus: research on foreign language teaching and learning in Turkey 2005-2009. *Language Teaching, 44(3),* 328-353.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Antonenko, P. D. (2015). The instrumental value of conceptual frameworks in educational technology research. *Educational Technology Research and Development*, *63*(1), 53-71.

Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y., & Schady, N. (2016). Teacher quality and learning outcomes in kindergarten. *The Quarterly Journal of Economics*, *131*(3), 1415-1453.

Atav, E., & Sönmez, S. (2013). Öğretmen adaylarının kamu personeli seçme sınavı (KPSS)'na ilişkin görüşleri. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, Özel Sayı*, *1*, 01-13.

Atkinson, P. A., & Coffey, A. (1997). Analysing documentary realities. In D. Silverman (Ed.), *Qualitative research: Theory, method and practice*, London: Sage, 45–62.

Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher, 36*, 258–267.

Averill, R. (2012). Caring teaching practices in multiethnic mathematics classrooms: Attending to health and well-being. *Mathematics Education Research Journal, 24*(2), 105–128.

Aydın, A., & Başkan, G. A. (2005). The problem of teacher training in Turkey. *Biotechnology & Biotechnological Equipment*, 19(2), 191-197.

Bachman, L.F. (1990). Fundamental considerations in language testing. Oxford: OUP.

Bachman, L. F., & Palmer, A. S. (1996). Language testing in practice. Oxford, England: Oxford University.

Bağçeci, B., Yıldırım, İ. Kara, K., & Keskinpalta, D. (2015). Pedagojik formasyon ve eğitim fakültesi öğrencilerinin öğretmenlik mesleğine yönelik tutumlarının karşılaştırılması. *Erzincan Üniversitesi Eğitim Fakültesi Dergisi*, *17*(1), 307-324.

Bailey, K.M. (2006). *Language teacher supervision: A case-based approach*. United States: Cambridge University.

Baldauf, R.B., Kaplan, R.B., Kamwangamalu, N., & Bryant, P. (2011). Success or failure of primary second/foreign language programmes in Asia: what do the data tell us? *Current Issues in Language Planning, 12*(2), 309-323.

Ballou, D. (1996). Do public schools hire the best applicants?. *The Quarterly Journal of Economics*, *111*(1), 97-133.

Balter, D., & Duncombe, W. D. (2008). Recruiting highly qualified teachers: Do district recruitment practices matter? *Public Finance Review, 36*(1), 33-62.

Barnes, B. D., & Lock, G. (2010). The attributes of effective lecturers of English as a foreign language as perceived by students in a Korean University. *Australian Journal of Teacher Education, 35*(1), 139-152.

Bauer, J., & Kenton, J. (2005). Toward technology integration in the schools: Why it isn't happening. *Journal of Technology and Teacher Education, 13*(4), 519-546.

Beard, C., & Wilson, J. P. (2013). *The Power of experiential learning: A Handbook for education, training and coaching* (3rd ed.). London: Kogan Page.

Beauchamp, G., Clarke, L., Hulme, M., & Murray, J. (2013). *Research and teacher education: The BERA-RSA inquiry. Policy and practice within the United Kingdom*. Project report. London, UK: British Educational Research Association.

Berk, R. A. (2002). *Humour as an ınstructional defibrillator: evidence-based techniques in teaching and assessment*. Sterling, VA: Stylus.

Berry, B., Daughtrey, A., & Wieder, A. (2009). Teaching effectiveness and the conditions that matter most in high-needs schools: a policy brief. *Center for Teaching Quality*.

Beyer, B. M., & Johnson, E. S. (2005). *Special programs and services in schools: Creating options, meeting needs*. Lancaster, PA: Pro Active Publications.

Blömeke, S., Olsen, R. V., & Suhl, U. (2016). Relation of student achievement to the quality of their teachers and instructional quality. In *Teacher Quality, Instructional Quality and Student Outcomes* (pp. 21-50). Springer, Cham.

Borghans, L., Duckworth, A. L., Heckman, J. J., & Ter Weel, B. (2008). The economics and psychology of personality traits. *Journal of Human Resources*.

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative research journal, 9*(2), 27-40.

Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2005). Explaining the short careers of high-achieving teachers in schools with low-performing students. *American economic review*, *95*(2), 166-171.

Brant, J. (2006). Subject knowledge and pedagogic knowledge: ingredients for good teaching? An English perspective. *Edukacja*, *94*(2), 60-77.

Brewer, C., Knoeppel, R. C., & Lindle, J. C. (2015). Consequential validity of accountability policy: Public understanding of assessments. *Educational Policy*, *29*(5), 711-745.

Bringer, J. D., Johnston, L. H., & Brackenridge, C. H. (2004). Maximizing transparency in a doctoral thesis1: The complexities of writing about the use of QSR* NVIVO within a grounded theory study. *Qualitative research*, *4*(2), 247-265.

Brosh, H. (1996). Perceived characteristics of an effective language teacher. *Foreign Language Annals, 29(2)*, 125-138.

Brown, H. D. (2001). *Teaching by principles: An interactive approach to language and pedagogy*. White Plains, NY: Longman.

Brown, D. J., & Lord, R. G. (1999). The utility of experimental research in the study of transformational/charismatic leadership. *Leadership Quarterly, 10*(4), 531-539.

Brutti, Z., & Sánchez, F. (2017). Does better teacher selection lead to better students? evidence from a large scale reform in Colombia. Documento de trabajo CEDE .

Bryman, A. (2006). Integrating quantitative and qualitative research: How is it done? *Qualitative research, 6*(1), 97-113.

Buyukkantarcioğlu, N. (2004). A sociolinguistic analysis of the present dimensions of English as a foreign language in Turkey. *International Journal of the Sociology of Language 165*, 33–58.

Büyüköztürk, S. (2011). *Sosyal Bilimler İçin Veri Analizi El Kitabı: İstatistik, Araştırma Deseni SPSS Uygulamaları ve Yorum.* Ankara: Pegem Akademi.

Caena, F. (2011). Education and Training 2020, Thematic Working Group 'Professional. Development of Teachers. *Literature review Quality in Teachers' continuing professional development. EUROPEAN COMMISSION (EU), Directorate-General for Education and Culture.*

Cain-Caston, M. (1999). A survey of opinions of North Carolina school administrators regarding factors considered most important in hiring teachers for their first teaching position. *Journal of Instructional Psychology*, *26*(2), 69-69.

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.

Canbulat, M., & Canbulat, A. N. K. (2015). Avusturya ve Türkiye'de Ögretmen Adayi Belirleme Süreci. *Pegem Egitim ve Ogretim Dergisi (Pegem Journal of Education and Instruction)*, *5*(1), 113.

Cantrell, S., & Kane, T. J. (2013). Ensuring fair and reliable measures of effective teaching: Culminating findings from the MET project's three-year study. MET Project Research Paper. Retrieved from https://files.eric.ed.gov/fulltext/ED540958.pdf

Casey, C. E., & Childs, R. A. (2007). Teacher Education Program Admission Criteria and What Beginning Teachers Need to Know to Be Successful Teachers. *Canadian Journal of Educational Administration and Policy, 67*, 1-24.

Castillo-Montoya, M. (2016). Preparing for interview research: The interview protocol refinement framework. *The Qualitative Report, 21*(5), 811-831.

Chamberlain, G. E. (2013). Predictive effects of teachers and schools on test scores, college attendance, and earnings. *Proceedings of the National Academy of Sciences*, 201315746.

Chambless, K.S. (2012). Teachers' oral proficiency in the target language: Research on its role in language teaching and learning. *Foreign Language Annals, 45* (1), 141–162.

Chen, Y.-H., Li, I. Y, & Chason, W. (2018). Construct underrepresentation. In Frey, B. (Ed.), *The SAGE encyclopedia of educational research, measurement, and evaluation.* Thousand Oaks, CA: SAGE Publications.

Chen, Y.-J., & Lin, S.-C. (2009). Exploring characteristics for effective EFL teachers from the perceptions of junior high school students in Tainan. *STUT Journal of Humanities and Social Sciences, 2*, 219-249.

Cheng, L. (2005). *Changing language teaching through language testing: A washback study*. Cambridge: CUP.

Cheng, L. (2014). Consequences, impact and washback. In A. J. Kunnan (Ed.), *The companion to language assessment: Evaluation, Methodology and Interdisciplinary themes* (pp. 1130–1145). Singapore: Wiley Blackwell.

Cheng, L., Watanabe, Y., & Curtis, A. (2004). *Washback in language testing*. New Jersey: Lawrence Erlbaum Association, Inc.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review, 104*(9), 2633-79.

Cho, Y., & So, Y. (2014). Construct-irrelevant factors influencing young English as a foreign language (EFL) learners' perceptions of test task difficulty. *Research Memorandum No. RM-14-04). Princeton, NJ: Educational Testing Service*.

Choi, I. (2008). The impact of EFL testing on EFL education in Korea. *Language Testing, 25*(1), 39–62

Cizek, G. J., Rosenbert, S. L., & Koons, H. H. (2008). Sources of validity evidence for educational and psychological tests. *Educational and Psychological Measurement, 68*, 397-412.

Clement, M. C. (2009). Hiring highly qualified teachers begins with quality interview. *Kappan, 9*(2), 22–24.

Clement, M.C., Kistner, W., & Moran, W. (2005). A question of experience: an interviewing technique that requires candidates to summarize and draw conclusions

from their experience can be a good predictor of their effectiveness in the classroom. *Principal Leadership 5*(9), 58-62.

Cohen, D. K. (2010). Teacher quality: An American educational dilemma. In M. M. Kennedy (Ed.), *Teacher assessment and the quest for teacher quality: A handbook* (pp. 375-401). San Francisco, CA: Jossey-Bass.

Collins, J. (2001). *Good to great: Why some companies make the leap and others don't*. Colorado: Harper Business.

Collins, K., Onwuegbuzie, A., & Sutton, I. (2006). A model incorporating the rationale and purpose for conducting mixed-methods research in special education and beyond. Learning Disabilities: *A Contemporary Journal, 4*(1), 67-100.

Coniam, D., & Falvey, P. (2007). High-stakes testing and assessment. In *International handbook of English language teaching* (pp. 457-471). Springer, Boston, MA.

Cook, V. (2001). *Second Language Learning and Language Teaching*. New York: Oxford University Press.

Cook, D. A., & Lineberry, M. (2016). Consequences validity evidence: evaluating the impact of educational assessments. *Academic Medicine, 91*(6), 785-795.

Council of Europe (2011). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University.

Council, B. (2015). The state of English in higher education in Turkey. A baseline study. Retrieved from https://www.britishcouncil.org.tr/sites/default/files/he_baseline_study_book_web_-_son.pdf

Creemers, B., & Kyriakides, L. (2008). *The dynamics of educational effectiveness: a contribution to policy, practice and theory in contemporary schools*. New York: Routledge.

Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approach* (2nd ed.). Thousand Oaks, CA: Sage Publications, Inc

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage.

Creswell, J. W. (2012). *Educational research: Planning conducting and evaluating quantitative and qualitative research (4th ed.).* Boston, MA: Pearson

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research (2nd ed.)*. Thousand Oaks, CA: Sage.

Cripps Clark, J., & Walsh, J. (2002, January). Elements of a model of effective teachers. In *AARE 2002: Problematic futures: educational research in an era of uncertainty; AARE 2002 conference papers* (pp. 1-11). Australian Association for Research in Education.

Crow, G., & Wiles, R. (2008). Managing anonymity and confidentiality in social research: The case of visual data in community research. Retrieved from http://eprints.ncrm.ac.uk/459/

Cruickshank, D.R., & Haefele, D. (2001) Good teachers, plural. *Educational Leadership, 58*(5), 26-30.

Çakır, İ. (2017). The washback effects of secondary education placement examination on teachers, school administrators and parents with specific reference to teaching English as a foreign language. *Turkish Journal of Teacher Education, 6*(2) , 61-73.

Çakıroğlu, E., & Çakıroğlu, J. (2003). Reflections on teacher education in turkey. *European Journal of Teacher Education, 26*(2), 253-264.

Çelik, S. (2011). Characteristics and competencies for teacher educators: Addressing the need for improved professional standards in turkey. *Australian Journal of Teacher Education*, *36*(4), 73-87.

Çelik, S., Arıkan, A., & Caner, M. (2013). In the eyes of Turkish EFL learners: What makes an effective foreign language teacher? *Porta Linguarum, 20*, 287-297.

Çılgın, A. (2009). Öğretmen seçim politikaları ve aday öğretmenin işe kabulü için sahip olması gereken nitelikler (teacher selection policies and qualifications required for teacher recruitment). Unpublished Dissertation. Yeditepe University

Çimen, O., & Yılmaz, M. (2011). Biyoloji öğretmen adaylarının KPSS ile ilgili görüşleri. *Gümüşhane Üniversitesi Sosyal Bilimler Elektronik Dergisi, 4*, 159-172.

Danielson, C. (1996). *Enhancing Professional Practice: A Framework for Teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.

Danielson, C. (2007). *Enhancing Professional Practice: A Framework for Teaching*. (2nd Ed.) Alexandria, VA: Association for Supervision and Curriculum Development.

Danielson, C. (2011). Evaluations that help teachers learn. *Educational Leadership, 68*(4), 35-39.

Daresh, J.C. (2001). *Supervision as a proactive leadership (3rd ed.).* Prospect Heights, IL: Waveland.

Darling-Hammond, L. (2000). How teacher education matters. *Journal of Teacher Education, 51*(3), 166-173.

Darling-Hammond, L. (2010a). Teacher education and the American future. *Journal of Teacher Education*, *61*(1-2), 35-47.

Darling-Hammond, L. (2010b). Recruiting and retaining teachers: Turning around the race to the bottom in high-need schools. *Journal of Curriculum and Instruction*, *4*(1), 16-32.

Darling-Hammond, L. (2012). *Powerful teacher education: lessons from exemplary programs*. San Francisco: JosseyBass.

Darling-Hammond, L. (2014). One piece of the whole: Teacher evaluation as part of a comprehensive system for teaching and learning. *American Educator, 38*(1), 4-13

Darling-Hammond, L., & Rothman, R. (2011). Lessons learned from Finland, Ontario, and Singapore. L. Darling---Hammond and R. Rothman (Eds.), in *Teacher and leader effectiveness in high performing education systems* (p. 1---11). Washington: Alliance for Excellent Education and Stanford: Stanford Center for Opportunity Policy in Education.

Darling-Hammond, L., & Wei, R. C. (2012). Teacher preparation and teacher learning: A changing policy landscape. In *Handbook of education policy research* (pp. 629-652). Routledge.

Darling-Hammond, L., & Wentworth, L. (2010). Benchmarking learning systems: Student performance assessment in international context. *Stanford Center for Opportunity Policy in Education, Stanford University, CA*.

Darling-Hammond, L., & Youngs, P. (2002). Defining "highly qualified teachers": What does "scientifically-based research" actually tells us? Education Researcher, 31(9), 13-25.

Davies, A. (2010). Test Fairness: a response. *Language Testing, 27(2)*, 171-176.

Davies, A., Brown, A., Elder, C., Hill, K., Lumley, T., and McNamara, T. (1999). *Dictionary of language testing* (Vol. 7). Cambridge University Press.

Davies, R. S. (2011). Understanding technology literacy: A framework for evaluating educational technology integration. *TechTrends*, *55*(5), 45.

Daymon, C., & Holloway, I. (2003). *Qualitative Research Methods in Public Relations and marketing communications*. London: Rout ledge.

De Búrca, S., & McLoughlin, D. (1996). *The Grounded Theory Alternative in Business Network Research.* Dublin City University Business School.

Demircan, Ö. (1988). *Dünden bugüne Türkiye'de yabanci dil* [Foreign languages in Turkey: From past to today]. Istanbul: Remzi.

Demirel, Ö. (1991). Türkiye'de yabancı dil öğretmeni yetiştirmede karşılaşılan güçlükler. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 6,* 25-39.

DeArmond, M., Gross, B., & Goldhaber, D. (2010). Is it better to be good or lucky? Decentralized teacher selection in 10 elementary schools. *Educational Administration Quarterly*, *46*(3), 322-362.

Denzin, N.K., & Lincoln, Y. S. (2011). *The Sage Handbook of Qualitative Research.* Thousand Oaks: Sage.

Dhanasobhon, S. (2006). English language teaching dilemma in Thailand. Retrieved from http://www.curriculumandinstruction.org/index.php?lay=show&ac=article&Id=539134523&Ntype=7

Díaz-Rico, L. T. (2008). *A Course for Teaching English Learners*. Boston, MA: Allyn and Bacon.

Dikilitaş, K., Dollar, Y. K., & Mede, E. (2018). English language teachers' perceptions of involvement in high-stakes testing in Turkey. *Teacher Involvement in High-Stakes Language Testing*, 233-244.

Doğançay-Aktuna, S. (1998). The spread of English in Turkey and its current sociolinguistic profile. *Journal of Multilingual and Multicultural Development 19(1),* 24-39.

Doğançay-Aktuna, S., & Kiziltepe, Z. (2005). English in Turkey. *World Englishes*, *24*(2), 253-265.

Donaldson, M. L. (2013). Principals' approaches to cultivating teacher effectiveness: Constraints and opportunities in hiring, assigning, evaluating, and developing teachers. *Educational Administration Quarterly*, *49*(5), 838-882.

Doughty, C., and M. H. Long, eds. 2003. *Handbook of Second Language Acquisition.* New York: Basil Blackwell.

Dounay, J. {2000). High stakes testing is high-stress too. *Education Digest, 65*(9), 9-13.

Doyle, S. (2007). Member checking with older women: A framework for negotiating meaning. *Health care for women international*, *28*(10), 888-908.

Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative and mixed methodologies*. Oxford: Oxford University.

Educational Testing Service. (2009). ETS guidelines for fairness review of assessments. Retrieved from https://www.ets.org/Media/About_ETS/pdf/overview.pdf

Educational Testing Service, (2012). Relationships between Big Five and Academic and Workforce Outcomes. Retrieved from https://www.ets.org/s/workforce_readiness/pdf/21334_big_5.pdf

Educational Testing Service. (2014). *ETS standards for quality and fairness*. Princeton, NJ: Educational Testing Service.

Ellis, C., Skidmore, S. T., & Combs, J. P. (2017). The hiring process matters: The role of person–job and person–organization fit in teacher satisfaction. *Educational Administration Quarterly*, *53*(3), 448-474.

Enever, J. (2011). *ELLiE: Early language learning in Europe*. London: British Council.

Enever, J., & Moon. J. (2009). New global contexts for teaching primary ELT: Change and challenge. In J. Enever, J. Moon, & U. Raman (Eds.), *Young Learner English language policy and implementation: International perspectives* (pp. 5-21). Reading: Garnet Education. https://doi.org/10.1111/j.1467-971X.2012.01766.x

Engel, M., & Finch, M. A. (2015). Staffing the classroom: How urban principals find teachers and make hiring decisions. *Leadership and Policy in Schools*, *14*(1), 12-41.

Erdem, E. & Soylu, Y, (2013). Öğretmen adaylarının KPSS ve alan sınavına ilişkin görüşleri. *Sosyal Bilimler Enstitüsü Dergisi, 4*(1): 223-236

European Commission (2013a). Supporting Teacher Competence Developments for Better Learning Outcomes. Retrieved from http://ec.europa.eu/dgs/education_culture/repository/education/policy/school/doc/teachercomp_en.pdf

European Commission. (2013b). *EU performance and first inferences regarding education and training policies in Europe*. Brussels, Belgium: Author. Retrieved from http://ec.europa.eu/dgs/education_culture/repository/education/policy/strategic-framework/ doc/pisa2012_en.pdf

Eurydice. (2012). *Key data on teaching languages at school in Europe, Brussels: Education, audiovisual and culture executive agency*. Retrieved from http://eacea.ec.europa.eu/education/eurydice/documents/key-dataseries/143en.pdf

Faez, F., & Karas, M. (2017). Connecting language proficiency to (self-reported) teaching ability: A Review and analysis of research. *RELC Journal,* 48(1), 135-151.

Ferguson, R.F. (1991). Paying for public education: New evidence on how and why money matters. *Harvard Journal on Legislation 28*, 465-498.

Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs—principles and practices. *Health services research, 48*(6pt2), 2134-2156.

Field, A. (2009). *Discovering statistics using SPSS (and sex and drugs and rock 'n' roll) (Third edition)*. London: SAGE.

Fox, J., & Cheng, L. (2007). Did we take the same test? Differing accounts of the Ontario Secondary School Literacy Test by first and second language test-takers. *Assessment in Education: Principles, Policy and Practice*, *14*(1), 9–26.

Fraenkel, J., Wallen, N., & Hyun, H. (2011). *How to design and evaluate research in education.* (8th ed.) New York, NY: McGraw-Hill.

Freeman, D. (2017). The Case for teachers' classroom English proficiency. *RELC Journal*, 48 (1), 31-52.

Friedman, J. (1994). *Cultural Identity and Global Process* (London: Sage Publications).

Garcia Laborda, J., Litzler, M. F., Gonzalez Such, J., Bakieva, M., & Otero de Juan, N. (2012). Washback factors and inference predictors in the Spanish university examination: The OPENPAU Project (FFI2011-22442). *Online Submission*, *46*, 5914-5918.

Garton, S. (2014). Unresolved issues and new challenges in teaching English to young learners: The case of South Korea, *Current Issues in Language Planning*, *15*(2), 201-219. https://doi.org/10.1080/14664208.2014.858657

Gasper, M. F., & Vieira, C. R., (2013). PLENATTITUDE teacher evaluation for teacher Effectiveness and well-being with neuro-linguistic programming. *US-China Education Review, 3*(1), 1-17.

Gass, S. M., & Selinker. L. (2008). *Second Language Acquisition: An Introductory Course*. New York: Routledge.

Gay, L.R., Mills, G.E., & Airasian, P.W. (2011). *Educational research: Competencies for analysis and applications.* (10th ed.) Boston, MA: Pearson.

Ghasemi, B., & Hashemi, M. (2011). The study of the characteristics of successful English language teachers from the view point of the English language students of Islamic Azad University, Hamedan Branch, in *Procedia-Social and Behavioral Sciences*, *28*, 411-415.

Glaser, B. G., & Strauss, A. (1967). *The Discovery of Grounded Theory. Strategies for Qualitative Research*. Chicago: Aldine

Goe, L., Bell, C., & Little, O. (2008). Approaches to Evaluating Teacher Effectiveness: A Research Synthesis. *National Comprehensive Center for Teacher Quality*.

Goldhaber, D., & Anthony, E. (2007). Can teacher quality be effectively assessed? National board certification as a signal of effective teaching. *Review of Economics and Statistics, 89*(1), 134–150.

Goldhaber, D., & Brewer, D. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis, 22*(2), 129-145.

Goldhaber, D., Grout, C., & Huntington-Klein, N. (2014, December). Screen twice, cut once: Assessing the predictive validity of teacher selection tools (CEDR Working Paper No. 2014-9). Seattle, WA: Center for Education Data and Research.

Goldhaber, D., & Hansen, M. (2013). Is it just a bad class? Assessing the long-term stability of estimated teacher performance. *Economica, 80*(319), 589–612.

Gökçe, F. (2013). Lisans öğrencileri ile formasyon programına devam eden kursiyerlerin KPSS konusundaki görüşleri. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, Özel Sayı (1)*, 171-190

Graneheim, U. H., & Lundman, B. (2004). Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. *Nurse education today*, *24*(2), 105-112.

Green, D. R. (1998). Consequential aspects of the validity of achievement tests: A publisher's point of view. *Educational Measurement: Issues and Practice*, *17*(2), 16-19.

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational evaluation and policy analysis, 11*(3), 255-274.

Grinyer, A. (2009). The anonymity of research participants: Assumptions, ethics, and practicalities. *刊名: Pan-Pacific Management Review*. Retrieved from http://ir.lib.isu.edu.tw/ir/retrieve/33373/11933.pdf

Grossman, P., Hammerness, K., & McDonald, M. (2009). Redefining teaching, re-imagining teacher education. Teachers and Teaching: *Theory and Practice, 15*, 273– 289.

Guba, E. G. (1981). Criteria for Assessing the Trustworthiness of Naturalistic Inquiries. *Educational Communication and Technology, 29*(2), 75-91. Retrieved from https://pdfs.semanticscholar.org/8d32/23ed3c76cc4066ec894b5aca51c4f4028b7e.pdf

Guha, R., Hyler, M. E., & Darling-Hammond, L. (2017). The teacher residency: A practical path to recruitment and retention. *American Educator*, *41*(1), 31.

Gupta, K. (2011). *A practical guide to needs assessment*. John Wiley & Sons.

Gurtler, L. (2002). 'Humour in educational contexts'. Paper presented at the 110th Annual Meeting of the American Psychological Association. Chicago, 23 August.

Güneş, B. (2009). *1945-1980 arası Türkiye'de İngilizce eğitimi*. Unpublished master's dissertation, Süleyman Demirel University, Isparta.

Günüç S. (2009), *İnternet bağımlılık ölçeğinin geliştirilmesi ve bazı demografik değişkenler ile internet bağımlılığı arasındaki ilişkilerin incelenmesi.* Unpublished master's dissertation. Yüzüncü Yıl University.

Gürsoy, E., Korkmaz, Ş. Ç., & Damar, E. A. (2013). Foreign language teaching within 4+4+4 education system in Turkey: Language teachers' voices. *Eurasian Journal of Educational Research*, *53*(A), 59-74.

Gürsoy, E., Korkmaz, S. C., & Damar, E. A. (2017). English Language Teaching within the New Educational Policy of Turkey: Views of Stakeholders. *International Education Studies*, *10*(4), 18.

Güven, I. 2008. Teacher education reform and international globalization hegemony: Issues and challenges in Turkish teacher education. *World Academy of Science, Engineering and Technology 16*, 625–634.

Haladyna, T. M., & Downing, S. M. (2004). Construct-irrelevant variance in high-stakes testing. *Educational Measurement: Issues and Practice*, *23*(1), 17-27.

Hanushek, E. (2010). The economic value of higher teacher quality. Cambridge: National Bureau of Economic Research. Working Paper 16606. Retrieved from http://www.nber.org/papers/w16606

Hanushek, E. A., & Woessmann, L. (2012). Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *Journal of economic growth*, *17*(4), 267-321.

Hanushek, E. A., & Rivkin, S. G. (2012). The distribution of teacher quality and implications for policy. *Annu. Rev. Econ.*, *4*(1), 131-157.

Hanushek, E. A., & Rivkin, S. G. (2010). Generalizations about using value-added measures of teacher quality. *American Economic Review, 100*(2), 267–271.

Hanushek E. A., & Rivkin S.G. (2006). Teacher quality. In *Handbook of Economics of Education, Vol. 2*, Hanushek E, Welch F (eds.). Elsevier: Amsterdam; 1051–1075.

Hansen, K. (2003). *Writing in the social sciences: A rhetoric with readings.* Boston: Pearson Custom.

Harding, J. (2013). *Qualitative data analysis: From start to finish*. Thousand Oaks, CA: Sage Publications.

Harris, D. N., & Sass, T. R. (2014). Skills, productivity and the evaluation of teacher performance. *Economics of Education Review*, *40*, 183-204.

Harris, D. N., Ingle, W. K., & Rutledge, S. A. (2014). How teacher evaluation methods matter for accountability: A comparative analysis of teacher effectiveness ratings by principals and teacher value-added measures. *American Educational Research Journal, 51*(1), 73-112.

Harris, D. N., & McCaffrey, D. (2010). Value-added: Assessing teachers 'contributions to student achievement. In M. M. Kennedy (Ed.), *Teacher assessment and the quest for teacher quality: A handbook* (pp. 251-282). San Francisco, CA: Jossey-Bass.

Hatipoğlu, Ç. (2013). English Language Testing and Evaluation (ELTE) Teaching Policies and Needs in Turkey. Paper presented at *the 19th Conference of the International Association for World Englishes: World Englishes in the Multilingual World*, Arizona State University, Temple, Arizona, USA, 16-18 November 2013, p. 25

Hatipoğlu, Ç. (2016). The impact of the university entrance exam on EFL education in Turkey: Pre-service English language teachers' perspective. *Procedia-Social and Behavioral Sciences, 232*, 136-144.

Haycock K., & Jerald, C. (2002). Closing the gap. *Principal, 82*(2), 20-23.

Head, B., Ferguson, M., Cherney, A., & Boreham, P. (2014). Are policy-makers interested in social research? Exploring the sources and uses of valued information among public servants in Australia. *Policy and Society, 33*(2), 89-101.

Heneman, H. G., Milanowski, A., Kimball, S. M., & Odden, A. (2006). Standards-based teacher evaluation as a foundation for knowledge- and skill-based pay (CPRE Policy Brief No. RB-45). Philadelphia: Consortium for Policy Research in Education. Retrieved from https://files.eric.ed.gov/fulltext/ED521219.pdf

Henning, G. (1987). *A guide to language testing: Development, evaluation, research. Rowley, MA: Newbury House.*

Henry, M. (2000). History and humour: A natural partnership. *OAH Magazine of History 14* (2), pp. 64–65.

Herrmann, Z. (2018). Rethinking teacher recruitment. *Educational Leadership*, *75*(8), 18-23.

Hsieh, F-J., Ling, P-J., Chao, G., & Wang, T-Y. (2013). Preparing teachers of mathematics in Chinese Taipei. In J. Schwille, L. Ingvarson, & R. Holdgreve-Resendez (Eds.), *TEDS-M encyclopaedia: A guide to teacher education context, structure, and quality assurance in 17 countries* (pp. 71–85). Amsterdam, Netherlands: International Association for the Evaluation of Educational Achievement.

Hitchcock, J. H., Onwuegbuzie, A. J., & Khoshaim, H. B. (2015). Examining the consequential validity of standardized examinations via public perceptions: A review of mixed methods survey design considerations. *International Journal of Multiple Research Approaches, 9*(1), 24-39.

Holloway, I., & Wheeler, S. (2010). *Qualitative research in nursing and healthcare*. Oxford, UK: Blackwell.

Howitt, D., & Cramer, D. (2011). *Introduction to SPSS statistics in psychology: For version 19 and earlier (Fifth edition)*. London: Pearson Education Limited.

Huck, S. (2012). *Reading statistics and research. Boston*, MA: Pearson

Hughes, A. 1989. *Testing for Language. Teachers*. Cambridge: Cambridge University.

Hughes, A. (2003). *Testing for Language Teachers* (2nd.ed.). UK: Cambridge University.

Ingle, K., Rutledge, S., & Bishop, J. (2011). Context matters: principals' sensemaking of teacher hiring and on-the-job performance. *Journal of Educational Administration*, *49*(5), 579-610.

Ingersoll, R. (2007). *A comparative study of teacher preparation and qualifications in 6 nations*. Philadelphia: University of Pennsylvania, Consortium for Policy Research in Education

Ingersoll, R. M. (2001). *Teacher turnover, teacher shortages, and the organization of schools* (Document R-01-1). Seattle: University of Washington, Center for the Study of Teaching and Policy.

Ingersoll, R. & Merrill, L. (2010). Who's teaching our children? *Education Leadership 67*(8), 14 -20.

Ingersoll, R. M., & Smith, T. M. (2003). The wrong solution to the teacher shortage. *Educational leadership*, *60*(8), 30-33.

Ingvarson, L., & Rowley, G. (2017). Quality assurance in teacher education and outcomes: A study of 17 countries. *Educational Researcher*, *46*(4), 177-193.

Isoré, M. (2009). Teacher evaluation: Current practices in OECD countries and a literature review", *OECD Education Working Papers*, No. 23, OECD Publishing. Retrieved from http://dx.doi.org/10.1787/223283631428

Jackson, C. K. (2012). *Non-cognitive ability, test scores, and teacher quality: Evidence from 9th grade teachers in North Carolina* (No. w18624). National Bureau of Economic Research.

Jackson, C.K. (2014). Teacher quality at the high-school Level: The importance of accounting for tracks, *Journal of Labor Economics, 32*(4): 645-684.

Jackson, J. (2002). Reticence in second language case discussions: Anxiety and aspirations. *System*, *30*(1), 65-84.

Jacob, B. A. (2016). The power of teacher selection to improve education. *Evidence Speaks Reports, 1*(12), 1-7.

Jacob, B., Rockoff, J. E., Taylor, E. S., Lindy, B., & Rosen, R. (2016). *Teacher applicant hiring and teacher performance: Evidence from DC public schools* (No. w22054). National Bureau of Economic Research.

James Popham, W., C. Berliner, D., M. Kingston, N., H. Fuhrman, S., M. Ladd, S., Charbonneau, J., & Chatterji, M. (2014). Can today's standardized achievement tests yield instructionally useful data? Challenges, promises and the state of the art. *Quality Assurance in Education*, *22*(4), 303-318.

Jensen, M.S. (1989). Leading the instructional staff. In Smith, S.C., & Piele, P.K. *School leadership. Handbook for excellence. 2nd ed*. Washington, D.C.: Office of Educational Research and Improvement. ERIC: EA 020 974

Johnston, L. (2006). Software and method: Reflections on teaching and using QSR NVivo in doctoral research. *International Journal of Social Research Methodology*, *9*(5), 379-391.

Johnstone, R. (2004). Language teacher education. In A. Davies & C. Elder (eds.), *The handbook of applied linguistics* (pp. 649-671). Malden, MA: Blackwell Publishing Ltd. http://dx.doi.org/10.1002/9780470757000.ch26

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational researcher, 33(7)*, 14-26.

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research, 1*(2), 112-133.

Johnson, S. M., & Birkeland, S. E. (2003). Pursuing a "sense of success": New teachers explain their career decisions. *American Educational Research Journal*, *40*(3), 581-617.

Kachru, B.B. (1992). The other side of English and the 1990s. In B. B. Kachru (ed.), *The other tongue: English across cultures* (pp. 27-54). Urbana/Chicago: University of Illinois Press

Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp.17–64). Westport, CT: American Council on Education.

Kang, N. H., & Hong, M. (2008). Achieving excellence in teacher workforce and equity in learning opportunities in South Korea. *Educational Researcher, 37*(4), 200-207.

Karaca, E. (2011). Öğretmen adaylarinin kamu personeli seçme sinavi'na (KPSS) yönelik tutumlari (Student teachers' attitudes towards the public service selection examination). *Akademik Bakış Dergisi. Vol. 1* pp. 1-18

Karahan, F. (2007). Language attitudes of Turkish students towards the English language and its use in Turkish context. *Journal of arts and sciences*, *7*(1), 73-87.

Keçik, İ., & Aydın, B. (2011). Achieving the impossible? Teaching practice component of a pre-service distance English language teacher training program in Turkey, *Australian Journal of Teacher Education, 36*(4), 73-83.

Kennedy, M. M. (2008). Sorting out teacher quality. *Phi Delta Kappan, 90*(1), 59–63.

Kersten, T. (2008). Teacher hiring practices: Illinois principals' perspectives. *The Educational Forum, 72*(4), 355-368.

Khamkhien, A. (2010). Teaching English speaking and English speaking tests in the Thai context: A reflection from Thai perspective. *English Language Teaching, 3*(1), 184-190.

Kılıçkaya, F., & Krajka, J. (2010). Language teachers' views and suggestions on the central teacher selection and recruitment exam in Turkey. *The New Educational Review, 20*(3), 253-260.

Kılıçkaya, F., & Krajka, J. (2013). A review of language teacher selection examination and recruitment in Turkey and Poland. *International Journal of Research Studies in Language Learning, 2*(5), 93-104

Kılıçkaya, F. 2016. Washback effects of a high-stakes exam on lower secondary school English teachers' practices in the classroom. *Lublin Studies in Modern Languages and Literature, 40*(1), 116-134.

Kırkgöz, Y. (2005). English Language Teaching in Turkey: Challenges for the 21st Century, in G. Braine (ed.), *Teaching English to the World: History, Curriculum, and Practice* (Mahwah: Lawrence Erlbaum Associates): 159-75.

Kırkgöz, Y. (2007). English language teaching in Turkey: Policy changes and their implementations. *RELC Journal*, *38*(2), 216-228.

Kırkgöz, Y. (2009) Globalization and English language policy in Turkey. *Educational Policy, 23(5)*, 663-684.

Kirkgöz, Y. (2016). Türkiye'deki yabancı dil Eğitim politikaları ve müfredat programı. *Türkiye'deki yabancı dil eğitimi üzerine görüş ve düşünceler* Akcan, S. & Bayyurt, Y. (Haz.). Istanbul: Boğaziçi Üniversitesi Matbaası.

Kim, E., & Han, Y. (2002). *Attracting, developing, and retaining effective teachers: Background report for Korea*. Seoul: Korean Educational Development Institute

Kim, K., Kim, G., Kim, S., Kim, J., Kim, J., Ryoo, I., ... & Choi, S. J. (2010). OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes.

Kiany, G. R., Shayestefar, P., Samar, R. G., & Akbari, R. (2013). High-rank stakeholders' perspectives on high-stakes University entrance examinations reform: priorities and problems. *Higher Education*, *65*(3), 325-340.

Klassen, R. M., Durksen, T., Kim, L., Patterson, F., Rowett, E., Warwick, J., ... & Wolpert, M. A. (2017). Developing a proof-of-concept selection test for entry into primary teacher education programs. *International Journal of Assessment Tools in Education*, 96-114.

Kocaoluk, F., & Kocaoluk, M.Ş. (2001). İlköğretim Okulu programı 1999–2000 (Primary Education Curriculum 1999–2000; İstanbul: Kocaoluk Publishers).

Koedel, C. (2008). Teacher quality and dropout outcomes in a large, urban school district. Journal of Urban Economics.

Koedel, C., Mihaly, K., & Rockoff, J. E. (2015). Value-added modeling: A review. *Economics of Education Review*, *47*, 180-195.

Korstjens, I., & Moser, A. (2018). Series: Practical guidance to qualitative research. Part 4: Trustworthiness and publishing. *European Journal of General Practice*, *24*(1), 120-124.

Kuhlman, N., & Knezevle, B. (2013). The TESOL guidelines for developing EFL standards. Retrieved from https://www.tesol.org/docs/default-source/pdf/the-tesol-guidelines-for-developing-efl-standards.pdf?sfvrsn=0

Kumar, R. (2011). *Research methodology: A step-by-step guide for beginners* (3rd Edition). London: Sage.

Kumar, K., Roberts, C., Bartle, E., & Eley, D. S. (2018). Testing for medical school selection: What are prospective doctors' experiences and perceptions of the GAMSAT and what are the consequences of testing? *Advances in Health Sciences Education*, 1-14.

Kunnan, A. J. (2005, May). Towards a model of test evaluation: Using the test fairness and test context frameworks. In *Multilingualism and assessment: Achieving transparency, assuring quality, sustaining diversity. Proceedings of the ALTE Berlin Conference* (pp. 229-251).

Külekçi, E. (2016). A concise analysis of the Foreign Language Examination (YDS) in Turkey and its possible washback effects. *International Online Journal of Education and Teaching, 3*(4), 303-315.

Kwon, O. N., & Ju, M. K. (2012). Standards for professionalization of mathematics teachers: policy, curricula, and national teacher employment test in Korea. (2), 211-222.

Kvale, S. (2007). *Doing interviews*. Thousand Oaks, CA: Sage.

Kyriacou, C. (2009). *Effective Teaching in Schools* (3rd edn). Cheltenham: Nelson Thornes.

Laczko-Kerr, I., & Berliner, D. C. (2002). The effectiveness of "Teach for America" and other under-certified teachers on student academic achievement: A case of harmful public policy. *Education Policy Analysis Achieves*, *10*(37), 1-53

Loumbourdi, L. (2014). T*he Power and Impact of Standardised Tests: Investigating the Washback of Language Exams in Greece*. Frankfurt am Main: Peter Lang GmbH

Larsen-Freeman, D., & Anderson, M. (2013). *Techniques and Principles in Language Teaching 3rd Edition-Oxford Handbooks for Language Teachers*. Oxford university.

Leshem, S., & Bar-Hama, R. (2007). Evaluating teaching practice. *ELT journal, 62*(3), 257-265.

Leech, N., & Onwuegbuzie, A. (2009). A typology of mixed methods research designs. *Qualitative Quantitative, 43*, 265-275.

Leedy, P. D., & Ormrod J, E. (2005). *Practical Research: Planning and Design*. Upper Saddle River, New Jersey, Pearson

Levine, J. (ed.). (2006). *Motivation in humour*. Edison, NJ: Aldine Transaction.

Lightbown, P. M., & Spada. N. (2006). *How Languages Are Learned*. Oxford: Oxford University

Lincoln, Y.S. & Guba, E. G (1985). *Naturalistic inquiry*. California: Sage Publications.

Liu, E., & Johnson, S. M. (2006). New teachers' experiences of hiring: Late, rushed, and information-poor. *Educational Administration Quarterly, 42*(3), 324-360.

Liu, S., Liu, C., Stronge, J., & Xu, X. (2016). Teacher characteristics for success in the classroom: Chinese principals' perceptions for hiring decisions. *Asia Pacific Education Review*, *17*(1), 107-120.

Liu, S., & Zhao, D. (2013). Teacher evaluation in China: latest trends and future directions. Educational Assessment. *Evaluation and Accountability, 25*(3), 231–250.

Lovorn, M. (2008). 'Humour in the home and in the classroom: The benefits of laughing while we learn'. *The Journal of Education and Human Development* 2 (1), 1–12.

Lovorn, M. (2009). Three easy ways to bring humour into the social studies classroom. *The Leader* 23 (1), 15–16.

Mackay, A., & Gass, S. (2005). *Second Language Research: Methodology and Design*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Mahalingappa, L. J., & Polat, N. (2013). English language teacher education in Turkey: policy vs academic standards. *European Journal of Higher Education*, *3*(4), 371-383.

Mahmoudi, L. (2015). The washback effect of the Iranian national university entrance exam (INUEE) on pre-university students' English learning PROCESS. *British Journal of English Linguistics*, *3*(3), 34-49.

Mason, J. (2002). *Qualitative researching* (2nd ed.). London: SAGE.

Mason, R.W., & Schroeder, M.P. (2010). Principal hiring practices: toward a reduction of uncertainty. *The Clearing House, 83*(5), 186-193.

Mattavarat, S., Viseshsiri, P., & Siribanpitak, P. (2017). Proposed policy for preparation of high-quality primary school teachers in Thailand. *Kasetsart Journal of Social Sciences*, *38*(2), 105-110.

Maxwell, J. A. (2013). *Qualitative research design: An interactive approach* (3rd ed.). Thousand Oaks, CA: Sage Publications.

McCrae, R. R., & Costa, P. T. (1987). Validation of the Five Factor Model of Personality across instruments and observers. *Journal of Personality and Social Psychology*, *52*, 81-90.

McNamara, D. (1991). Subject knowledge and its application: problems and possibilities for teacher educators. *The Journal of Education for Teaching 17*(2), 113–28.

MEB, (1997). İlköğretim ve Eğitim Kanunu. Resmi Gazate. No. 22084. Ağustos 18. Ankara: General Directorate of the Turkish Ministry

MEB, (2006). English Language Curriculum for Primary Education Grades 4, 5, 6, 7 and 8, Ankara: T.C. Milli Eğitim Bakanlığı

MEB, (2013). İlköğretim Kurumları (İlkokullar ve Ortaokullar) İngilizce Dersi (2.-8. Sınıflar) Öğretim Programı. Ankara: T.C. Milli Eğitim Bakanlığı

MEB, (2017a). Teacher strategy paper. Ankara: MEB. Retrieved from https://oygm.meb.gov.tr/meb_iys_dosyalar/2018_05/25170118_Teacher_Strategy_ Paper_2017-2023.pdf

MEB, (2017b). General competencies for teaching profession. http://oygm.meb.gov.tr/meb_iys_dosyalar/2018_06/29111119_TeachersGeneralCo mpetencies.pdf

Medley, D. M., Coker, H., & Soar, R. S. (1984). *Measurement-based evaluation of teacher performance.* New York, NY: Longman.

Meier, D. (2002). Standardization versus standards. *Phi Delta Kappan, 811*(3), 100-198.

Menken K. (2017). High-stakes tests as de facto language education policies. In: Shohamy E., Or I., May S. (eds) *Language Testing and Assessment. Encyclopedia of Language and Education (3rd ed.*). Springer, Cham.

Mentiş Taş, A. (2004). Sosyal bilgiler öğretmenliği eğitimi program standartlarının belirlenmesi. *Eğitim Bilimleri Fakültesi Dergisi. (37)*1, 28-51.

Merriam, S.B. (1998). *Qualitative research and case study applications in education: revised and expanded from case study research in education*. CA: Jossey-Bass.

Mertz, N. T. (2010). Teacher selection and school leader effects. *Journal of School Leadership, 20*(2), 184–207.

Mertz, N.T., and McNeely, S.R. (2001, April). *Blind man's bluff: Instructional leadership, teacher selection and rational decision-making.* Paper presented at the annual meeting of the American Educational Research Association. Seattle, WA.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). New York, NY: Macmillan.

Messick, S. (1995). Validity of psychological assessment: validation of inferences from person's responses and performances as scientific inquiry into score meaning. *American Psychologist, 50(9)*, 741-749.

Mete, Y. A. (2013). Güney Kore, Japonya, Yeni Zelanda ve Finlandiya'da öğretmen yetiştirme ve atama politikaları. *International Periodical for the Languages, Literature and History of Turkish or Turkic, 8*(12), 859-878.

Miles, M. B., & Huberman, M. (1994). *An expanded sourcebook: Qualitative data analysis* (2nd ed.). Newbury Park, CA: SAGE.

Mitchell, R., & Barth, P. (1999). Not Good Enough: A Content Analysis of Teacher Licensing Examinations. How Teacher Licensing Tests Fall Short. *Thinking K-16*, *3*(1), 1-20.

McGrath, I. (2006). Teachers' and learners' images for coursebooks. *ELT journal*, *60*(2), 171-180.

Mohamed, Z., Valcke, M., & De Wever, B. (2017). Are they ready to teach? Student teachers' readiness for the job with reference to teacher competence frameworks. *Journal of Education for Teaching, 43*(2), 151-170.

Moon, T. R., Brighton, C. M., Jarvis, J. M., & Hall, C. J. (2007). *State standardized testing programs: Their effects on teachers and students*. Storrs: University of Connecticut, The National Research Center on the Gifted and Talented.

Moore, A (2004). *The Good Teacher: dominant discourses in teaching and teacher education*. London: Routledge.

Morgan, G. B., & Hodge, K. J. (2014). The stability of teacher performance and effectiveness: Implications for policies concerning teacher evaluation. *Education Policy Analysis Archives*, 22(95), 1-21.

Mourão, S., & Lourenço, M. (2015). *Early years second language education: International perspectives on theory and practice*. Abingdon: Routledge.

Mourshed, M., Chijioke, C., & Barber, M. (2010). *How the world's most improved school systems keep getting better*. McKinsey.

Musset, P. (2010). Initial teacher education and continuing training policies in a comparative perspective: Current practices in OECD countries and a literature review on potential effects. OECD Education Working Papers, 48, OECD Publishing.

Nation, I. S. P., & Macalister, J. (2010). *Language curriculum design*. New York: Taylor and Francis.

National Commission on Teaching & America's Future. (1996). What matters most: Teaching for America's future. New York, NY: Teachers College, Columbia University.

National Council on Teacher Quality. (2013). Teacher preparation review report. Retrieved from https://www.bushcenter.org/publications/articles/2013/06/-national-council-on-teacher-quality-publishes-new-report-on-teacher-preparation-programs.html

NCEE (National Centre for Entrepreneurship in Education) (2013). South Korea: instructional systems. Retrieved from http://www.ncee.org/programs-affiliates/center-on-international-education-benchmarking/top-performing-countries/south-korea-overview/

Nghia, T. L. H. (2015). Vietnamese students' perception of English teacher qualities: Implications for teacher professional development. *International Journal of Academic Research in Education and Review, 3*(1), 7-19.

Nguyen, H. T. M. (2011). Primary English language education policy in Vietnam: Insights from implementation. *Current Issues in Language Planning*, *12*(2), 225-249. https://doi.org/10.1080/14664208.2011.597048

Noam, G. (1996). Assessment at a crossroads: Conversation. *Harvard Educational Review, 66*(3), 631–657.

Noddings, N. (2013). *Caring: A relational approach & moral education* (2nd ed.). Berkley, CA: University of California Press.

Nunan, D. (1992). *Research methods in language learning*. Cambridge University.

Nunan, D., & Bailey, K. M. (2009). Exploring second language classroom research: A comprehensive guide. Heinle ELT.

OBEC, (2015), Number of people who applied for teaching positions in OBEC schools, 2010-2014, Office of the Basic Education Commission, Bangkok.

OECD, (2005). Teachers matter: Attracting, developing and retaining effective teachers. Paris: Organisation for Economic Co-operation and Development. Retrieved from http://www.oecd.org/edu/school/48627229.pdf

OECD, (2009). Education at a glance. OECD Indicators. *OECD Publishing*. doi: 10.1787/eag-2009-en

OECD, (2013a). Education policy outlook: Turkey. Retrieved from http://www.oecd.org/edu/EDUCATION%20POLICY%20OUTLOOK%20TURKEY_EN.pdf

OECD, (2013b). Southeast Asian Economic Outlook 2013:With Perspectives on China and India, *OECD Publishing*. Retrieved from http://dx.doi.org/10.1787/saeo-2013-en.

OECD, (2013c). Lessons from PISA 2012 for the United States: Strong performers and successful reformers in education. Paris, France: Author. http://dx.doi.org/10.1787/9789264207585-en

OECD, (2014). Education policy outlook: Spain. Retrieved from http://www.oecd.org/education/EDUCATION%20POLICY%20OUTLOOK%20SPAIN_EN.pdf

OECD, (2016). Education Policy Outlook: Korea, Retrieved from www.oecd.org/education/policyoutlook.htm

OECD/UNESCO, (2016). Education in Thailand: An OECD-UNESCO perspective, Reviews of National Policies for Education, *OECD Publishing*, Paris. Retrieved from http://dx.doi.org/10.1787/9789264259119-en

Onwuegbuzie, A. J., Bustamante, R. M., & Nelson, J. A. (2010). Mixed research as a tool for developing quantitative instruments. *Journal of Mixed Methods Research*, *4*(1), 56-78.

Onwuegbuzie, A. J., & Collins, K. M. (2007). A typology of mixed methods sampling designs in social science research. *The qualitative report, 12*(2), 281-316.

Onwuegbuzie, A. J., Dickinson, W. B., Leech, N. L., & Zoran, A. G. (2010). Toward more rigor in focus group research in stress and coping and beyond. *Toward a broader understanding of stress and coping. Mixed methods approaches*, 243-285.

ÖSYM, (2017). 2017-KPSS ÖABT Sonuçlarına İlişkin Sayısal Bilgiler. Retrieved from http://www.osym.gov.tr/TR,13307/2017-kpss-oabt-sonuclarina-iliskin-sayisal-bilgiler.html

Özmen, K. S. (2011a). Analyzing washback effect of SEPPPO on prospective English teachers. *The Journal of Language and Linguistic Studies, 7*(2), 24-52.

Özmen, K. S. (2011b). Washback effects of the inter-university foreign language examination on foreign language competencies of candidate academics. *Novitas-ROYAL (Research on Youth and Language), 5*(2), 215-228.

Öztürk, C. (2005). Türkiye'de dünden bugüne öğretmen yetiştiren kurumlar. *İstanbul: Millî Eğitim Bakanlığı Yayınları*.

Pan, Y. C., & Roever, C. (2016). Consequences of test use: a case study of employers' voice on the social impact of English certification exit requirements in Taiwan. *Language Testing in Asia*, *6*(1), 1-21

Pantić, N., & Wubbels, T. (2010). Teacher competencies as a basis for teacher education–Views of Serbian teachers and teacher educators. *Teaching and Teacher Education*, *26*(3), 694-703.

Papa, F. Jr. & Baxter, I. (2008). Hiring teachers in New York's public schools: Can the principal make a difference? *Leadership and Policy in Schools, 7*(1), 87-117.

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*(2), 236-248.

Parsi, A., & Darling-Hammond, L. (2015). Performance Assessments: How State Policy Can Advance Assessments for 21st Century Learning. White Paper. National Association of State Boards of Education.

Paterson, J. (2006). 'Did you hear the one about…? Humour in the classroom'. *Middle Ground, 10*(2), 43–45.

Patton, M. Q. (2014). *Qualitative research and evaluation methods*. Thousand Oaks, CA: Sage.

Pérez-Granados, L. (2018). Using specific tests to access for initial teacher training. *Journal for Educators, Teachers and Trainers*, *9*(1). 55 – 72.

Pervin L. A. (2002). *The Science of Personality* 2nd Edn New York: Oxford University.

Pettis, J. (1997). Developing our professional competence: Some reflections. *TESL Canada Journal*, *16*(2), 67-71.

Phillipson, R. (1992). *Linguistic imperialism*. Oxford: Oxford University.

Pillay, N. (2014). Teachers' knowledge of legislation and education law specifically and its influence on their practice. Unpublished Doctoral dissertation, University of Pretoria. Retrieved from https://repository.up.ac.za/handle/2263/43224

Pillsbury, P. (2005). Only the best: Hiring outstanding teachers. *Leadership, 35*(2), 36–38.

Polesel, J., Rice, S., & Dulfer, N. (2014). The impact of high-stakes testing on curriculum and pedagogy: A teacher perspective from Australia. *Journal of Education Policy, 29*(5), 640-657.

Polk, J. A. (2006). Traits of effective teachers. *Arts Education Policy Review*, *107*(4), 23-29.

Popham, W. J. (1997). Consequential validity: Right concern-wrong concept. *Educational measurement: Issues and practice*, *16*(2), 9-13.

Posnick-Goodwin, S. (2009). Laughter makes you smarter. *California Educator 13*(4), 16–20.

Qi, L. (2005). Stakeholders' conflicting aims undermine the washback function of a high-stakes test. *Language Testing, 22*(2), 142–173

Quirke, P. (2007). *A coherent approach to faculty appraisal*. In C. Coombe, M. Al-Hamly, P. Davidson, and S. Troudi (Eds.), Evaluating teacher effectiveness in EFL/ESL contexts (pp. 89-106). University of Michigan.

Ralph, E. G., Kesten, C., Lang, H., & Smith, D. (1998). Hiring new teachers: What do school districts look for?. *Journal of Teacher Education*, *49*(1), 47-56.

Ramazani, M. (2014). Mismatches in beliefs between teachers and students, and characteristics of effective English teacher: An Iranian context. *Procedia-Social and Behavioral Sciences*, *98*, 1518-1527.

Ramírez Carpeño, E., & Mekochi, Y. (2015). Initial teacher education in Japan and Spain. A comparative study. *Revista Española de Educación Comparada*, *25*, 101-127.

Rice, J. K. (2008). From highly qualified to high quality: An imperative for policy and research to recast the teacher mold. *Education Finance and Policy*, *3*(2), 151-164.

Richards, J. C. (2017a). Introduction. *RELC Journal*, *48*(1), 3-5.

Richards, J. C. (2017b). Teaching English through English: Proficiency, pedagogy and performance. *RELC Journal*, *48*(1), 7-30.

Richards, J. C. (2011). Exploring teacher competence in language teaching. *LANGUAGE TEACHER, 35*(4), 3.

Richards, H., Conway, C., Roskvist, A., & Harvey, S. (2013). Foreign language teachers' language proficiency and their language teaching practice. *The Language Learning Journal, 41*(2), 231–46.

Richards, J. C. (2007). *Curriculum development in language teaching*. New York: Cambridge University.

Richards, K. (2003). *Qualitative Inquiry in TESOL*. Houndsmill, UK: Palgrave, MAcmillan.

Richards, L. (2002). Qualitative computing--a methods revolution?. *International Journal of Social Research Methodology*, *5*(3), 263-276.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2001). *Teachers, schools, and academic achievement*. Working Paper No. 6691. Washington, DC: National Bureau of Economic Research.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, *73*(2), 417-458.

Robins, K. (1996). Interrupting Identities: Turkey/Europe. *Questions of cultural identity*, 61-86.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, *94*(2), 247-252.

Roellke, C., and Rice, J. K. (2008). Responding to teacher quality and accountability mandates: The perspective of school administrators and classroom teachers. *Leadership and Policy in Schools*, *7*(3), 264-295.

Rose, D., English, A., & Finney, T. (2014). *Hire better teachers now: using the science of selection to find the best teachers for your school*. Cambridge, MA. Harvard Press.

Rutledge, S. A., Harris, D. N., Thompson, C. T., & Ingle, W. K. (2008). Certify, blink, hire: An examination of the process and tools of teacher screening and selection. *Leadership and Policy in Schools*, *7*(3), 237-263.

Rowan, B., Correnti, R., & Miller, R. J. (2002). What large-scale survey research tells us about teacher effects on student achievement: Insights from the Prospects study of elementary schools. *Teachers College Record, 104*(8), 1525-1567.

Rubin, H. J., & Rubin, I. S. (2005). *Qualitative interviewing: The art of hearing data (2nd Ed.)*. Thousand Oaks, CA: Sage.

Rubin, H. J., & Rubin, I. S. (2012). *Qualitative interviewing: The art of hearing data.* Los Angeles: SAGE.

Ryan, D. E. (2014). Consider the candidate: Using test-taker feedback to enhance quality and validity in language testing. *e-Teals: an e-journal of Teacher Education and Applied Language Studies*, (*5*), 1-23.

Ryan, F., Coughlan, M., & Cronin, P. (2009). Interviewing in qualitative research: The one-to-one interview. *International Journal of Therapy and Rehabilitation*, *16*(6), 309-314.

Safari, P. (2016). Reconsideration of language assessment is a MUST for democratic testing in the educational system of Iran. *Interchange, 47*(3), 267-296.

Sakurai, Y. (2012). Learners' Perceptions of" Good" Foreign Language Teachers: A Quantitative Analysis between Native and Non-Native Teachers. *Electronic Journal of Foreign Language Teaching*, *9*(1), 46-60

Sanders, W. L., Wright, S. P., & Horn, S. P. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education, 11*(1), 57-67.

Sands, M. K., & Stevens, D. D. (2004). Teacher education in Turkey: Issues and trends. In *Research on Education in Africa, the Caribbean, and the Middle East*, edited by K. Mutua and C. S. Sunal, 267–281. Greenwich: Information Age Publishing.

Santelices, M. V., Valencia, E., Gonzalez, J., & Taut, S. (2017). Two teacher quality measures and the role of context: evidence from Chile. *Educational Assessment, Evaluation and Accountability*, *29*(2), 111-146.

Sautelle, E., Bowles, T., Hattie, J., & Arifin, D. N. (2015). Personality, resilience, self-regulation and cognitive ability relevant to teacher selection. *Australian Journal of Teacher Education*, *40*(4), 4.

Sayın, B. A., & Aslan, M. M. 2016. The negative effects of undergraduate placement examination of English (LYS-5) on ELT students in Turkey. *Participatory Educational Research, 3*(1), 30-39.

Scheurich, J. J., & Skrla, L. (2001). Continuing the conversation on equity and accountability: Listening appreciatively, responding responsibly. *Phi Delta Kappan, 811*(4), 322-326.

Schussler, D. L., Stooksberry, L. M., & Bercaw, L. A. (2010). Understanding teacher candidate dispositions: Reflecting to build self-awareness. *Journal of Teacher Education*, *61*(4), 350-363.

Sebüktekin, H. (1981). Yüksek Öğretim Kurumlarımızda Yabancı Dil İzlenceleri. (Foreign language curricula in institutes of higher education). Istanbul: Boğaziçi University.

Seferoğlu, G. (2006). Teacher candidates' reflections on some components of a pre-service English teacher education programme in Turkey. *Journal of education for Teaching*, *32*(4), 369-378.

Sevimli, S. (2007). The washback effect of foreign language foreign language component of the university entrance examination on the teaching and learning context of English language groups in secondary education (A case study). Unpublished MA Thesis, Gaziantep University.

Sezgin, F., & Duran, E. (2011). Kamu personeli seçme sınavı'nın (KPSS) öğretmen adaylarının akademik ve sosyal yaşantılarına yansımaları. Paper presented at 5. Ulusal Sosyal Bilimler Kongresi

Shih, C. M. (2008). Critical language testing: A case study of the General English Proficiency Test. 英語教學期刊 *(English Teaching Journal) 32*(3), 1-34.

Shiyab, S. M. (2008). Humour as a teaching strategy, in Mantero, M., Miller, P. C. & Wateke, J. L. (eds.), *Reading in Language Studies: Language across Disciplinary Boundaries* 1, 613–626.

Shohamy, E. (1996). Language testing: Matching assessment procedures with language knowledge. In *Alternatives in assessment of achievements, learning processes and prior knowledge* (pp. 143-160). Springer Netherlands.

Shohamy, E. (2001a). Democratic assessment as an alternative. *Language testing*, *18*(4), 373-391.

Shohamy, E. (2001b). *The power of tests: a critical perspective on the uses of language tests.* London: Longman.

Shohamy, E (2006) *Language policy: Hidden agendas and new approaches.* London: Routledge.

Shohamy E. (2017) Critical language testing. In: Shohamy E., Or I., May S. (eds) *Language Testing and Assessment. Encyclopedia of Language and Education (3rd ed.).* Springer, Cham

Smith, K. (2005). New methods and perspectives on teacher evaluation: Who evaluate what and for which purpose? In D. Beijaard, P.C. Meijer, G. Morine-Dershimer,& H. Tillema (Eds.), *Teacher professional development in changing conditions* (pp. 95-114). Netherlands: Springer.

Smith, J. A., Flowers, P., & Larkin, M. (2009). *Interpretative Phenomenological Analysis: Theory, Method and Research*. Sage Publications, London.

Solak, E. (2016). English teacher training programs in Denmark, Sweden and Turkey. *Procedia-Social and Behavioral Sciences*, *232*, 439-443.

Staiger, D. O., & Kane, T. J. (2015). Making Decisions with imprecise performance measures. In T. Kane, K. Kerr, R. Pianta, & Bill and Melinda Gates Foundation (Eds.), *Designing Teacher Evaluation Systems* (pp. 144-169). San Francisco, CA: Jossey-Bass.

Staiger D. O. and Rockoff J. E. (2010). Searching for effective teachers with imperfect information. *Journal of Economic Perspectives 24*(3): 97–117.

Stiell, B., Willis, B., Culliney, M., & Coldwell, M. (2018). International teacher recruitment: understanding the attitudes and experiences of school leaders and teachers. Research Report.

Stiggins, R. J., & Duke, D. L. (1988). *The case for commitment to teacher growth: Research on teacher evaluation*. Albany, NY: State University of New York.

Strauss, A., & Corbin, J. M. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Sage.

Stronge, J. H. (2010). *Teacher effectiveness = student achievement: What the research says*. Larchmont, NY: Eye On Education.

Stronge, J. H. (2007). *Qualities of effective teachers (2nd ed.).* Alexandria: Association for Supervision and Curriculum Development.

Stronge, J. H., & Hindman, J. (2006). *The teacher quality index: A protocol for teacher selection* (Association for Supervision and Curriculum Development Alexandria, Virginia, USA).

Stronge, J.H., Ward, T.J., Tucker, P.D., & Hindman, J.L. (2007). What is the relationship between teacher quality and student achievement? An exploratory study. *Journal of Personnel Evaluation in Education, 20*(3-4), 165-184.

Strübing, J. (2002). Just do it? *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, *54*(2), 318-342.

Şallı-Çopur, D. (2008). Teacher effectiveness in initial years of service: A case study on the graduates of METU foreign language education program. *Unpublished PhD Thesis). Middle East Technical University, Turkey*.

Şimşek, H., & Yıldırım, A. (2001). The reform of pre-service teacher education in Turkey. In *Challenge and Change in the Euro-Mediterranean Region*, edited by R. G. Sultana, 411–430. New York: Peter Lang.

Şişman, M. (2009). Teacher's competencies: A modern discourse and the rhetoric. Invited Paper ISSN, 1300, 2899.

Tabachnick, B. G. (2013). *Using multivariate statistics*. Northridge: Pearson Taylor, T. L. (2002). Living digitally: Embodiment in virtual worlds. In R. Schroeder (Ed.), *The social life of avatars: Presence and interaction in shared virtual environments* (pp. 40-62). London: Springer-Verlag.

Tabachnick, B.G., & Fidell, L.S. (2013). *Using Multivariate Statistics* (6th edt). Pearson, Boston

Thaman, K. H. (2007). Partnerships for progressing cultural democracy in teacher education in Pacific Island countries. In *Handbook of teacher education* (pp. 53-66). Springer, Dordrecht.

Taner, G. (2017). *Investigating perceived competencies of English language teachers in Turkey with regard to educational background and experience.* Unpublished Doctoral dissertation, Middle East Technical University.

Tashakkori, A., & Teddlie, C. (2003). *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage.

Tashakkori, A. & Teddlie, C. (Eds.). (2010). Sage handbook of mixed methods *in social & behavioral research* (2nd ed.). CA: SAGE Publications, Inc.

Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Thousand Oaks, CA: Sage.

Telli, S., Den Brok, P., & Çakiroglu, J. (2008). Teachers' and students' perceptions of the ideal teacher. *Egitim ve Bilim*, *33*(149), 118.

TESOL International Association (2010). *TESOL P–12 professional teaching standards* (2nd ed.). Alexandria, VA: Author.

Toch, T., & Rothman, R. (2008). Rush to Judgment: Teacher Evaluation in Public Education. Education Sector Reports. *Education Sector*.

Toksoz, I., & Kilickaya, F. (2017). Review of journal articles on washback in language testing in Turkey (2010-2017). *Lublin Studies in Modern Languages and Literature, 41*(2), 184-204.

Tongco, M. D. C. (2007). Purposive sampling as a tool for informant selection. *Ethnobotany Research and applications*, *5*, 147-158.

Tongliemnak, P. (2010). *Three essays on teacher labor markets in Thailand*. Unpublished PhD Thesis, Stanford University. Retrieved from

https://stacks.stanford.edu/file/druid:yw658sz0538/Pumsaran_Dissertation_Aug201 0_title_b-augmented.pdf

Tovar, D. (2011). *Teacher evaluation: Foreign-language teachers' and administrators' perspectives on effective evaluation strategies and processes*. Unpublished Doctoral dissertation. The University of San Francisco. Retrieved from https://repository.usfca.edu/diss/16/

Trimble, S. (2001). The Teacher Applicant Pool: What Top Administrators Seek. *Principal Leadership*, *1*(7), 44-47.

Trochim, W. M., & Donnelly, J. P. (2007). *Research methods knowledge base*. Mason, OH: Thomson Custom.

Tsagari, D. (2011). Washback of a high-stakes English exam on teachers' perceptions and practices. *Selected papers on theoretical and applied linguistics*, *19*, 431-445.

Tsang, A. (2017). EFL/ESL teachers' general language proficiency and learners' engagement. *RELC Journal, 48 (1),* 99- 113

Tsui, A.B.M. (2003). *Understanding Expertise in Teaching: Case Studies of ESL Teachers*. Cambridge: Cambridge University.

Tucker, P., & Stronge, J. (2005). *Linking teacher evaluation and student learning*. Alexandria, VA: Association for Supervision and Curriculum Development.

Uyulgan, M. A., & Akkuzu, N. (2015). Öğretmen adaylarının gözünden nitelikli öğretmen seçimi: ortaöğretim fen ve matematik alanları eğitimi bölümünde okuyan öğrencilerin durumu. *KÜ Kastamonu Eğitim Dergisi*, *23*(3), 917-940.

Wagner, T. (2003). School reform that works: Reinventing America's schools. *Phi Della Kappan, 811*(9), 665-668.

Walker, R. J. (2008). Twelve characteristics of an effective teacher: A longitudinal, qualitative, quasi-research study of in-service and pre-service teachers' opinions. *Educational Horizons 87* (1), pp. 61–68.

Wall, D. (1997). Impact and washback in language testing. *Encyclopedia of language and education*, *7*, 291-302.

Wall, D. & Horak, T. (2011). *The impact of changes in the TOEFL examination on teaching in a sample of countries in Europe. Phase 3. The Role of the Coursebook. Phase 4. Describing Change*. Manuscript, Lancaster University.

Walsh, K. (2002). Positive spin: the evidence for traditional teacher certification, reexamined. (Check the Facts). *Education Next*, *2*(1), 79-85.

Walsh, I., Holton, J. A., Bailyn, L., Fernandez, W., Levina, N., & Glaser, B. (2015). What grounded theory is… a critically reflective conversation among scholars. *Organizational Research Methods, 18*(4), 581-599.

Wang, A., Coleman, A., Coley, R., & Phelps, R. (2003). *Preparing teachers around the world*. Princeton, NJ: ETS.

Weir, C.J. (2005). *Language resting and validation*. New York: Palgrave Macmillan.

Welsh, E. (2002, May). Dealing with data: Using NVivo in the qualitative data analysis process. In *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research 3*(2).

Wichadee, S. (2010). Defining the effective English language teacher: Students' and teachers' perspectives, in A.M. Stoke (ed.), *JALT 2009 Conference Proceedings*. Tokyo: JALT.

Wiles, R., Crow, G., Heath, S., & Charles, V. (2008). The management of confidentiality and anonymity in social research. *International Journal of Social Research Methodology*, *11*(5), 417-428.

Van Canh, L., & Renandya, W. A. (2017). Teachers' English proficiency and classroom language use: A conversation analysis study. *RELC Journal*, *48*(1), 67-81.

Van Der Schaaf, M. F., Stokking, K. M., & Verloop, N. (2003). Developing performance standards for teacher assessment by policy capturing. *Assessment & Evaluation in higher education*, *28*(4), 395-410.

VanPatten, B., & Williams, J. (2007). *Theories in Second Language Acquisition: An Introduction*. Mahwah, NJ: Erlbaum

Yalçın-İncik, E. & Akay, C. (2015). Eğitim fakültesi ve pedagojik formasyon sertifika programlarında öğrenim gören öğretmen adaylarının öğretmenlik mesleği yeterliklerine yönelik görüşleri. *Journal of Kirsehir Education Faculty*. *16* (2), 179-197.

Yanow, D. (2007). Qualitative-interpretive methods in policy research. In F. Fischer, G. Miller & M. S. Sidney (Eds.), *Handbook of public policy analysis: theory, politics, and methods* (pp. 405-416). Boca Raton, FL: CRC/Taylor & Francis.

Yavuz, A., & Zehir Topkaya, E. (2013). Teacher educators' evaluation of the English language teaching program: A Turkish Case. *Novitas-royal (research on youth and language)*, *7*(1), 64-83.

Yeşilçınar, S., & Çakır, A. (2018). Continuing professional development process: English language teachers' views on the role of teacher research. *Journal of Language and Linguistic Studies*, *14*(2), 61-75.

Yıldırım, Ö. 2010. Washback effects of a high-stakes university entrance exam: Effects of the English section of the university entrance exam on future language teachers in Turkey. *The Asian EFL Journal Quarterly, 12*(2), 92-116.

Yıldırım, A. & Ok, A. (2002). Alternative teacher certification in Turkey: Problems and issues. In R.G. Sultana (Ed.), *Teacher education in the Mediterranean region: Responding to the challenges of society in transition* (pp. 259-276). USA: Peter Lang.

Yilmaz, M., & YAŞAR, Z. (2016). Öğretmen Adaylarının Kamu Personeli Seçme Sınavına (KPSS) Yönelik Algıları. *Bartin Üniversitesi Eğitim Fakültesi Dergisi*, *5*(2), 644-651.

YÖK, (1998). *Egitim fakültesi ögretmen yetistirme lisans programları*. Ankara: Yüksek Ögretim Kurulu Baskanlıgı.

YÖK, (1999). *Türk yükseköğretiminin bugünkü durumu: YÖK raporları*. Retrieved from http://www.yok.gov.tr/egitim/raporlar/raporlar.htm

YÖK, (2000). Minimum Requirements for Applying to Become Associate Professor. Retrieved from http://www.yok.gov.tr/uakpdfs/tablo1.pdf

YÖK, (2004). *Türk yükseögretiminin bugünkü durumu: YÖK raporları*. Retrieved from http://www.yok.gov.tr/egitim/raporlar/ raporlar.htm

YÖK, (2007). *Teacher education and education faculties*. Ankara: YÖK Publications.

YÖK, (2010). *Student placement scores for undergraduate degree programs*. Retrieved from http:// www.yok.gov.tr/content/view/795/194/lang,tr/.

YÖK, (2018a). Yükseköğretim bilgi yönetim sistemi. Retrieved from https://istatistik.yok.gov.tr/

YÖK, (2018b). İngilizce öğretmenliği lisans programı (English language teaching undergraduate program). Retrieved from http://www.yok.gov.tr/ogretmenlik_lisans_yeni_duzenleme

Zhang, L. F. (2007). Do personality traits make a difference in teaching styles among Chinese high school teachers?. *Personality and Individual Differences*, *43*(4), 669-679.

YÖK, (2018c). Yükseköğretim kurumlari sinavi hakkinda genel bilgi. Retrieved from http://www.yok.gov.tr/documents/10279/36772609/Yuksekogretim_Kurumlari_Sinavi_Hakk%C4%B1nda_Genel_Bilgi.pdf

APPENDICES

APPENDIX A: Average Scores of Reading Skills of Countries and Economies Participating in PISA 2015

	Ortalama puanı istatistiksel olarak anlamlı şekilde OECD ortalamasının üzerinde olan ülkeler
	Ortalama puanı istatistiksel olarak anlamlı şekilde OECD ortalamasından farklı olmayan ülkeler
	Ortalama puanı istatistiksel olarak anlamlı şekilde OECD ortalamasının altında olan ülkeler

Ortalama Puan	Ülke	Ortalama Puan	Ülke
535	Singapur	479	İsrail
527	Hong Kong (Çin)	472	Litvanya
527	Kanada	470	Macaristan
526	Finlandiya	467	Yunanistan
521	İrlanda	459	Şili
519	Estonya	453	Slovakya
517	Güney Kore	447	Malta
516	Japonya	443	Güney Kıbrıs
513	Norveç	437	Uruguay
509	Yeni Zelanda	434	Romanya
509	Almanya	434	Birleşik Arap Emirlikleri
509	Makao (Çin)	432	Bulgaristan
506	Polonya	431	Malezya
505	Slovenya	*428*	*Türkiye*
503	Hollanda	427	Kosta Rika
503	Avustralya	427	Trinidad ve Tobago
500	İsveç	427	Kazakistan
500	Danimarka	427	Karadağ
499	Fransa	425	Arjantin
499	Belçika	425	Kolombiya
498	Portekiz	423	Meksika
498	Birleşik Krallık	416	Moldova
497	Tayvan - Çin	409	Tayland
497	Amerika	408	Ürdün
496	İspanya	407	Brezilya
495	Rusya Federasyonu	405	Arnavutluk
494	Çin Halk Cumhuriyeti	402	Katar
492	İsviçre	401	Gürcistan
488	Letonya	398	Peru
487	Çek Cumhuriyeti	397	Endonezya
487	Hırvatistan	361	Tunus
487	Vietnam	358	Dominik Cumhuriyeti
485	Avusturya	352	Makedonya
485	İtalya	350	Cezayir
482	İzlanda	347	Kosova
481	Lüksemburg	347	Lübnan

Taken from Taş, Arıcı, Ozarkan and Özgürlük (2016).

APPENDIX B: Semi-Structured Interviews Questions with the Student Teachers

I-   Demographic Questions

Why did you choose this department?

What do you want to do after graduating from this department?

Are you joining any course for TAE?

How long have you been studying for TAE?

How do you study for ÖABT?

II-   Main Questions

1) How does the current TAE model affect your life? (both IN and OUTSIDE the faculty)

2) Does preparing for TAE have any contribution to you in terms of the teaching profession?

3) Is KPSS together with its components (e.g. ÖABT, Interview) a must? Why? Why not?

4) What do you think about the components of the current TAE model?
   a) Which component(s) is/are a must?
   b) Which component(s) is/are unnecessary one(s)?
   c) Which component(s) should be added?

5) Do you think the nature of the current TAE model is parallel with the courses you have taken?

6) Do you ask teacher trainers for giving lessons according to TAE model?

7) In your opinion, how should an appropriate TAE model be?

8) Do you think that the current TAE model include all competencies that an effective EFL teacher should have? If not, what can be done to enhance it?

9) Do you think that the EFL teachers who are chosen by the current TAE model can teach both primary and secondary schools?

10) In your opinion, what is sine qua non of being an EFL teachers?

11) What do you think about EFL teachers who are not graduates of ELT programs and, therefore, must join a pedagogical formation programs (e.g. graduates of English language and literature)?

12) In your opinion, is there a relationship between how teachers are evaluated and how they teach English to their students?

13) Do you think that skills and attitudes should be assessed (for example, their performance, technological literacy, motivational dispositions)? If yes, how?

APPENDIX C: Semi-Structured Interviews Questions with the In-Service Teachers

I-   Demographic Questions

How long have you been teaching?

Which grades do you teach?

Which university and department did you graduate from?

Why did you choose that department?

What was your goal after graduating from this department? (to be a teacher or…?)

How long did you study for TAE?

How did you study for TAE?

II-   Main Questions

1) How did the TAE model affect your life? (both IN and OUTSIDE the faculty)

2) Did preparing for TAE have any contribution to you in terms of the teaching profession?

3) Is KPSS together with its components (e.g. ÖABT, Interview) a must? Why? Why not?

4) What do you think about the components of the current TAE model?
   a) Which component(s) is/are a must?
   b) Which component(s) is/are unnecessary one(s)?
   c) Which component(s) should be added?

5) Do you think the nature of the current TAE model was parallel with the courses you had taken?

6) Did you ask teacher trainers for giving lessons according to TAE model?

7) In your opinion, how should an appropriate TAE model be?

8) Do you think that the current TAE model include all competencies that an effective EFL teacher should have? If not, what can be done to enhance it?

9) Do you think that the EFL teachers who are chosen by the current TAE model can teach both primary and secondary schools?

10) In your opinion, what are sine qua nons of being an EFL teachers?

11) What do you think about EFL teachers who are not graduates of ELT programs and, therefore, must join a pedagogical formation programs (e.g. graduates of English language and literature)?

12) In your opinion, is there a relationship between how teachers are evaluated and how they teach English to their students?

13) Do you think that skills and attitudes should be assessed (for example, their performance, technological literacy, motivational dispositions)? If yes, how?

APPENDIX D: Semi-Structured Interviews Questions with the Teacher Trainers

1) Would you please share your opinions about the current TAE model?

2) Does the introduction of ÖABT and Interview change your students' attitudes towards the courses?

3) Does preparing for TAE model contribute student teachers in terms of the teaching profession?

4) Is KPSS together with its components (e.g. ÖABT, Interview) a must? Why? Why not?

5) What do you think about the components of the current TAE model?
   a) Which component(s) is/are a must?
   b) Which component(s) is/are unnecessary one(s)?
   c) Which component(s) should be added?

6) Do you think the nature of TAE model should be parallel with the courses that student teachers have taken?

7) Do you revise your syllabus with the aim of helping student teachers prepare for TAE?

8) In your opinion, how should an appropriate TAE model be?

9) Do you think that the current TAE model include all competencies that an effective EFL teacher should have? If not, what can be done to enhance it?

10) Do you think that the EFL teachers who are chosen by the current TAE model can teach both primary and secondary schools?

11) In your opinion, what is sine qua non of being an EFL teachers?

12) What do you think about EFL teachers who are not graduates of ELT programs and, therefore, must join a pedagogical formation programs (e.g. graduates of English language and literature)?

13) In your opinion, is there a relationship between how teachers are evaluated and how they teach English to their students?

14) Do you think that skills and attitudes should be assessed (for example, their performance, technological literacy, motivational dispositions)? If yes, how?

APPENDIX E: Semi-Structured Interviews Questions with the Education Experts

1) Would you please share your opinions about the current TAE model?

2) Does the introduction of ÖABT and Interview change students' attitudes towards the courses?

3) Does preparing for TAE model contribute student teachers in terms of the teaching profession?

4) Is KPSS together with its components (e.g. ÖABT, Interview) a must? Why? Why not?

5) What do you think about the components of the current TAE model?
   a) Which component(s) is/are a must?
   b) Which component(s) is/are unnecessary one(s)?
   c) Which component(s) should be added?

6) Do you think the nature of TAE model should be parallel with the courses that student teachers have taken?

7) Do you revise your syllabus with the aim of helping student teachers prepare for TAE?

8) In your opinion, how should an appropriate TAE model be?

9) Do you think that the current TAE model include all competencies that an effective EFL teacher should have? If not, what can be done to enhance it?

10) Do you think that the EFL teachers who are chosen by the current TAE model can teach both primary and secondary schools?

11) In your opinion, what are sine qua nons of being an EFL teachers?

12) What do you think about EFL teachers who are not graduates of ELT programs and, therefore, must join a pedagogical formation programs (e.g. graduates of English language and literature)?

13) In your opinion, is there a relationship between how teachers are evaluated and how they teach English to their students?

14) Do you think that skills and attitudes should be assessed (for example, their performance, technological literacy, motivational dispositions)? If yes, how?

APPENDIX F. Semi-Structured Interviews Questions with the Policymakers

1) What kind of policies do you follow when revising the TAE model?

2) Is KPSS together with its components (e.g. ÖABT, Interview) a must? Why? Why not?

3) What do you think about the components of the current TAE model?

    a)   Which component(s) is/are a must?
    b)   Which component(s) is/are unnecessary one(s)?
    c)   Which component(s) should be added?

4) In your opinion, how should an appropriate TAE model be?

5) Do you think that the current TAE model include all competencies that an effective EFL teacher should have? If not, what can be done to enhance it?

6) In your opinion, what are sine qua nons of being an EFL teachers?

7) What do you think about EFL teachers who are not graduates of ELT programs and, therefore, must join a pedagogical formation programs (e.g. graduates of English language and literature)?

8) In your opinion, is there a relationship between how teachers are evaluated and how they teach English to their students?

9) Do you think that skills and attitudes should be assessed (for example, their performance, technological literacy, motivational dispositions)? If yes, how?

APPENDIX G: İngilizce Öğretmeni Seçme ve Atama Modeline Yönelik Tutum Ölçeği

Bu araştırma İngilizce öğretmenlerinin atama sistemine yönelik tutumlarının saptanması ve geleceğe yön veren öğretmenlerimizin daha adil bir şekilde seçilip atanmasına yardımcı olabilecek bir **atama modeli** önerisi açısından önemlidir. Maddeleri sadece kendi duygu ve düşüncelerinizi yansıtacak şekilde içtenlikle cevaplandırmanız çok önemlidir. Bu nedenle, lütfen sorulara neyin doğru neyin yanlış olduğuna göre değil, gerçekte sizin ne yaşadığınıza ve ne düşündüğünüze göre cevap veriniz. Size en yakın gelebilecek seçeneğin olduğu ilgili kutucuğa herhangi bir işaret bırakmanız (ör., X, √) yeterli olacaktır.

Ölçekle ilgili herhangi bir sorunuz varsa veya sonuçlar hakkında bilgi almak istiyorsanız, lütfen iletişime geçmekten çekinmeyiniz.

Sabahattin YEŞİLÇINAR	Prof. Dr. Abdulvahit ÇAKIR
Gazi Eğitim Fakültesi, İngiliz Dili Eğitimi	Gazi Eğitim Fakültesi, İngiliz Dili Eğitimi
sycinar@gazi.edu.tr	vahit@gazi.edu.tr

Araştırmaya katıldığınız için çok teşekkür ederiz.

		Kesinlikle katılmıyorum	Katılmıyorum	Kararsızım	Katılıyorum	Kesinlikle Katılıyorum
	Cinsiyet .... **Atandığınız Yıl** .... **Ders verdiğiniz sınıflar** .... **Mezun olduğunuz üniversite:**... **Mezun olduğunuz bölüm:**... **Bölümü seçme nedeniniz:** a) sevme b) yönlendirme c)öğretmenimden dolayı d)(diğer)................ **Hayalinizdeki meslek:** a) öğretmen b) akademisyen c) (diğer)................					
1.	Mevcut atama modelinin öğretmenlik yapamayan adayları da seçtiğine inanırım	1	2	3	4	5
2.	Mevcut atama modelinin yeterli alan bilgisine sahip öğretmen adaylarını seçtiğini düşünürüm	1	2	3	4	5
3.	Mevcut atama modelinin değiştirilmesi gerektiğine inanırım	1	2	3	4	5
4.	Mevcut atama modelinin istekli öğretmen adaylarını seçtiğini düşünürüm	1	2	3	4	5
5.	Mevcut atama modelinin kaldırılmasından mutluluk duyarım	1	2	3	4	5
6.	Mevcut atama modelinin motivasyonu yüksek öğretmen adaylarını seçtiğini düşünürüm	1	2	3	4	5
7.	Mevcut atama modelinin paydaşlarla (örn. Öğrenci, meslektaş, veli) iletişim kurabilen öğretmen adaylarını seçtiğini düşünürüm	1	2	3	4	5
8.	Mevcut atama modeliyle ilgili Türkiye'deki gelişmeleri takip ederim	1	2	3	4	5
9.	Mevcut atama modelinin seçme kriterlerinin adaletli düzenlendiğine inanırım	1	2	3	4	5
10.	Mevcut atama modelinin her kademede (örn. İlkokul, ortaokul, lise) İngilizce öğretebilecek öğretmen adaylarını seçebildiğine inanırım	1	2	3	4	5
11.	Mevcut atama modeliyle ilgili dünyadaki gelişmeleri takip ederim	1	2	3	4	5
12.	Mevcut atama modelinin kendini sürekli geliştirebilen öğretmen adaylarını seçtiğini düşünürüm	1	2	3	4	5
13.	Mevcut atama modeliyle ilgili tartışmalar ilgimi çeker	1	2	3	4	5
14.	Mevcut atama modelinin değişmesinden mutluluk duyarım	1	2	3	4	5
15.	Mevcut atama modelinin yeterli teknolojik bilgiye sahip öğretmen adaylarını seçtiğini düşünürüm	1	2	3	4	5
16.	Mevcut atama modeliyle ilgili yazılı kaynakları (örn., kitap, dergi, haber) okurum	1	2	3	4	5
17.	Mevcut atama modelinin adil olmadığı kanısındayım	1	2	3	4	5
18.	Mevcut atama modelinin yeterli mesleki bilgiye sahip öğretmen adaylarını seçtiğini düşünürüm	1	2	3	4	5
19.	Mevcut atama modeliyle ilgili elektronik kaynakları (örn., blog, paylaşım siteleri) takip ederim	1	2	3	4	5
20.	Mevcut atama modeliyle ilgili bilgimi artırmak için meslektaşlarımla tartışırım	1	2	3	4	5
21.	Mevcut atama modelinin öğrencilerin ihtiyacını göz önünde bulundurarak İngilizce öğretmen adayı seçtiğini düşünürüm	1	2	3	4	5
22.	Mevcut atama modeliyle ilgili her şey dikkatimi çeker	1	2	3	4	5
23.	Boş zamanlarımda mevcut atama modeliyle ilgili bilgiler edinmeye çalışırım	1	2	3	4	5

APPENDIX H: English Teachers' Attitudes towards Recruitment System (ETARS)

This research is important in terms of identifying the attitudes of English language teachers towards the way English language teachers assessed and evaluated in Turkey, and suggesting a fair **teacher assessment and evaluation (TAE) model** for future teachers. It is very important that you respond the items sincerely so it reflects your own feelings and thoughts. For this reason, please answer the questions according to your actual experiences and thoughts rather than what is right or wrong. Please, leave a mark (e.g., X, √) in the relevant box.

Please do not hesitate to contact the researchers if you have any question or if you want to be aware of the results.

Sabahattin YEŞİLÇINAR
Gazi Faculty of Education,
English Language Teaching
sycinar@gazi.edu.tr

Prof. Dr. Abdulvahit ÇAKIR
Gazi Faculty of Education,
English Language Teaching
vahit@gazi.edu.tr

Thanks for your participation.

Gender ………. Recruitment year………… Grades taught……………
University that you graduated from…………………………………………
Department that you studied…………………………………………………
Reason behind choosing the department? a) love b) guidance
  c) because of English teacher d) (other)……………
Your desired occupation: a) teacher b) academician c) (other)………

		Strongly disagree	Disagree	Neutral	Agree	Strongly agree
1.	I believe that the current TAE model selects also the student teachers who cannot teach.	1	2	3	4	5
2.	I think that the current TAE model selects student teachers with required field knowledge.	1	2	3	4	5
3.	I believe that the current TAE model needs to be changed.	1	2	3	4	5
4.	I think that the current TAE model selects student teachers who love teaching profession	1	2	3	4	5
5.	I become happy if the current TAE model is removed.	1	2	3	4	5
6.	I think that the current TAE model selects highly motivated student teachers	1	2	3	4	5
7.	I think that the current TAE model selects student teachers who can communicate well with the stakeholders (such as student, colleague, parent)	1	2	3	4	5
8.	I follow the recent developments in Turkey related to the current TAE model	1	2	3	4	5
9.	I believe that the selection criteria of the current TAE model is fair.	1	2	3	4	5
10.	I believe that the current TAE model can select student teachers who can teach English at every level (e.g. primary school, junior high school, high school)	1	2	3	4	5
11.	I follow the recent developments in the world related to the current TAE model	1	2	3	4	5
12.	I think that the current TAE model selects student teachers who can constantly improve themselves	1	2	3	4	5
13.	I am interested in discussions related to the current TAE model.	1	2	3	4	5
14.	I become happy if the current TAE model is changed.	1	2	3	4	5
15.	I think that the current TAE model selects student teachers with required technological knowledge.	1	2	3	4	5
16.	I read written sources (e.g. books, magazines, news) about the current TAE model.	1	2	3	4	5
17.	I think that the current TAE model is unfair.	1	2	3	4	5
18.	I think that the current TAE model selects student teachers with required professional knowledge.	1	2	3	4	5
19.	I follow electronic sources (e.g. blogs, forums) about the current TAE model.	1	2	3	4	5
20.	I discuss with colleagues to increase my knowledge of the current TAE model.	1	2	3	4	5
21.	I think that the current TAE model selects English student teachers according to the needs of the students.	1	2	3	4	5
22.	Everything about the current TAE model attracts my attention.	1	2	3	4	5
23.	I try to get information about the current TAE model in my spare time.	1	2	3	4	5

APPENDIX I. Participant Consent Form

I hereby confirm that I take part in this research study voluntarily. I was assured that only the researcher (Sabahattin YEŞİLÇINAR) will have access to the information I will provide without revealing my name and personal information.

Title of the Study: AN EVALUATION OF EFL TEACHER ASSESSMENT AND EVALUATION: A SUGGESTED MODEL

The purpose of this research study: This study aims to develop a teacher selection and recruitment model based on the views of student teachers, teachers, teacher trainers, and policy makers.

Ethical considerations and confidentiality: There is no risk involved in this study. All the information provided by you will remain confidential. Nobody except the researcher will have access to any of the data. The results of the study may be published in academic journals and elsewhere without giving your identity.

Available Sources of Information

If you have any further inquiries, you may contact me from my mobile phone- 0532 509 30 16- or through my e-mail at sycinar@yahoo.com

Authorization

I have read and understood this consent form, and I volunteer to participate in this research study. I am ensured that in case of my request to be excluded from the study during the course of research, informing the researcher Sabahattin YEŞİLÇINAR will be sufficient.

Participant's Name:

Date:

Signature:

APPENDIX J. Permission of Gazi University Ethics Commission

T.C.
GAZİ ÜNİVERSİTESİ
Etik Komisyonu

90 yıl
GAZİ
ÜNİVERSİTESİ

* B E N U K K C C L *

Sayı : 77082166-302.08.01-
Konu : Bilimsel ve Eğitim Amaçlı

EĞİTİM BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜNE

İlgi : 13/10/2017 tarihli ve 80287700-302.08.01- 145341 sayılı yazı.

İlgi yazınız ile göndermiş olduğunuz, Yabancı Diller Eğitim Anabilim Dalı İngilizce Öğretmenliği Bilim Dalı **Doktora Öğrencisi Sabahattin YEŞİLÇINAR'ın, Prof.Dr.Abdulvahit ÇAKIR** danışmanlığında yürüttüğü *"İngilizce Öğretmeni Seçme ve Atanmasının Değerlendirilmesi: Bir Model Önerisi"* adlı tez çalışması ile ilgili konu Komisyonumuzun **19.12.2017** tarih ve **10** sayılı toplantısında görüşülmüş olup,

İlgilinin çalışmasının, yapılması planlanan yerlerden izin alınması koşuluyla yapılmasında etik açıdan bir sakınca bulunmadığına oybirliği ile karar verilmiş ve karara ilişkin imza listesi ekte gönderilmiştir.

Bilgilerinizi ve gereğini rica ederim.

e-imzalıdır
Prof. Dr. Alper CEYLAN
Komisyon Başkanı

Araştırma Kod No: 2017-499

Ek:1 Liste

Bu belge 5070 sayılı Elektronik İmza Kanununun 5. Maddesi gereğince güvenli elektronik imza ile imzalanmıştır.

APPENDIX K. Permission of the Ministry of National Education

T.C.
ANKARA VALİLİĞİ
Milli Eğitim Müdürlüğü

Sayı : 14588481-605.99-E.9272174   10.05.2018
Konu : Araştırma İzni

GAZİ ÜNİVERSİTESİNE
(Eğitim Bilimleri Enstitüsü)

İlgi: a) MEB Yenilik ve Eğitim Teknolojileri Genel Müdürlüğünün 2017/25 nolu Genelgesi.
        b) 03/05/2018 Tarihli ve 749 sayılı yazınız.

Enstitünüz, Yabancı Diller Eğitimi Anabilim Dalı, İngilizce Öğretmenliği Bilim Dalı Doktora Programı öğrencisi Sabahattin YEŞİLÇINAR'ın **"The Assessment of EFL Teacher Selection and Recruitment: A Suggested Model"** konulu tez çalışması kapsamında uygulama talebi Müdürlüğümüzce uygun görülmüş ve uygulamanın yapılacağı İlçe Milli Eğitim Müdürlüğüne bilgi verilmiştir.

Görüşme formunun (8 sayfa) araştırmacı tarafından uygulama yapılacak sayıda çoğaltılması ve çalışmanın bitiminde bir örneğinin (cd ortamında) Müdürlüğümüz Strateji Geliştirme (1) Şubesine gönderilmesini rica ederim.

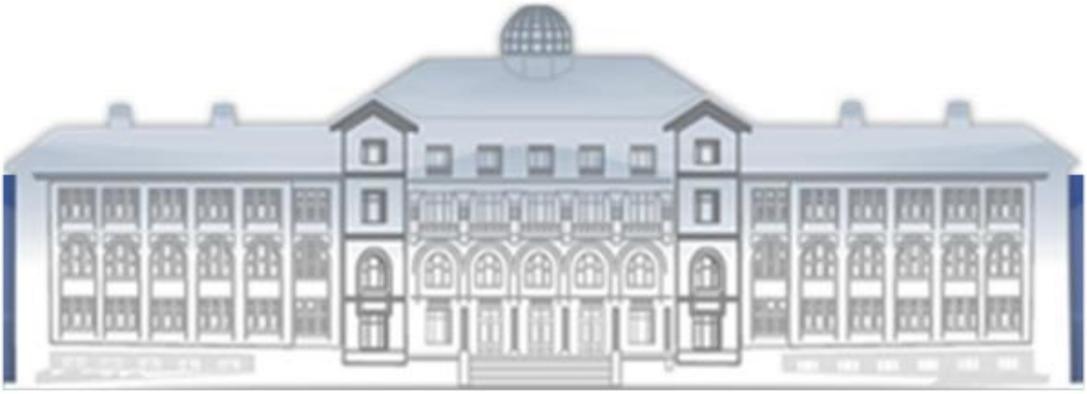Vefa BARDAKCI
Vali a.
Milli Eğitim Müdürü

Güvenli Elektronik İmzalı
Aslı İle Aynıdır
........../201....
11 Mayıs 2018

Adres: Alparslan Türkeş cad. Emniyet Mah.4/A   Bilgi için: A.ARDA
Yenimahalle/ANKARA
Elektronik Ağ: ankara.meb.gov.tr   Tel: 0 (312) 221 02 17
e-posta: istatistik06@meb.gov.tr   Faks: 0 (312) 221 02 16

Bu evrak güvenli elektronik imza ile imzalanmıştır. https://evraksorgu.meb.gov.tr adresinden 6a74-5cad-36ea-af90-ac22 kodu ile teyit edilebilir.

GAZİLİ OLMAK AYRICALIKTIR...